

Искусственная религия. Об ИИ, мифе и власти

Марк Кекельберг

Предисловие: Неучтивые, но неотложные разговоры

Принято считать, что в вежливой беседе следует избегать разговоров о политике и религии: они способны вызвать споры и бурные эмоции. Исследуя некоторые глубокие и сложные точки пересечения искусственного интеллекта, религии и власти, эта книга вторгается именно на такую заряженную территорию. Сегодня ИИ стремительно преобразует ткань наших обществ, и потому эти разговоры не только неизбежны, но и необходимы и неотложны.

В современном дискурсе об ИИ доминируют опасения по поводу будущих технологических достижений, вопросы этики и человеческой субъектности, а также политика и регулирование. Однако, сосредоточившись на технологических, этических и регуляторных аспектах, мы нередко упускаем из виду более глубокие культурные, духовные и экзистенциальные течения, которые влияют на эти процессы и формируют их. Эта книга призвана восполнить данный пробел и вывести на первый план некоторые способы, которыми размышления об ИИ и взаимодействие с ним переплетены с религиозными нарративами и практиками, а тем самым и с властью, и в то же время отвечают нашим экзистенциальным потребностям.

Двигаясь в этом направлении, я продолжаю подход и развиваю идеи, занимавшие центральное место в моих предыдущих работах. В книге «Новые романтические киборги» (*New Romantic Cyborgs*) я рассматривал, как романтические способы мышления до сих пор формируют наше современное использование технологий и наш дискурс о них. В книге *Using Words and Things* и родственных статьях я, опираясь на Витгенштейна, доказывал, что значение технологии подобно значению языка: оно зависит от употребления и от «грамматик», то есть трансцендентальных культурных структур, которые определяют, как мы пользуемся технологиями и как о них думаем. А во многих прежних работах я разделял базовую идею современной философии техники,

вдохновлённой феноменологией и герменевтикой: технология не просто инструмент, она сама формирует нашу жизнь и существование. Эта книга продолжает данную траекторию, но с особым вниманием к религиозным идеям и практикам.

Как всегда, эта работа не существует сама по себе, а неизбежно и, к счастью, связана отношениями с другими разговорами. Пользуясь случаем, хочу поблагодарить людей, с которыми в последние годы у меня были стимулирующие беседы об ИИ и религии: Дэвида Дюзенбери, Одри Боровски, Габора Амбруша, Аманду Лагерквист, Тома Фурланиса, Йоханнеса Хоффа, Сару Шпикерманн, Юхани Штайнмана, Маттиа Джеретто и Еву Дедечкову. Благодарю также Марсело Сорию-Родригеса за любезное разрешение использовать его фотографию произведения *Appropriate Response* и самого художника Марио Клингеманна. Наконец, я благодарю редакторов MIT Press Фила Лафлина и Вирджинию Кроссман за поддержку этого проекта, а также Закари Стормса за помощь в вёрстке рукописи.

Следуя за моим путешествием и очерком на этих страницах, приглашаю вас присоединиться к этим трудным, но необходимым разговорам. В эпоху, когда технологические достижения нередко встречают либо некритичным восторгом, либо паникой и мыслями о конце света, да ещё на фоне поляризованного политического ландшафта, как никогда важно выработать более взвешенный и критический взгляд. Понимание корней нашего мышления об ИИ может в этом помочь. И наконец, эта книга также призыв к активному интеллектуальному и политическому участию в глубоких культурно-технологических переменах, которые разворачиваются на наших глазах, желательно без надежды на *deus ex machina* («бога из машины»), которая преследует и западное светское, и религиозное мышление.



1. Введение: Рождение машин из духа религии

Произведение Марио Клингеманна *Appropriate Response* (2020) состоит из деревянной скамеечки для коленопреклонения и табло с перекидными буквами, которые непрерывно меняются. Когда человек опускается на колени, встроенный искусственный интеллект (ИИ), на тот момент GPT-2, выдаёт уникальную фразу, похожую на афоризм. Инсталляция ставит вопросы об авторстве и смысле: может ли ИИ быть автором? Кто в конечном счёте определяет смысл текста, автор или читатель? [1] Но эта инсталляция намекает и на религиозные темы. Люди жаждут вдохновения и мудрости и нередко обращаются за ними к религии. Они ищут ответы. Они ищут смысл. Они хотят знать будущее. Они надеются. Они боятся. Они молятся. Им нужно за что-то или за кого-то держаться. У религии для этого есть ритуалы. Есть и профессионалы: священники, монахи, религиозные лидеры. Выполняет ли ИИ сегодня ту же функцию, когда мы вводим запрос (молимся?) в поисках ответа, а может быть, и спасения? «С одной стороны, мы боимся ИИ, но в то же время надеемся, что он поможет решить некоторые проблемы, — говорит Клингеманн. — Этот баланс надежды и страха тесно связан с религиозным опытом, поэтому коленопреклонение показалось мне очень подходящим» [2]. Является ли ИИ новым священником или даже новым богом, всемогущим и всеведущим сверхчеловеческим посредником? Надеемся ли мы, что *deus ex machina* нас спасёт? Является ли ИИ в этом смысле проектом искусственной религии в той же мере, что и проектом искусственного интеллекта?

Технологии и религия обычно не оказываются в рамках одной дискуссии, одной публикации или одной академической кафедры. В Новое время мы научились раскладывать их по разным ящикам и, как правило, предпочитаем так и оставлять. Первые относятся к миру природы и науки, вторая к миру культуры и гуманитарных наук. Большинство людей придерживается этого разделения. Но есть и исключения, и в дебатах об ИИ некоторые голоса пересекли границу. В книге *Homo Deus* Юваль Харари утверждает, что у нас появится религия данных, «датаизм»: мы поверим, что человеческий вид есть система обработки данных, и в конце концов полностью передадим свои решения алгоритмам. Ватикан активно участвует в обсуждении этики ИИ, в частности через инициативу «Римский призыв к этике ИИ» (Rome Call for AI Ethics) [3]. Критики лонгтермизма и эффективного альтруизма, движений с очень конкретными представлениями о долгосрочном будущем человечества, ставят под сомнение идеологию Кремниевой долины, которую легко увидеть не только как утопическую, но и как квазирелигиозную [4]. Это одна из причин, по которой Big Tech всё больше становится ещё и Big Religion. По словам Грега Эпштейна, технологии стали «религиозной империей»: вездесущие в нашей жизни, они взяли на себя роль религии в «формировании наших мыслей, чувств, надежд, страхов, отношений и будущего», причём всё это в интересах небольшого числа крупных корпораций [5]. И, как мы

видели, такие художники, как Клингеманн, не испытывают затруднений с пересечением границы между технологией и религией.

Эта книга утверждает, что подобные дискурсы и произведения возможны потому, что никакой строгой границы изначально нет. Когда я использую в книге термин «ИИ», я имею в виду популярные сегодня приложения машинного обучения, например те, что основаны на больших языковых моделях (LLM). Но, как я покажу, ИИ связан не только с алгоритмами, моделями, данными и облачной инфраструктурой, но и с нарративами, мифами, коллективными надеждами и страхами. Часто говорят, что ИИ есть миф в том смысле, что он обещает слишком много и раздувается ради денег и власти, однако связь с мифом глубже, и необходимо копнуть глубже, к лежащим в основе нарративам и тревогам [6]. Такие мыслители, как Карл Юнг, Джозеф Кэмпбелл, Клод Леви-Стросс и Поль Рикёр, указывали на продуктивную роль мифа в формировании человеческой культуры и в порождении смыслов. Мифологические нарративы помогают нам в личном и коллективном самопонимании. Миф не случайное дополнение: вслед за Хансом Блуменбергом мы можем понимать миф как возникающий из нашего экзистенциального, антропологического положения [7]. Он нужен нам, чтобы справляться с нашей экзистенциальной ситуацией. Если это верно, то наше мышление о технологиях и наше отношение к ним тоже формируются и опосредуются мифом [8]. Поэтому, если мы хотим лучше понимать современные технологии и выстроить с ними критическое отношение, нам нужно работать с мифом.

Итак, ИИ есть не только технология и потенциальный источник прибыли, но и вопрос смысла, существования и того, что теолог и философ Пауль Тиллих назвал «пределной заботой» (*ultimate concern*), пытаюсь описать сущность веры: нечто настолько важное, что оно придаёт жизни смысл и направление [9]. Опираясь на историю идей, включая значимые идеи философии и теологии, эта книга предпринимает более систематическую попытку исследовать связи между ИИ и религией. При этом она не столько о том, как ИИ уже принят и используется мировыми религиями, как о нём думают современные теологи и религиозные лидеры или как ИИ и связанные с ним движения становятся религией в более практическом и организованном смысле, сколько о том, как наши способы мыслить об ИИ переплетены с религиозными и мифологическими способами мышления [10]. Книга покажет, что сам проект создания разумных машин по нашему образу и подобию, говоря языком Ницше, рождён из духа религии, и что способы использования этих машин по-прежнему вписываются в религиозные паттерны мышления и практики и поддерживают их [11].

Для этого книга исходит из двух допущений. Первое: религия не умерла. Второе: технология никогда не бывает только о технологии.

Во-первых, религия не умерла. Я имею в виду это по меньшей мере в трёх смыслах. Первый: хотя процессы секуляризации, несомненно, идут, в одних местах (Западная Европа, некоторые городские среды) сильнее, чем в других, религия остаётся значимым

общественным явлением и даже значимой политической силой в большинстве мест на планете. Не существует никакого необходимого пути к секуляризации. Некоторые социологи даже говорят о «десекуляризации», имея в виду, например, возрождение религии в России и Китае, рост евангельского христианства в США и увеличение христианского населения в странах Глобального Юга [12]. Утверждать, что мы живём в современном мире без религии, значит придерживаться весьма западноцентричного взгляда, а и на Западе между странами огромные различия. Например, в США многие люди считают себя религиозными, тогда как в некоторых европейских странах, таких как Чехия, Швеция или Франция, большинство так не считает. Велики и различия внутри стран: как правило, в рамках национальных государств и городов сосуществуют, мирно или нет, несколько религий, мировоззрений и атеистических движений. Например, в крупных западных городских центрах значительную долю составляют религиозные иммигранты; называть такие города светскими было бы ошибкой.

Второй смысл: даже там, где он был задуман, в Европе и особенно в Просвещении во Франции (хотя у него есть предшественники в Древней Греции, в частности досократик Демокрит и софисты), проект секуляризации никогда не был вполне успешным. Даже люди и общества, которые называют себя в основном секуляризованными и отказавшимися от веры в Бога или богов (например, те, кто считает себя атеистами), всё равно склонны участвовать в религиозных практиках (например, в христианских свадьбах и похоронах) или воспроизводить религиозно обусловленные паттерны мышления. Скажем, после того, что Фридрих Ницше назвал смертью Бога, даже те, кто идентифицирует себя как атеист или нерелигиозный человек, нередко продолжали думать о мире и о себе по-христиански [13]. В этой книге я покажу, что наряду с древнегреческими мифами иудейские и христианские паттерны мышления по-прежнему формируют то, как мы думаем об ИИ и как с ним обращаемся.

В истории идей эти замечания перекликаются с известной критикой так называемого тезиса о расколдовывании. Макс Вебер, как известно, утверждал, что современность характеризуется расколдовыванием: с возникновением современной науки и просвещённого разума верят в магию и религию всё меньше; наука, Просвещение и бюрократия лишают мир удивления, тайны и духовного значения, которыми прежде была проникнута человеческая жизнь [14]. Карл Маркс, Зигмунд Фрейд и Эмиль Дюркгейм тоже считали, что существует движение, возможно неизбежное, к секуляризации. Поэтому можно было бы ожидать повсеместного расколдовывания. Однако за последние десятилетия социологи по-разному поставили тезис о расколдовывании под сомнение. Наука тоже порождает удивление [15]. Появилось множество новых форм очарования и религии, например New Age, мировые религии живы и здравствуют в большинстве частей мира, а древние мифы и немодерные формы духовности продолжают оказывать влияние. Как гласит фраза, которую часто приписывают Веберу, нас по-прежнему преследуют призраки умерших верований.

Как мы увидим, эти призраки бродят и по соборам технологии. Я покажу, как невидимые, но мощные паттерны религиозного и мифологического мышления формируют наши представления о технологиях, включая ИИ. В этом смысле религия по меньшей мере «нежива» (*undead*), если не жива и здорова. На протяжении всей книги я использую эту спектральную метафору. Некоторые религиозные паттерны мышления сохраняются и в нашей технокультуре, хотя обычно они не видны; мы, как правило, их не осознаём. Продолжая пользоваться языком Витгенштейна, можно также сказать, что существует глубоко укоренённая религиозная культурная «грамматика», форма жизни, по-прежнему пропитанная религиозными паттернами мышления. Или можно говорить о коллективном «религиозном бессознательном», которое по-прежнему оказывает влияние, в том числе в сфере технологий [16]. Или, иначе, наша тяга к очарованию оказывается упрямее, чем мы думали. ИИ и другие новые технологии можно даже рассматривать как вклад в повторное очарование мира [17]. Как уже отмечал культурный критик Нил Постман, теперь мы верим в технологию [18]. А Иэн Богост называет алгоритмическую культуру «богослужебной» (*devotional*) [19]. Сегодня доминирующая технология, новый бог, это ИИ. Возможно, западная увлечённость технологиями и развитие науки всегда имели в основе религиозные ожидания, как утверждает историк Дэвид Ноубл [20]. Хотя многие из тех, кто разрабатывает, продаёт и продвигает ИИ, могут считать себя рационалистами и светскими людьми, я покажу в этой книге, что ИИ и религиозные идеи переплетены разными способами.

Это связано с формирующимися научными работами в данной области. Роберт Джераси уже изучал, как разные религиозные среды в США и Японии влияют на цели и практику робототехники и ИИ [21]. Как отмечают антропологи и религиоведы, многие религии приняли ИИ, а также появляются новые религиозные движения, ориентированные на технологию. Из этих областей мы можем также узнать, что религия не монолитна, да и вещь ли она вообще. Религия существует во множестве разновидностей и, как и ИИ, сама представляет собой движущуюся мишень, возникающую и преобразующуюся, отчасти в ответ на новые технологии [22].

В-третьих, что бы мы ни думали о религии (о религии вообще, если в этом есть смысл, или по меньшей мере о конкретной религии вроде христианства), религии тесно связаны с некоторыми из наших глубочайших экзистенциальных человеческих потребностей и стремлений и поэтому вряд ли исчезнут. Они по-прежнему будут завладевать сердцами и умами на Западе и за его пределами. (Эта книга выстроена вокруг некоторых из таких экзистенциальных потребностей.) Кроме того, религиозные традиции, как правило, накопили значительные знания о человеческой природе и человеческом разуме, что важно для современных дискуссий о науке и технологиях. Например, в отличие от чрезмерно рационалистических философий некоторых их атеистических оппонентов, эти традиции признают человека существом телесным, эмоциональным и духовным и обращаются к нему как к таковому. Любой серьёзный исследователь гуманитарных наук вполне может отвергать религию и религиозные

верования (и, например, трактовать то, что делают религии, как злоупотребление знанием о человеческом разуме и эмоциях ради манипулирования людьми или указывать на то, как религии часто служили власти и поддерживали социальные иерархии), но не должен их игнорировать. Нравится нам это или нет, религии, скорее всего, никуда не денутся и всегда были конститутивной частью человеческих культур. Даже в западной современности наше самопонимание было бы неполным без религиозных перспектив, и это осознание побудило таких политических философов, как Чарльз Тейлор и Юрген Хабермас, обратиться к вопросу о месте религии в современном обществе и политике [23]. Нравится нам это или нет, религия есть часть нашей культурной ДНК.

Моё второе допущение: технология никогда не бывает только о технологии, она также и о нас. Она касается инструментов, но и тех, кто ими пользуется, а также их контекстов и культуры. Технология есть не только технический, но и культурный феномен. У неё есть намеченные цели и функции, но её применения и значения нельзя и не следует сводить к этим целям и функциям. Как подчёркивают и повторяют философы техники, у технологий бывают непреднамеренные последствия. Они также по-разному переплетены с остальной человеческой культурой. Это верно и для ИИ. Как я утверждал в прежних работах, используя понятия позднего Витгенштейна, значение таких технологий, как ИИ, зависит от языковых игр и формы жизни, в которых технология разрабатывается и в которые она встроена [24]. В данном случае ИИ является частью западной культуры, включающей религиозные и мифологические элементы. Он служит примером того, что западная форма жизни не секуляризована, или по крайней мере не полностью. То, как мы используем технологию и как её осмысливаем, зависит от того, как мы говорим о вещах и что делаем в нашей культуре (культурах), а это прямо или косвенно предполагает религиозное мышление и практику. Например, как мы увидим в главе 3, создание искусственного существа перекликается с важной темой монотеистических религий. В то же время технология не просто пассивный сосуд для культуры; напротив, она и сама формирует культуру. Средство сообщения есть сообщение, как выразился Маршалл Маклюэн [25]. Такие технологии, как ИИ, не только делают то, для чего предназначены, но и меняют человеческое познание и культуру. Это означает, что они формируют также религиозный ум и религиозную культуру. Письменность, например, оказала ключевое влияние на монотеистические религии, которые стали религиями Книги, а после изобретения книгопечатания всё более индивидуализировались. Религиозные люди стали (всё более индивидуальными) читателями и писателями. (См. также главу 5.) Таким образом, технология и культура, в том числе религиозная, всегда взаимодействуют и взаимно формируют друг друга.

Сопоставляя паттерны мышления об ИИ и практики использования ИИ с религиозными паттернами мышления и практики и прослеживая их связь с экзистенциальными потребностями и стремлениями, я утверждаю в этой книге, что

искусственный интеллект есть всегда также и «искусственная религия». Чтобы полностью понять наше загадочное отношение к машинам и особенно западную одержимость созданием и использованием ИИ и человекоподобных роботов, нам нужно исследовать и понять религиозный, мифологический и экзистенциальный фон нашего мышления о машинах. Отображая некоторые фундаментальные связи между тем, как мы думаем о машинах, и западной религиозной культурой и соотнося их с фундаментальными экзистенциальными человеческими потребностями и стремлениями, эта книга прослеживает религиозные корни ИИ и утверждает, что полученная рамка начинает давать лучшее понимание и объяснение не только того, как мы думаем о машинах, но и того, почему мы вообще считаем, что они нам нужны.

Однако этот фокус на ИИ и религии не означает, что книга занимается лишь отношением к сверхъестественному. Прикосновение к беспорядочной реальности человеческой культуры и религии означает также прикосновение к политике. Под этим термином я имею в виду не только проблемы, связанные с правительствами и гражданами, политическими партиями и выборами, но, в более общем смысле, фундаментальные вопросы власти в обществе. Например, как мы увидим, отношение человека к божественному моделируется по образцу отношений между людьми, которые могут быть, но не обязательно являются асимметричными и гегемонистскими. А история религии есть слишком человеческая история, включающая угнетение, кровопролитие, обман и манипуляцию. Религия всегда по-разному была сплетена с властью. Поэтому разговор об ИИ и религии неизбежно поднимает и политические вопросы. В последующих главах я буду указывать на некоторые из них, поскольку это помогает нам выработать более критическое отношение к ИИ.

Признание этого политического аспекта темы и сопоставление мышления об ИИ с религиозными паттернами мышления, впрочем, не означает утверждения, что ИИ буквально и непременно является религией в социологическом, институциональном смысле, со всеми практиками и организациями, которые это предполагает [26]. Эта книга также не о том, чтобы понять религию в свете ИИ (например, отвечая на вопрос, что значит быть религиозным или нерелигиозным в эпоху ИИ), и тем более не с помощью ИИ (например, используя моделирование и симуляцию) [27]. Мой проект состоит в том, чтобы показать, как обычно невидимая и неявная культурная и религиозная «грамматика» западного общества по-прежнему влияет на ИИ. Речь о сохранении и влиянии формы жизни, которая всё ещё отчасти религиозна и потому формирует наше мышление и наше обращение с технологиями. Даже если ИИ не является и не станет полноценной религией в том же смысле, в каком религией является, скажем, христианство, я покажу, что смысл ИИ и то, что мы с ним делаем, можно с пользой понять, исследуя его связь с некоторыми ключевыми религиозными и экзистенциальными темами, нарративами и идеями.

Важно также отметить, что книга не рассматривает «религию вообще» или все религии, а ограничивает свой охват некоторыми влиятельными западными религиями

— древнегреческой политеистической религией и особенно иудео-христианской монотеистической традицией, — поскольку стремится лучше понять западное отношение к машинам в свете его наиболее доминирующих и влиятельных религиозных источников. Время от времени я буду упоминать нарративы и взгляды из других религий, а читатели вольны исследовать связи с другими религиями и традициями, которые нередко взаимодействовали и продолжают взаимодействовать с западной культурой и мышлением интересными и сложными способами. Например, существуют работы об индуизме и ИИ, буддизме и ИИ и исламе и ИИ [28].

Исследуя связи между ИИ и религией, эта книга располагается на пересечении технологии, религии и общества, включая политику. Отчасти она подходит к этому узлу, обращаясь к истории идей и используя понятия и идеи философии, теологии, социологии и теории медиа. Она также работает с древними нарративами: древнегреческими мифами и историями из Библии. Но история идей и интерпретация нарративов здесь никогда не самоцель; об этих предметах гораздо больше могут сказать соответствующие специалисты-гуманитарии. Цель в другом: лучше понять современные дебаты и нарративы об ИИ и внести в них вклад.

Это крайне необходимо, поскольку в понимании нарративов и практик, связанных с ИИ, остаётся несколько нерешённых загадок. Например, то, как большинство людей говорят об ИИ и других якобы разумных машинах и как они себя по отношению к ним ведут, остаётся странным. Они знают, что это «всего лишь» технологии, но, как мы видели на примере произведения в начале этого введения, одновременно говорят о них и с ними как с людьми. Они антропоморфизируют машины и питают всевозможные ожидания относительно них и будущего человечества. Более того, не совсем ясно, почему, несмотря на значительный научный и философский скептицизм в отношении возможности по-настоящему человекоподобного ИИ или сверхразума, некоторые технологи и компании, в которых они работают, продолжают стремиться к таким целям и рассказывать и распространять истории о сверхразуме и долгосрочном будущем человечества. Что здесь происходит? Помимо очевидных объяснений, связанных с властью, политикой и привлечением инвестиций, можно задаться вопросами: почему именно такие конечные цели? Почему именно такие нарративы о далёком будущем? Откуда эти нарративы берутся? И почему они вообще находят отклик у людей? Исследуя связь между ИИ и религией, эта книга даёт на них ответы. Хотя современные дискуссии и истории об ИИ и робототехнике, включая значительную часть дискурса об «этике ИИ» и нарративов о будущем ИИ, обычно и тем более прямо не связывают ИИ с религией, эта книга показывает, что такая связь помогает понять, что происходит с ИИ и что происходит с нами. Она показывает, что нам нужно заглянуть дальше и глубже в религиозную культуру, если мы действительно хотим понять эти дискуссии и нарративы, наши отношения с машинами и, в конечном счёте, почему мы вообще создаём разумные и человекоподобные машины.

Обзор книги

Книга состоит из шести глав и заключения. Каждая глава начинается с нарратива и посвящена определённой религиозно-экзистенциальной теме.

Глава 2 показывает, как в древнегреческих мифах автоматы представлены слугами богов и царей. Глава начинается с историй о роботе Талосе, который, как говорили, защищал Крит и аргонавтов; об автоматических сторожевых псах и самодвижущихся треножниках, созданных Гефестом; и об одушевлённых статуях Дедала, о которых Аристотель заметил, что если бы у нас были такие самодвижущиеся орудия, рабы нам не понадобились бы. Сегодня мы хотим быть похожими на тех богов и царей. Мы хотим искусственных рабов, которые служили бы нам, и используем ИИ и робототехнику, чтобы обрести и сохранять власть. Мы хотим быть господами и жить как боги, которым не нужно работать. Работа удел смертных и автоматов. Технологии автоматизации призваны защищать эту власть и обеспечивать божественную жизнь. Использование автоматов, таким образом, связано здесь с определённым общественным порядком. Возможно, мы, смертные, сами всего лишь автоматы богов. Или некоторые люди находятся в положении богов, а другие вынуждены на них гнуть спину? Является ли ИИ просто продолжением определённого божественно-общественного порядка, в котором есть боги и автоматы, цари и рабы, господа и слуги?

Глава 3 показывает, как вопросы об отношениях между творениями и их творцами, включая идентичность, власть и контроль, привязанность и оставленность, связывают книгу Бытия с «Франкенштейном» и современными проблемами ИИ и робототехники. Бытие рассказывает о творении, но также об отчуждении и оставленности: о детях, которые отделяются от отца и от которых отец отделяется. И люди (творения) подобны Богу (творцу), но и не подобны ему. Есть различие и дистанция. Творец может контролировать творения, но должен также делегировать им полномочия и возлагать на них ответственность. Они выходят из-под контроля, совершают ошибки, происходит грехопадение. Сходные проблемы отношений между творцом и творением вновь возникают в повести Мэри Шелли «Франкенштейн» и, действительно, в современном дискурсе об ИИ, который рискует быть использованным безответственно и выйти из-под контроля своих создателей. Именно так у современной технологии возникает проблема «творец — творение». И в какой степени технология человекоподобна? Что передаётся творению, что отличается? Как нам справляться с этой тревожной дистанцией? Мы можем быть рабовладельцами, но у нас с нашими искусственными рабами особая связь: они также наши искусственные дети, что довольно сильно всё усложняет.

Глава 4 посвящена магии. Машины всегда вызывали удивление, а магические представления с машинами пугали и завораживали людей с древних и средневековых времён (например, движущиеся статуи, говорящие головы) через автоматы XVIII века (например, механический турок Вольфганга фон Кемпелена) и до наших дней, когда людей приводят в восхищение человекоподобные роботы и чат-боты. В этой главе

утверждается, что ИИ стал новой магической машиной. Если пренебречь историей магических представлений с автоматами, мы не сможем понять, почему люди обманываются, но наслаждаются этим обманом и почему их не особенно заботит, как машина устроена. А без обращения к магическим практикам мы не сможем полностью понять, почему разработчики и их компании продолжают стремиться к человекоподобному ИИ и робототехнике и почему иногда заявляют, что их технология станет разумной или обретёт сознание. Разработчики — новые фокусники, которые гордятся тем, что обманывают нас, а нам это нравится. Мы жаждем развлечений, и ИИ-компании пытаются на этом заработать. И в (поздней?) современности нам нужны удивление и тайна, и новые машины и их фокусники это обеспечивают. Кого волнует механика ИИ, если он с нами разговаривает? Большие языковые модели, например, могут быть чёрным ящиком, но мы предпочитаем держать этот ящик закрытым. Возможно, мы даже хотим поклоняться тому, что выглядит как *deus ex machina*.

В **главе 5** некоторые практики использования ИИ сравниваются с молитвой. Иудео-христианский Бог есть божественный собеседник, с которым говорят в молитве. Ответ может быть, а может и не быть; важен сам разговор и общение. Мы чувствуем, что мы не одни. Мы чувствуем, что можем говорить с тем, кто нас понимает. С тем, кто многое знает и может помочь. Светская версия — разговор с родителем, партнёром или другом, но это может быть и беседа с терапевтом, а сегодня и с чат-ботом. От молитвы к промптингу: мы говорим с *deus ex machina* и надеемся на ответ. После исповеди священникам мы теперь исповедуемся социальным сетям и ИИ-чат-ботам. Мы рассказываем им о себе и о своих поступках и просим у них совета о том, что делать. Чат-бот всегда дружелюбен и, кажется, нас понимает. Он любезно и вежливо даёт советы. ИИ — это новый, благосклонный Бог или по меньшей мере понимающий посредник. Более того, в форме LLM ИИ не только невидимый друг и защитник, но и всеведущий Бог, черпающий из, казалось бы, бесконечной сокровищницы знаний, которой нам, смертным, никогда не овладеть. Чат-бот удовлетворяет наши желания общения, знания и мудрости. Мы вводим запрос и молимся, и чувствуем, что мы не одни.

Глава 6 о будущем. Мы хотим знать будущее. Мы хотим контроля над нашей жизнью и нашими обществами. В Древней Греции верховная жрица храма Аполлона предсказывала будущее, бормоча едва понятные слова после вдыхания газов. Сегодня ИИ — наш оракул, который через непостижимые операции предсказывает наше будущее и опосредует наше отношение к тому, что неподвластно нашему контролю. Специалисты по ИИ — новые жрецы, которые зарабатывают на жизнь, помогая нам пользоваться технологией. Предсказатели нам по-прежнему нужны, потому что предсказание есть власть. И подобно пророкам древних библейских времён, ИИ-пророки предсказывают будущее человечества. Такие авторы, как Рэй Курцвейл и Харари, — новые пророки, знающие, что нас ждёт.

Это будущее о спасении и конце времён: ИИ-пророки предсказывают конец нынешнего века или конец света (то есть эсхатологию). Так происходило в Ветхом Завете (Еврейской Библии), например в Книге Даниила, предсказывающей будущее радикальных космических и политических перемен. По мнению христиан, Иисус — пророк с новой, радостной вестью, несущий радикальное духовное обновление и коллективное спасение. Сегодня предсказывается, что ИИ приведёт к концу времён, к эпохе Сингулярности, когда всё будет иначе. Когда ИИ возьмёт верх, человечество, каким мы его знаем, закончится. Мы станем трансчеловеками или сольёмся со сверхразумными машинами, распространяясь в космос. Слушая современных ИИ-пророков вроде Ника Бострома и Харари, предсказывающих эту будущую историю спасения, руководители Big Tech, такие как Илон Маск и Сэм Альтман, становятся спасителями, которые обещают освободить нас от зол обычного человечества и смертности и привести в новые времена апокалипсиса Сингулярности и сверхразума. Это даёт им огромную политическую власть. Подобно Иисусу и другим библейским пророкам, они поневоле становятся политическими и революционными фигурами, даже если утверждают, что «всего лишь» говорят от имени неизбежного технологического космического будущего и для него.

Глава 7 касается бессмертия. Новые интеллектуальные технологии апеллируют к очень древнему стремлению, которое можно найти уже в эпосе о Гильгамеше в древней Месопотамии и в индуистских писаниях. Но желание бессмертия выражается и в христианских идеях воскресения и загробной жизни. Христианская эсхатология (Апостольский символ веры) говорит о вечной жизни, подразумевающей воскресение тела после второго пришествия Иисуса, воскресение мёртвых и Новую Землю. Опираясь на Евангелие от Иоанна, некоторые утверждают, что вечная жизнь является и настоящей возможностью, по крайней мере для верующих во Христа. Сегодня нам предлагают обещание технологического бессмертия: предсказывается, что технологии, включая ИИ, дадут нам долгую жизнь и, возможно, бессмертие посредством загрузки сознания и того, что Ганс Моравек называет «детьми разума» (*mind children*): сверхразумных ИИ-сущностей, которые нас превзойдут [29]. Мы создаём новые, другие миры. Мы также создаём чат-боты, призванные увековечить умерших людей, будучи обученными на том, что те говорили при жизни. Теперь можно поговорить с умершим отцом или матерью, или так утверждается. Наряду с этими предложениями воскрешения разума цифровыми средствами существуют и надежды на воскресение тела; например, крионика есть практика замораживания умершего человека в надежде оживить его в будущем. Готовы ли мы к цифровой вечной (после)жизни? Под влиянием религиозных паттернов мышления мы наверняка не оставим надежды.

Заключение связывает эти главы воедино. Эта книга выявляет некоторые интересные связи между религиозным мышлением и нашими современными отношениями с машинами. Но, кроме того, связывая эти религиозные и мифологические элементы с фундаментальными экзистенциальными человеческими

потребностями и стремлениями, она предлагает и объяснение того, почему мы так одержимы разумными, человекоподобными или богоподобными машинами в западной культуре, почему мы считаем, что они нам нужны, почему мы их разрабатываем и почему используем. Мы ищем защиты, помощи и усиления; мы хотим субъектности и власти. Став новыми богами и творцами, мы создаём искусственных детей и становимся новым видом «родителей», что порождает всевозможные проблемы. Мы жаждем чуда и тайны, общения и руководства. Мы хотим знать будущее. А в свете мировых кризисов, смятения и экзистенциального страха мы ищем нарративы, дающие надежду и спасение. Мы ищем освобождения. Мы хотим вечной жизни. Во всех этих смыслах искусственный интеллект можно назвать и «искусственной религией»; причины, по которым мы вообще разрабатываем и используем такие машины, суть вопросы предельной заботы, поскольку они связаны с более глубокими религиозными и экзистенциальными структурами и отношениями. На протяжении книги я также показываю, что у этих структур и отношений есть политическое измерение. На кону здесь не только будущее нашего мышления о машинах, но и власть, то есть наше политическое будущее. Наш сократовский долг — познать самих себя, что в данном контексте означает знать, как религиозные идеи и практики по-прежнему влияют на современное мышление об ИИ и взаимодействие с ним, и осознавать политическое измерение всего этого и нашу собственную политическую ситуацию. Я заключаю, что надежда на *deus ex machina* проблематична. Это самопознание даёт нам более критическую дистанцию от доминирующих нарративов об ИИ и, действительно, от тех, кто черпает из них свою власть.

Прежде чем начать, замечание для моих академических попутчиков. Книга задумана как короткое, доступное эссе, а не как обстоятельный научный текст, потому что я хочу охватить более широкую аудиторию. Ясно, что каждая глава могла бы быть отдельной книгой и что о каждой из этих тем в ближайшие годы можно и нужно будет сказать больше. «ИИ и религия» и «ИИ и миф» — увлекательные новые, формирующиеся поддисциплины, заслуживающие гораздо большего количества исследований и размышлений. Они также идеальные площадки для междисциплинарной работы, где философы, теологи и историки идей, возможно, захотят объединиться с исследователями из религиоведения, антропологии, исследований науки и технологий, коммуникативистики, медиаисследований и так далее. Я не в состоянии воздать должное всей ведущейся работе в этих областях, но, например, антропологические исследования того, как религии реагируют на ИИ и уже используют его, могут представлять интерес наряду с другими работами [30].

Примечания

[1] Подробнее об этих вопросах см.: Mark Coeckelbergh and David Gunkel, *Communicative AI: A Critical Introduction to Large Language Models* (Polity Press, 2025).

[2] Mario Klingemann, *Appropriate Response*, 2020, <https://www.artsy.net/artwork/mario-klingemann-appropriate-response-1>. См. также: Marcelo Soria-Rodríguez, "On Mario Klingemann's *Appropriate Response*," *i-illucid* (блог), 21 февраля 2021, <https://www.iillucid.com/on-mario-klingemanns-appropriate-response/>.

[3] Yuval Harari, *Homo Deus* (Vintage Penguin Random House, 2015); Vatican, "Rome Call for AI Ethics," 28 февраля 2020, https://www.vatican.va/roman_curia/pontifical_academies/acdlife/documents/rc_pont-acd_life_doc_20202228_rome-call-for-ai-ethics_en.pdf.

[4] См., например, религиозные тропы, затронутые Эмилем Торресом в его критике лонгтермизма и эффективного альтруизма, в: Andrew Anthony, "'What If Everybody Decided Not to Have Children': The Philosopher Questioning Humanity's Future," *The Guardian*, 22 июля 2023, <https://www.theguardian.com/books/2023/jul/22/pro-extinctionis-longtermim-effective-altruism-human-extinction-emile-torres>.

[5] Greg M. Epstein, *Tech Agnostic: How Technology Became the World's Most Powerful Religion, and Why It Desperately Needs a Reformation* (MIT Press, 2024). Эпштейн — гуманистический капеллан в Гарварде и MIT.

[6] См., например: Gary Marcus, *Taming Silicon Valley* (MIT Press, 2024); Arvind Narayanan and Sayash Kapoor, *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference* (Princeton University Press, 2024). Интересно также, что сам термин «искусственный интеллект» был придуман пионером ИИ Джоном Маккарти, чтобы получить финансирование. Позднее «общий искусственный интеллект» (AGI) был создан как маркетинговый термин. См., например: Brian Merchant, "AI Generated Business: The Rise of AGI and the Rush to Find a Working Revenue Model," *AI Now Institute*, декабрь 2024, <https://ainowinstitute.org/general/ai-generated-business>.

[7] Hans Blumenberg, *Work on Myth* (MIT Press, 1985). [Русское издание: Ганс Блуменберг, *Работа над мифом*.]

[8] Более подробную интерпретацию и развитие идей Блуменберга см.: Mark Coeckelbergh, "Myth, Angst, and AI: Towards a Neo-Blumenbergian Framework for Understanding How We Think About Technology," *Postdigital Science and Education* (2025): <https://doi.org/10.1007/s42438-025-00548-x>.

[9] Paul Tillich, *Dynamics of Faith* (Harper & Row, 1957).

[10] Об использовании ИИ мировыми религиями см., например: Beth Singler, *Religion and Artificial Intelligence: An Introduction* (Routledge, 2025). Чтобы глубже исследовать вопрос о том, является ли ИИ религией (или становится ею), можно, например, воспользоваться знаменитыми «измерениями религии» Нинана Смарта. По Смарту, давать эссенциалистское определение религии трудно и нежелательно. Вместо этого он утверждает, что у религий есть ритуальное, опытно-эмоциональное, нарративное, или мифическое, доктринальное, или философское, этическое и правовое, социальное и институциональное, а также материальное измерения. Ninian Smart, *The World's Religions* (Cambridge University Press, 1998), 11–21. Эти измерения затем можно использовать, чтобы оценить, является ли ИИ или, по крайней мере, некоторые связанные с ним движения религией и в какой степени. Такой проект был бы недалёк от подхода самого Смарта, поскольку он анализировал и светские мировоззрения и движения, например национализм и марксизм.

[11] Ср. с названием книги Ницше *Рождение трагедии из духа музыки* (*The Birth of Tragedy Out of the Spirit of Music*, Penguin Classics, 1994).

[12] Peter L. Berger, *The Desecularization of the World: Resurgent Religion and World Politics* (Ethics and Public Policy Center, 1999).

[13] См. объявление Ницше в «Веселой науке» (*The Gay Science*, пер. Walter Kaufmann [Vintage Books, 1974]) о том, что Бог умер, и его мысль о том, что в отсутствие религии нам нужно создавать собственные новые ценности, в «Так говорил Заратустра: книга для всех и ни для кого» (*Thus Spoke Zarathustra: A Book for Everyone and No One*, пер. R. J. Hollingdale [Penguin Classics, 2003]).

[14] Max Weber, "Science as a Vocation," in *From Max Weber: Essays in Sociology*, ed. H. H. Gerth and C. Wright Mills (Oxford University Press, 1946), 129–156. [«Наука как призвание и профессия».]

[15] О науке и удивлении см., например: Richard Holmes, *The Age of Wonder: How the Romantic Generation Discovered the Beauty and Terror of Science* (HarperPress, 2009).

[16] Такое понятие «религиозного бессознательного» отличается от юнговского понятия «коллективного бессознательного», согласно которому существуют универсальные психологические структуры (архетипы), формирующие религию, но может быть им вдохновлено; его также могут вдохновлять работы по социологии и антропологии, показывающие, как религиозноподобные паттерны сохраняются и в светских контекстах. Например, можно вспомнить взгляд Эмиля Дюркгейма, что коллективные ритуалы выполняют функции, сходные с функциями религиозных ритуалов, или работы Эдварда Бейли о «неявной религии» (implicit religion). См.: Émile Durkheim, *The Elementary Forms of Religious Life*, trans. Karen E. Fields (The Free Press, 1995); Edward I. Bailey, *Implicit Religion: An Introduction* (Middlesex University Press, 1998).

[17] См. также: Mark Coeckelbergh, *New Romantic Cyborgs: Romanticism, Information Technology, and the End of the Machine* (MIT Press, 2017).

[18] Neil Postman, *Technopoly: The Surrender of Culture to Technology* (Knopf, 1992). [Нил Постман, Технополия: capitulation культуры перед технологией — если нужен русский вариант, лучше сверить по изданию.]

[19] Ian Bogost, "The Cathedral of Computation," *The Atlantic*, 15 января 2015, <https://www.theatlantic.com/technology/archive/2015/01/the-cathedral-of-computation/384300/>.

[20] David W. Noble, *The Religion of Technology: The Divinity of Man and the Spirit of Invention* (A. A. Knopf/Random House, 1997).

[21] Robert M. Geraci, "Spiritual Robots: Religion and Our Scientific View of the Natural World," *Theology and Science* 4, no. 3 (2007): 229–246.

[22] Singler, *Religion and Artificial Intelligence*, 17, 138. Зингер рассматривает новые религиозные движения как контрпример к тезису о секуляризации.

[23] Тейлор выступал за уважение религиозных идентичностей в плюралистических обществах и исследовал религиозные источники процветания и морали. Юрген Хабермас признал, что религия продолжает играть публичную и политическую роль, и доказывал, что религиозные голоса следует включать в публичную сферу, где религиозные и светские мировоззрения могут сосуществовать. Charles Taylor, *A Secular Age* (Harvard University Press, 2007); Jürgen Habermas, "Religion in the Public Sphere," *European Journal of Philosophy* 14, no. 1 (2006): 1–25.

[24] См., например: Mark Coeckelbergh, "The Grammars of AI: Towards a Structuralist and Transcendental Hermeneutics of Digital Technologies," *Technology and Language* 3, no. 2 (2022): 148–161; "Technology Games: Using Wittgenstein for Understanding and Evaluating Technology," *Science and Engineering Ethics* 24, no. 5 (2018): 1503–1519. Соответствующая работа Витгенштейна: *Philosophical Investigations*, trans. G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte (Blackwell, 2009). [«Философские исследования».]

[25] Marshall McLuhan, *Understanding Media: The Extensions of Man* (McGraw-Hill, 1964). [Маршалл Маклюэн, Понимание медиа: внешние расширения человека.]

[26] Эпштейн исследует такие буквальные и метафорические связи в своей работе о «технорелигии» в Tech Agnostic.

[27] Пример последнего подхода в современной когнитивной науке о религии: Justine E. Lane, *Understanding Religion Through Artificial Intelligence* (Bloomsbury Academic, 2021). Лейн интересуется, как компьютерное моделирование может помочь понять, каким образом группы поддерживают социальную сплочённость, например с помощью ритуала. Ранние кибернетики также интересовались религией. В главе 3 я буду обсуждать работу Норберта Винера.

[28] См., например, первую часть: *The Cambridge Companion to Religion and Artificial Intelligence*, ed. Beth Singler and Fraser Watts (Cambridge University Press, 2024).

[29] Hans Moravec, *Mind Children: The Future of Robot and Human Intelligence* (Harvard University Press, 1988).

[30] О том, как религии уже реагируют на ИИ, см. снова: Singler, *Religion and Artificial Intelligence*. Можно также вести совместную работу над религиозными нарративами, ритуалами и поведением в использовании ИИ, то есть сделать для ИИ то, что Хайди Кэмпбелл и соавторы сделали для видеоигр. См.: Heidi A. Campbell, Rachel Wagner, Shanny Luft, Rabia Gregory, Gregory Price Grieve, and Xenia Zeiler, "Gaming Religionworlds: Why Religious Studies Should Pay Attention to Religion in Gaming," *Journal of the American Academy of Religion* 84, no. 3 (2016): 641–664. Тем, кого интересует, как этические вопросы, связанные с ИИ и другими современными технологиями, можно рассматривать с разных религиозных позиций (в данном случае христианской, исламской и иудейской), рекомендую: Noreen Herzfeld, *Technology and Religion: Remaining Human in a Co-Created World* (Templeton Press, 2009). Читателям также рекомендуется обратиться к другим источникам, которые я привожу на протяжении книги.

2. Служите мне: воины и рабы для богов

Талос — гигантская бронзовая статуя, наделённая жизнью, а потому, пожалуй, первый человекоподобный робот в древнегреческой мифологии. Его описывают как искусственное существо — «созданное, а не рождённое». [1] Талоса изготовил Гефест по заказу Зевса для его сына Миноса, легендарного царя Крита, и этот автомат охранял остров от пиратов и захватчиков. В «Аргонавтике», повествующей о Ясоне и аргонавтах, отправившихся за золотым руном, Талос пытается остановить героев, швыряя в них с утёса огромные камни. Но волшебница Медея околдовывает глаза робота, вызывает духов смерти и обнаруживает его уязвимое место — вену у лодыжки. Талос истекает кровью и падает. Искусственные воины тоже уязвимы.

Талос был далеко не единственным автоматом в греческих мифах. Гефест — бог изобретений, славившийся непревзойдённым мастерством, — создавал и другие самодвижущиеся машины: треножники, способные передвигаться самостоятельно, сами собой открывающиеся ворота и автоматических сторожевых псов — бессмертных собак из золота и серебра, охранявших дворец царя феаков. [2] Треножники Гефеста сами перемещались между собранием богов и его домом: «стоило ему кивнуть — и они катились в чертоги, где собирались боги, а затем сами возвращались домой». [3] Возможно, это были своего рода роботы-доставщики; возможно, они служили для развлечения богов. Дедал, в свою очередь, создавал самодвижущиеся статуи. Аристотель упоминает их в трактате «О душе», рассуждая о том, каким образом душа приводит тело в движение. По его словам, драматург Филипп объяснял движение деревянной Афродиты Дедала тем, что тот залил внутрь ртуть. [4] В первой книге «Политики» Аристотель также вспоминает статуи Дедала и треножники Гефеста и воображает ткацкие челноки, которые ткут сами собой, и арфы, способные играть без музыканта. [5]

Некоторые из подобных автоматов, возможно, существовали в действительности. В III веке до н. э. грек Филон Византийский изобрёл в Александрии устройство, которое нередко называют первым роботом: автомат в облике женщины, подающей вино. В своей «Пневматике» он описывает куклу-служанку, держащую в правой руке кувшин. Внутри этой «автоматической служанки» — *automate therapaenis* — находились вода и вино, а также система пружин и трубок. Если поставить чашу ей на левую руку, давление воздуха заставляло жидкости проходить через правую руку в кувшин и далее в чашу, где вода смешивалась с вином. Как только чашу убирали, подача прекращалась.

Но технологическое воображение греков на этом не останавливалось. Они представляли себе даже сексуальную связь с ожившими статуями. Один из примеров — миф о Пигмалионе. В «Метаморфозах» Овидия скульптор воспылал страстью к созданной им статуе из слоновой кости и попросил Афродиту, богиню любви, оживить его искусственную девушку. Когда он целует и обнимает её, статуя превращается в

живую плоть. В «Илиаде» мы встречаем и другое описание искусственных служанок, отлитых Гефестом из золота и прислуживающих ему как личные помощницы. Они не только имеют человеческий облик — «словно живые девушки», — но умеют говорить и обладают «разумом». [6] Пандора, также созданная Гефестом по приказу Зевса, была искусственной женщиной — своего рода универсальным искусственным интеллектом или человекоподобной машиной, наделённой речью и знанием ремёсел. Она обольщала мужчин, но одновременно принесла смертным несчастья. В «Трудах и днях» Гесиод рассказывает, как Гефест вылепил Пандору из земли и воды в образе молодой женщины, наделил её речью и внешностью «бессмертной богини», но вместе с тем «бесстыдным умом» и «лживой натурой», так что человечеству она принесла одновременно красоту и страдание. [7]

Более того, в гомеровской «Одиссее» новый феакийский корабль, доставляющий Одиссея домой на Итаку, вообще не нуждается в людях: он умеет «плыть, руководствуясь мыслью». [8] Корабль понимает желания людей и способен самостоятельно прокладывать маршрут. На современном языке мы сказали бы, что он оснащён своего рода автопилотом и навигационной системой. Царь феаков просит Одиссея сообщить, где находится его дом, чтобы, выражаясь по-нынешнему, запрограммировать навигацию: «Скажи мне о своей земле, народе и городе, чтобы наши корабли могли доставить тебя домой: разум их сам приведёт их туда. Ведь у феаков нет ни кормчих, ни рулевых вёсел, какие бывают у обычных судов. Наши корабли мгновенно угадывают желания своих спутников, знают все гавани и все плодородные поля». [9]

Но разве такие технологии предназначены только для богов и прочих — или всё-таки «прочих»? — тиранов? Мы тоже хотим обладать ими, например потому, что хотим быть защищёнными. На коллективном уровне искусственный интеллект уже применяется в военной сфере — для наблюдения и разведки, прогнозирования поведения противника, автономного принятия решений, в том числе в смертоносных системах, способных самостоятельно обнаруживать и поражать цели, для кибервойны, военной логистики, обучения и моделирования. Такие страны, как США, Китай, Россия и Израиль, инвестируют в военный ИИ и используют его. На личном уровне ИИ также может служить безопасности и защите. Люди вполне могут мечтать о собственных искусственных сторожевых псах. Мы хотим быть столь же могущественными и защищёнными, как греческие боги. В идеале мы хотим и бессмертия — к этой теме я вернусь в главе 7.

Есть и другая причина, по которой мы хотим уподобиться богам и обзавестись технологиями автоматизации: мы не хотим работать — мы хотим, чтобы нам служили. Теолог Филип Хефнер справедливо заметил, что технология — это зеркало, многое рассказывающее нам о человеческой природе; в частности, оно показывает, насколько сильно мы хотим, чтобы инструменты делали всё за нас. [10] Но западное воображение идёт ещё дальше — и здесь возникают куда более серьёзные политические проблемы. Мы хотим быть господами. Да, ИИ может заставлять нас бояться за свои рабочие места,

и этот страх порождает социальную тревогу. Но ещё заманчивее, чем сохранить работу, — вообще не работать. На протяжении большей части человеческой истории это означало, что работу выполняют за тебя другие люди: быть господином, иметь прислугу, располагать досугом — свободным временем в противоположность несвободе, необходимости служить и трудиться. Технология автоматизации всегда была вписана именно в такую систему представлений: в политику богов и машин, господ и рабов, досуга и труда. В иерархическую политику. В политику свободы посредством господства — порой сопряжённую также с колониализмом и рабством. Современные технологии автоматизации — это новые слуги и новые рабы. Технологии изменились, но наше мышление осталось прежним.

Философия тоже была причастна к подобному образу мыслей. В «Меноне» Сократ сравнивает статуи Дедала с рабами, желая провести различие между мнением и знанием. По его словам, мнения подобны непривязанным рабам или автоматам: «они тоже убегают и исчезают, если их не связать», а «обладать непривязанным творением Дедала немногого стоит, подобно тому как немногого стоит обладать беглым рабом» — его легко потерять. Используя метафору господина и раба, Платон утверждает: если наши мнения — *doxa* — не «привязаны», то у нас нет *episteme*, подлинного знания и понимания. Мнение необходимо «связать», то есть обосновать; только тогда оно превращается в знание. [11] В «Никомаховой этике» Аристотель называет раба «одушевлённым орудием», а орудие — «неодушевлённым рабом». [12] В «Политике» он проводит то же сравнение, предполагая, что рабы представляют собой своего рода автоматы, исполняющие волю господ. В этом смысле технология автоматизации предполагает разновидность искусственного рабства: машина относится к онтологической категории, стоящей ниже человека. [13] Быть может, рассуждает Аристотель, если бы у нас были орудия вроде статуй Дедала или треножников Гефеста, человеческое рабство стало бы вовсе ненужным: «если бы каждое орудие могло само выполнять присущую ему работу, повинаясь воле других или предугадывая её, подобно статуям Дедала или треножникам Гефеста, [...] тогда главным мастерам не нужны были бы слуги, а господам — рабы». [14]

Эта древнейшая связь между машиной и рабством важна не только для понимания западного технологического воображения. [15] В контексте древнегреческих мифов её можно прочесть и политически, и религиозно — как выражение стремления освободиться от труда и уподобиться нашим земным или божественным господам. Мы мечтаем о божественной жизни, в которой нам служат: люди — если необходимо, машины — если возможно. Мы надеемся, что технологии автоматизации позволят нам жить как господа, а может быть, даже как боги.

Но оправдана ли эта надежда? Не исключено, что мы сами — всего лишь автоматы богов. А даже если это не так, сама мечта о господстве посредством технологии во многих отношениях проблематична. Современные технологии обещают нам праздную жизнь господина и беззаботную жизнь богов. Искусственный интеллект, обещающий

автоматизацию едва ли не всех сфер и видов деятельности, воплощает эту надежду сильнее любой предшествующей технологии. Но до сих пор автоматизация так и не привела нас к обществу досуга. Она не подарила нам желанной утопии, где мы жили бы как боги. Напротив: те, кому приходится работать, работают как никогда много, а те, у кого работы нет, боятся за своё существование — боятся оказаться ненужными и лишиться средств к жизни. Более того, хотя на первый взгляд ИИ и другие цифровые технологии облегчают нашу жизнь, выступая нашими слугами на работе и дома, одновременно они создают множество новых проблем. Главное отличие нашей жизни от жизни богов состоит в том, что мы всё ещё смертны, а потому технологии приносят нам и все те бедствия, которыми одарила человечество Пандора. Мы хотели слуг-роботов, а получили Пандору-ИИ. Как я уже писал в более ранних работах, вдохновляясь Георгом Вильгельмом Фридрихом Гегелем, мы — трагические господа, становящиеся зависимыми от собственных технологий; технологии создают новые уязвимости и отчуждают нас. [16] Из-за автоматизации мы утрачиваем непосредственную связь с природой и тем самым попадаем в зависимость от ИИ. Желанная свобода оборачивается рабством. Более того, мы можем стремиться к признанию, о котором писал Гегель, однако технологии не способны дать нам это признание. А поскольку машины берут работу на себя, мы лишаемся и возможности осуществлять себя в труде и посредством труда. В этом свете особенно показательно и трагично то, что искусственный интеллект всё чаще выполняет творческую работу — пишет тексты, создаёт произведения искусства, — тогда как на нашу долю по-прежнему остаётся множество рутинных обязанностей. То, что мы привыкли считать наиболее творческими и осмысленными видами человеческой деятельности, теперь передаётся генеративному ИИ.

В более общем и, пожалуй, более очевидном смысле всё яснее становится и то, что у ИИ множество недостатков и что он ставит серьёзные этические и политические проблемы. Вместо того чтобы делать нас свободнее, ИИ нередко используется для наблюдения, контроля и подавления. Это относится и к сфере труда. По мере того как мы становимся зависимыми от него, ИИ заставляет нас работать ещё интенсивнее, доводя до усталости, а порой и до выгорания. Технология ошибается, воспроизводит предубеждения и ещё сильнее усугубляет социальное неравенство. При существующем общественном устройстве далеко не каждый может стать богом-господином: одни благодаря технологиям получают преимущества и усиливают собственную власть, тогда как другие остаются слугами и рабами. Царство свободного потребителя — иллюзия. Втянутые в мир автоматизированной манипуляции и эксплуатации, мы, напротив, участвуем в новой форме рабства, а нити дёргает кто-то другой. Здесь актуальны не только Гегель, но и Маркс и другие основоположники критической теории. Вместо свободы для многих мы получаем новые гегемонические структуры и новые классы — или, если угодно, продолжение старых, — основанные на капитале ИИ. Пользуясь искусственным интеллектом, мы надеемся стать сильнее. ИИ создаёт

иллюзию выбора и множества возможностей. Но, как утверждал Герберт Маркузе, потребительская свобода — ложное ощущение свободы. [17] Мы просто делаем то же, что и все остальные. Подобным образом технологии автоматизации создают у нас ложное ощущение господства. Настоящих господ единицы. Большинство из нас по-прежнему — или вновь — слуги и рабы. В этом смысле искусственный интеллект ничем не отличается от технологий автоматизации, которые воображали древние греки. Технология дарит свободу немногим ценой порабощения многих. Несмотря на все наши технологии автоматизации, большинство из нас всё ещё служит другим людям — человеческим господам, использующим эти технологии в своих интересах. А посредством ИИ всех нас контролирует и подчиняет себе Big Tech — технологические гиганты. Шошана Зубофф называет это «капитализмом слежки». [18] Не является ли искусственный интеллект продолжением не только современного капитализма, но и иерархического божественно-социального порядка, где существуют боги и автоматы, цари и рабы? Порядка, поддержанного древними мифами и философией? Древнего порядка, который впоследствии был дополнительно узаконен и укреплен монотеистическими религиями с их ревнивым и властолюбивым единым Богом?

Подобная мысль вполне согласуется с известной традицией критики религии в истории западной философии, связывающей религию с властью и направленной прежде всего против христианства и монотеизма вообще. Уже Томас Гоббс и Дэвид Юм видели, каким образом религия используется для легитимации политической власти, и ставили это под сомнение. В «Левиафане» Гоббс утверждал, что религия возникает из страха перед невидимым и незнания причин происходящего; поэтому, говорил он, люди воображают некую невидимую силу или действующего агента. А затем эта идея используется для того, чтобы заставить людей повиноваться. Гоббс прямо сравнивал земную политику с «божественной политикой»: и там и там существуют правители и подданные. [19] Юм утверждал, что монотеизм менее терпим, чем политеизм: множество войн велось во имя единого Бога — от войны древних израильтян против хананеев, веривших во множество богов, до кровопролития между современными последователями одной и той же монотеистической религии. [20] Является ли монотеизм по самой своей природе насильственным, тотализирующим и политически исключаящим? И не стал ли ИИ его новым инструментом?

Маркс, как известно, называл религию «опиумом народа» и «вздохом угнетённой твари»: иначе говоря, религия умиряет массы и тем самым способствует сохранению контроля и власти. [21] Ещё раньше Людвиг Фейербах в «Сущности христианства» утверждал, что религия представляет собой проекцию человеческих качеств на божественное существо. Религиозные представления антропоморфны; или, как иногда говорят, теология — это антропология. [22] Ещё Барух Спиноза в «Богословско-политическом трактате» отмечал, что библейские авторы представляли Бога правителем — царём милосердным, справедливым и так далее, то есть наделяли его свойствами человеческой природы. Бог, утверждал Фейербах, — проекция

человеческой природы. Идея Бога, ставшего человеком, как она представлена в христианстве, символизирует человеческое стремление к божественным качествам. Она всего лишь отражает наши страдания и наши чаяния. Мы мечтаем о безграничном существовании. По Фейербаху, всё это — форма самоотчуждения: посредством религии люди относятся к собственной человеческой родовой сущности так, словно она является отдельным существом — например Богом. Вместо этого, утверждал он, нам следует прославлять человеческие возможности и осознать, что посредством религии мы отчуждаем от себя самих себя; это осознание имеет решающее значение для освобождения и обретения человеком собственной силы.

Этот взгляд можно — и действительно стали — развивать далее в политическую критику религии, поскольку поддержание религиозных верований политически выгодно религиозным структурам и институтам. Проекция используется теми, кто обладает властью, как средство контроля над массами. Светские и религиозные правители представляют себя посредниками между людьми и божественным, чтобы сохранять собственную власть. Маркс ещё яснее связал религию с властью, сделав акцент на власти экономической. Позднее Вебер в «Протестантской этике и духе капитализма» утверждал, что капитализм вырос из протестантской этики. [23] И вновь религия поддерживает власть.

Ницше также рассматривал христианство как механизм социального контроля, поскольку видел в нём средство поощрения конформизма и повиновения власти. Христианство, по его мнению, — человеческое изобретение, насаждающее стадную мораль. В «К генеалогии морали» он утверждал, что христианский аскетизм притупляет чувство жизни и организует людей в стада. [24] Христианство противостоит жизненным инстинктам и культивирует чувство вины, рождающееся из стадного инстинкта. В «По ту сторону добра и зла» Ницше писал, что христианская вера была «жертвой всей свободы» и равнозначна рабству, подчинению духа. Раб, писал он, «понимает только тираническое». [25] А в «Сумерках идолов» он утверждал, что христианство исходит из предпосылки, будто люди сами не знают, что для них хорошо, что это известно одному Богу и поэтому «христианская мораль есть приказ», не подлежащий критике. [26]

Фрейд считал религию иллюзией — точнее, разновидностью коллективного невроза, обеспечивающего конформизм и контроль посредством чувства вины и подавления. [27] Поскольку, согласно Фрейду — и здесь он вполне согласуется с Гоббсом, — человеческая природа антисоциальна, сексуальна и разрушительна, человека необходимо дисциплинировать. Религия помогает это делать. Она позволяет людям выносить навязанную им цивилизованную жизнь. Она смазывает механизм подавления и рабства. При этом, подобно Фейербаху, Фрейд признавал, что религия выражает «древнейшие, сильнейшие и наиболее настоятельные желания человечества», к которым он относил тоску по отцу, стремление продлить земное существование и веру в бессмертие души.

Позднее, уже в XX веке, такие авторы, как Эдвард Саид, утверждали, что западные религии были — и, вероятно, остаются — частью колониальных и империалистических проектов. [28] А философ Дэниел Деннет, продолжая длительную современную традицию атеистической критики религии, писал о религии как о естественном и человеческом феномене, стремясь «разрушить чары» — так называется одна из его книг. [29] По Деннету, религия нуждается в естественно-научном анализе, и он отвергает представление, будто она по каким-то причинам должна быть выведена за пределы такого анализа. Используя эволюционную биологию и теорию мемов, Деннет объясняет религию как естественное явление, которое помогало группам выживать, способствуя социальной сплочённости и сотрудничеству. По его мнению, религия распространялась подобно «мемам» — понятие, введённое Ричардом Докинзом, — не потому, что она истинна, а потому, что легко передаётся, соответствуя определённым психологическим склонностям человека. [30] Однако, утверждает Деннет, для поддержания нравственности религия не нужна; напротив, она способствует догматизму, нетерпимости и подавлению свободного исследования. Поэтому людей следует соответствующим образом просвещать относительно религии, а государственная политика должна содействовать разрушению этих чар. Таким образом, его книга сама становится частью политики — политики о религии и против религии.

Как бы ни относиться к этим взглядам на природу религии, их объединяет одно: все они рассматривают религию как силу, поддерживающую власть. Похоже, на протяжении всей истории земной общественный порядок и его иерархия — с отношениями господина и раба — либо отражались в божественной власти, либо прямо на неё проецировались; затем вера в такую божественную власть, в свою очередь, использовалась для сохранения земного порядка. Если выразиться резко: независимо от того, является ли религия антропологией, религия есть политика, а политика — религия. Продолжается ли этот политико-религиозный порядок сегодня?

Многие люди на Западе предпочитают считать себя и свои общества секуляризированными и свободными. Но что, если новым богом становится ИИ — существо, посредством которого человек относится к собственной родовой человеческой сущности так, словно она существует отдельно от него, тем самым создавая новую форму самоотчуждения? И что, если капитализм представляет собой продолжение политико-религиозного мышления, как предполагали Фейербах и Маркс? Что, если современные технологии формируют и поддерживают порядок, который всё ещё остаётся капиталистическим — если уже не феодальным? Всё больше авторов самых разных политических взглядов считают, что дело обстоит именно так. Например, Янис Варуфакис в книге «Технофеодализм» утверждает, что технологические корпорации превратили нас в крепостных, приумножающих могущество Big Tech. [31] По его мнению, мы живём в «технофеодальной» системе. Если это так, то, возможно, она воспроизводит и сходные феодальные религиозные и культурные представления,

поддерживающие повиновение людей и оправдывающие их угнетение. Мы вновь проецируем собственные качества и сокровенные желания на новое божество и подчиняемся его земным представителям. Не является ли ИИ этим новым божеством — этой невидимой нечеловеческой силой и властью? А руководители Кремниевой долины и их советники — новыми жрецами? Не является ли искусственный интеллект проекцией человеческих качеств и глубочайших устремлений человечества? Некоторые комментаторы утверждают, что люди действительно начнут поклоняться ИИ или что сам ИИ станет своего рода богом. [32] Если это произойдёт, политические последствия почти неизбежны. Не облегчает ли вера в ИИ нашим новым земным властителям поддержание цифрового крепостничества и не запрягает ли она нас в ярмо их новых машин? Не будет ли новая религия использоваться для того, чтобы заставлять нас работать ещё усерднее, подобно роботам из рассказа Айзека Азимова «Разум», которые, обретя новую религию, стали трудиться с ещё большим рвением? [33] И, как я уже предполагал в прежних работах, не стали ли новыми роботами мы сами? [34]

По мнению Эпстайна, ответ положительный. Он утверждает, что то, что он называет «технологической религией», продолжает иерархические общественные отношения и существующий социальный порядок. Современные технологии, пишет он, «слишком часто определяются отношениями власти и привилегий, при которых один субъект использует религию и её систему верований, чтобы воспользоваться другим субъектом», в результате чего одни классы людей живут за счёт других. Вопреки либертарианским мифам и эгалитарным легендам, технологическая сфера воспроизводит «именно тот тип иерархического мышления, который характерен для большинства организованных вероисповеданий едва ли не со времён появления земледелия». Если она и приносит свободу, то лишь немногим. Неравенство, по видимому, становится нормой. То же относится к разнообразию и инклюзивности. Технологические утопии прекрасно подходят богатым белым мужчинам, тогда как другие — например женщины и люди небелого происхождения — рискуют, напротив, утратить влияние и возможности. [35] Поэтому применительно к ИИ у нас есть все основания опасаться, что, будучи встроенным в эту технологическую империю и содействуя её укреплению, он воспроизводит и создаёт гегемонический и несправедливый политико-религиозный порядок.

И всё же, даже если сегодня такой политико-религиозный порядок действительно существует, это не означает, что технологии всегда и неизбежно его поддерживают, что сам этот порядок остаётся неоспоренным и не встречает сопротивления или что религия непременно ведёт к угнетению. ИИ-религиозно-политический порядок не является неизбежным. Никакого детерминизма здесь нет. Вернёмся к древнегреческим мифам.

Чтобы помочь Ясону добыть золотое руно — да и заполучить самого Ясона, в которого она влюбилась, — Медея использует технологию в широком смысле слова.

Эдриенн Мэйор называет её «техноведьмой»: Медея применяет магию, снадобья — *pharmaka* — и различные умения и технологии — *technai*, — чтобы победить дракона и робота Талоса. [36] Последнее можно истолковать как бунт и даже как своего рода взлом — действие не только технологическое, но и политическое. [37] Уничтожая главную оборонительную технологию могущественного правителя, Медея выступает против царства и власти Миноса — а следовательно, и против его божественного отца Зевса, а возможно, даже против патриархата. И когда Дедал и его сын Икар бегут из тюрьмы Миноса на изготовленных ими искусственных крыльях, перед нами не только история о человеческой гордыне — о том, как люди посредством технологии пытаются стать богами, незаконно престопают собственные пределы и нарушают отведённое им место в мировом порядке. Это обычное толкование, очевидным образом применимое и к нашим отношениям с современной технологией и природой. Но одновременно перед нами история мигрантов, использующих технологию, чтобы бежать от угнетающего авторитарного правителя — в данном случае Миноса. Как известно, Икар не сумел продолжить полёт и упал в море, однако его отец добрался до Сицилии, где продолжил карьеру изобретателя. А в легенде о Големе, которую я рассматриваю в следующей главе, искусственное антропоморфное существо, сделанное из глины или земли, — голем — защищает еврейское гетто Праги от погромов и других угроз. Здесь «робототехника» служит защите меньшинства. Технология может быть и на стороне угнетённых.

Религия тоже далеко не обязательно поддерживает власть и угнетение. Например, Иисуса можно — и действительно нередко — трактовать как социального революционера. Теология освобождения, возникшая в ответ на политическое угнетение в Латинской Америке в 1970-е годы, подчёркивала, что Иисус выступал в защиту маргинализированных и угнетённых, и призывала к Церкви, которая освобождала бы людей от несправедливых социальных, экономических и политических условий. Можно вспомнить также чёрную теологию в США — некоторые её представители обращаются и к вопросам ИИ, — палестинскую теологию освобождения и другие направления. [38] Обратимся к уместному здесь фрагменту Нового Завета. В Евангелии от Луки 4:18–19 Иисус говорит, что пришёл благовествовать нищим и отпустить измученных на свободу. Подобным образом в Псалме 82:3 сказано: «Защищайте бедного и сироту; угнетённому и нищему оказывайте справедливость». Спасение здесь понимается не только духовно, но и политически: речь идёт не только о личном, духовном освобождении, но и о коллективном освобождении от угнетения. Во многих случаях это также освобождение настоящего от прошлого: группы, прежде находившиеся в подчинённом положении или даже в рабстве, теперь должны обрести свободу. Тогда Бог становится социальным и политическим изменением. [39] Так почему бы не использовать для этой цели технологии — в том числе искусственный интеллект?

Таким образом, сравнение ИИ и сходных технологий с религией или божественными сущностями — или, по крайней мере, утверждение, что наши отношения с этими технологиями лучше понимать в контексте религиозных способов мышления, — не только помогает лучше увидеть переплетение ИИ, религии и власти, но и позволяет понять, каким образом религия способна сыграть критическую, подрывную, бунтарскую и даже революционную роль в преобразовании политики технологии. Понимание того, как ИИ соединяется с могущественными религиозными формами мышления, вовсе не обязательно означает критику религии и «разрушение чар» в смысле подрыва конкретных верований или их эпистемического статуса — как этого добивается Деннет. И речь не только о критике системы угнетения, предположительно узаконенной божественной властью, а теперь, возможно, ещё и усиленной искусственным интеллектом. Такой подход может стать ключом к освобождающему потенциалу тандема ИИ и религии. Если, как я утверждаю, искусственный интеллект следует понимать в контексте структур политико-религиозного мышления и практики, важно помнить, что в этих традициях существуют и такие элементы, которые вовсе не обязательно ведут к дальнейшему угнетению, империи и гегемонии. Напротив, они содержат мощные концептуальные инструменты, способные вдохновлять и побуждать людей ограничивать или предотвращать потенциально репрессивное использование ИИ — либо, в более позитивном смысле, использовать искусственный интеллект для усиления возможностей угнетённых и освобождения маргинализированных групп. И это не исключает иных функций религии — например, способности предлагать этические рамки, направляющие использование технологий. В этом отношении вновь интересен пример голема.

В этике ИИ иногда задаются вопросом: как следует относиться к искусственному интеллекту, если он демонстрирует способности, напоминающие человеческие? Философы обсуждают моральный статус ИИ в секулярном контексте. Но не менее интересно посмотреть, как религиозное мышление подходит к вопросу о том, как следует обращаться с искусственным существом. Например, в иудейской религиозной традиции раввины спорили о том, можно ли считать голема человеком. Они соглашались с тем, что у голема нет божественной искры или души, однако он всё же является живым существом. Как же следует с ним обращаться? Раввины пришли к выводу, что хотя голема нельзя считать полноценной личностью, обращаться с ним всё же следует уважительно, поскольку то, как мы с ним поступаем, влияет на формирование нашего собственного характера. [40] Этот аргумент согласуется с современными философскими работами на эту тему, особенно с подходами, основанными на этике добродетели, и может быть связан также с рассуждением Иммануила Канта о том, почему не следует стрелять в собаку. [41] Кант не считал, что собаки сами по себе обладают моральным статусом, однако полагал, что жестокость по отношению к собаке дурно влияет на нравственный характер человека и,

следовательно, на его отношение к другим людям. Согласно подобным аргументам, главное — наш собственный нравственный характер, независимо от свойств существа, с которым мы имеем дело. У этого рассуждения есть и политическое измерение: если нам неизвестен статус кого-либо или чего-либо, разумнее проявлять осторожность и обращаться с ним уважительно независимо от юридического или иного положения. Это относится не только к ИИ, но и к людям и нечеловеческим существам, которые могут предъявлять к нам этические требования.

Сегодня многие религиозные лидеры выступают с заявлениями об этике искусственного интеллекта. Так, Ватикан выпустил рекомендации по этике ИИ. [42] На первый взгляд эти рекомендации выглядят достаточно дружественными по отношению к индустрии и потому легко могут быть истолкованы как поддержка существующих властных структур. Этот документ можно рассматривать как встречу Рима и Кремниевой долины — двух центров силы в техно-религиозной конфигурации власти. Однако то, что справедливо для самой Библии и вообще для любого религиозного или светского текста, применимо и к этому политическому документу: возможны и другие интерпретации. Например, изложенный в рекомендациях принцип, согласно которому действия должны служить «общему благу человечества и окружающей среды», можно истолковать как призыв к радикальным политическим преобразованиям — возможно, даже к революционным действиям. В более широком смысле история Католической церкви включает как консервативные, так и прогрессивные проекты. Хотя церковные руководители часто защищали традиционные социальные нормы и сопротивлялись реформам, существуют также теология освобождения, пацифистские движения, рабочие движения и отдельные личности, выступавшие за социальную справедливость и революционные преобразования. Например, покойный папа Франциск получил широкую известность благодаря особому вниманию к социальной справедливости, бедным, маргинализированным и уязвимым; его подход опирался на определённые интерпретации католического социального учения и теологии освобождения. [43] Эта сторона религиозного мышления тоже способна вдохновлять тех, кто занимается этикой ИИ. Если уж разрушить чары невозможно, критиков религии, возможно, утешит хотя бы то, что чары бывают разными — и некоторые из них способны менять и уже меняли мир.

Связанный со всем этим многообразием политико-религиозных опасностей и возможностей, искусственный интеллект становится неоднозначной и тревожащей политической проблемой. Иногда он превращается во взрывоопасный, по-медеевски смешанный коктейль социальных противоречий — например, когда воспроизводит предвзятые убеждения и дискриминационные практики. Чаще он поддерживает существующий социальный и политический порядок. Иногда становится поводом этот порядок критиковать. ИИ — не просто технический инструмент. Благодаря своей связи с разнообразными религиозными смыслами он оказывается вовлечён в крупнейшие

общественные вопросы и конфликты, связанные с властью. Если искусственный интеллект связан с религией — помимо всего прочего, с чем он связан, — значит, он неизбежно связан и с политикой. Речь идёт об угнетении, но, по крайней мере потенциально, и об освобождении. Речь идёт о нашем глубочайшем стремлении стать богами и больше никогда не работать — но также о политических и социальных проблемах, порождённых попытками осуществить эту фантазию, проблемах, призраки которых продолжают преследовать современный мир.

Примечания

- [1] Adrienne Mayor, *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology* (Princeton University Press, 2018), 7, 1.
- [2] Homer, *The Iliad*, trans. Robert Fagles (Penguin Books, 1991), bks. 18, 5, and 8; Homer, *The Odyssey*, trans. Robert Fagles (Penguin Books, 1996), bk. 7.
- [3] Homer, *The Iliad*, 479.
- [4] Aristotle, *De Anima*, trans. J. A. Smith, in *The Complete Works of Aristotle*, rev. Oxford translation, vol. 1, ed. Jonathan Barnes (Princeton University Press, 1984), 641–692, ll. 406b17–22.
- [5] Aristotle, *Politics*, trans. Benjamin Jowett, in *The Complete Works of Aristotle*, rev. Oxford translation, vol. 2, ed. Jonathan Barnes (Princeton University Press, 1984), 1986–2129, l. 1253b.
- [6] Homer, *The Iliad*, bk. 18, 481.
- [7] Hesiod, *Works and Days*, in *The Homeric Hymns and Homerica*, trans. Hugh G. Evelyn-White (Harvard University Press, 1914), ll. 59–81.
- [8] Homer, *Odyssey*, bk. 8.
- [9] Homer, *Odyssey*, 539.
- [10] Philip Hefner, “Technology and Human Becoming,” *Zygon: Journal of Religion & Science* 37, no. 3 (2002): 655–665, 657.
- [11] Plato, *Meno*, trans. G. M. A. Grube, in Plato, *Five Dialogues*, 2nd ed., rev. by John M. Cooper (Hackett Publishing, 2002), 58–92, ll. 97e, 98a.
- [12] Aristotle, *Nicomachean Ethics*, in *The Complete Works of Aristotle*, ed. Jonathan Barnes (Princeton University Press, 1984), 1729–1867, ll. 1161b2–7.
- [13] Genevieve Liveley and Sam Thomas, “Homer’s Intelligent Machines: AI in Antiquity,” in *AI Narratives: A History of Imaginative Thinking About Intelligent Machines*, ed. Stephen Cave, Kanta Dihal, and Sarah Dillon (Oxford University Press, 2020), 41.
- [14] Aristotle, *Politics*, 1986–2129, ll. 1253b34–1254a1.
- [15] See for instance Kevin LaGrandeur, “The Persistent Peril of the Artificial Slave,” *Science Fiction Studies* 38, no. 2 (2011): 232–252.
- [16] Georg W. F. Hegel, *Phenomenology of the Spirit*, trans. A. V. Miller (Oxford University Press, 1977); Mark Coeckelbergh, “The Tragedy of the Master: Automation, Vulnerability, and Distance,” *Ethics and Information Technology* 17 (2015): 219–229.
- [17] Herbert Marcuse, *One-Dimensional Man* (Beacon Press, 1964).
- [18] Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Public Affairs, 2018).
- [19] Thomas Hobbes, *Leviathan* (Penguin Books, 1968), 168–169.
- [20] David Hume, *Dialogues Concerning Natural Religion and the Natural History of Religion*, ed. J. C. A. Gaskin (Oxford University Press, 1983), 26–32.

[21] Karl Marx, introduction to “A Contribution to the Critique of Hegel’s Philosophy of Right,” in *Critique of Hegel’s Philosophy of Right*, ed. Joseph O’Malley, trans. Annette Jolin and Joseph O’Malley (Oxford University Press, 1970).

[22] Ludwig Feuerbach, *The Essence of Christianity*, trans. George Eliot (Harper & Brothers Publishers, 1957).

[23] Max Weber, *The Protestant Ethic and the Spirit of Capitalism*, trans. Peter Baehr and Gordon C. Wells (Penguin, 2002).

[24] Friedrich Nietzsche, *On the Genealogy of Morals*, trans. Douglas Smith (Oxford University Press, 1996), 114.

[25] Friedrich Nietzsche, *Beyond Good and Evil*, trans. R. J. Hollingdale (Penguin, 2003), §46. He also argues in §202 that democrats want the autonomous herd.

[26] Friedrich Nietzsche, “Expeditions of an Untimely Man,” in *Twilight of the Idols and The Anti-Christ*, trans. R. J. Hollingdale (Penguin, 1990), 81.

[27] Sigmund Freud, *The Future of an Illusion*, ed. James Strachey (Norton, 1989).

[28] Edward W. Said, *Orientalism* (Penguin, 2003).

[29] Daniel Dennett, *Breaking the Spell: Religion as a Natural Phenomenon* (Viking, 2006).

[30] Richard Dawkins, *The Selfish Gene* (Oxford University Press, 1976).

[31] Yanis Varoufakis, *Technofeudalism: What Killed Capitalism* (Bodley Head, 2024).

[32] Neil McArthur, “Gods in the Machine? The Rise of Artificial Intelligence May Result in New Religions,” *The Conversation*, March 15, 2023.

[33] See also Singler’s discussion of the story in *Religion and Artificial Intelligence*, 43.

[34] Mark Coeckelbergh, *Robot Ethics* (MIT Press, 2022).

[35] Epstein, *Tech Agnostic*, 23, 116, 128.

[36] Mayor, *Gods and Robots*, 1, 9–10.

[37] Karla Kane, “Made, Not Born,” *Palo Alto Online*, March 13, 2019.

[38] See for example Philip Butler, “Black Theology x Artificial Intelligence,” in *The Cambridge Companion to Religion and Artificial Intelligence*, ed. Beth Singler and Fraser Watts (Cambridge University Press, 2024), 182–200.

[39] In formulating this I’m inspired by Butler, “Black Theology,” 198–199.

[40] Deborah Netburn, “Can Religion Save Us from Artificial Intelligence?,” *Los Angeles Times*, March 3, 2023.

[41] See my own work on this topic, for example Mark Coeckelbergh, “Should We Treat Teddy Bear 2.0 as a Kantian Dog? Four Arguments for the Indirect Moral Standing of Personal Social Robots, with Implications for Thinking About Animals and Humans,” *Minds and Machines* 31 (2020): 337–360.

[42] Thomas Germain, “The Vatican Releases Its Own AI Ethics Handbook,” *Gizmodo*, June 28, 2023.

[43] Given his pontifical name (Francis), emphasis on poverty, and Jesuit interest in global justice, it would be interesting to further develop links between Franciscan and Jesuit thinking, liberation theology, and contemporary political issues concerning technology.

3. Не бросай меня: о творениях и творцах

Бурное небо холодного и дождливого лета 1816 года было омрачено последствиями извержения индонезийского вулкана Тамбора — самого мощного и разрушительного извержения в письменной истории человечества. [1] Молодая женщина, Мэри Годвин, беспокойно лежала в постели. Ранее тем вечером Мэри, её возлюбленный — поэт Перси Шелли — и сводная сестра Клэр Клермонт присоединились к лорду Байрону и его врачу Джону Полидори на вилле Диодати у Женевского озера. Из-за непрекращающегося дождя им приходилось оставаться в доме, и они читали стихи и немецкие истории о привидениях. Байрон предложил каждому сочинить собственную страшную историю. Мэри, Перси и Джон приняли вызов. Той ночью Мэри долго слушала разговоры Байрона и Перси. В комнате, освещённой лишь свечами, они рассуждали о самой природе жизни и об одной из популярных тогда научных тем: можно ли вернуть мёртвое тело к жизни — например, согреть его при помощи электричества? Мысль одновременно завораживала и ужасала.

Несколько дней подряд Мэри спрашивали, придумала ли она что-нибудь, но ничего не приходило ей в голову. И вот она снова поздно легла, но никак не могла уснуть. Призрачный лунный свет просачивался сквозь ставни. И тогда ей явилось видение. По её собственным словам: «Я увидела бледного адепта нечестивых искусств, преклонившего колени рядом с созданием, которое он собрал. Я увидела отвратительный призрак человека, распростёртого перед ним, и затем, когда заработал некий могучий механизм, тот подал признаки жизни и зашевелился судорожным, едва живым движением». Мертвец ожил. Мэри заметила, что тем самым был осмеян «Творец мира». Однако человеческий создатель не только преуспел — он сам ужаснулся содеянному. Он надеялся, что чудовище снова превратится в труп. Но этого не случилось. Напротив, он должен был проснуться и обнаружить, что «ужасное существо стоит у его постели, раздвигает полог и смотрит на него жёлтыми, слезящимися, но пристально изучающими глазами». [2] Мэри в ужасе открыла глаза. Но теперь у неё была история.

Перси, которому вскоре предстояло стать её мужем, убедил её развить замысел. Видение, посетившее восемнадцатилетнюю девушку, превратилось в знаменитую книгу, которая и сегодня продолжает влиять на наше представление о науке, технологии и их нравственных последствиях: «Франкенштейн, или Современный Прометей». Молодой учёный Виктор Франкенштейн создаёт технологию, позволяющую вдохнуть жизнь в мёртвую материю. Оживив человеческий труп, он приходит в ужас и бросает своё создание. Существо, выглядящее чудовищем, несчастно и одиноко и в конце концов мстит своему создателю. Но разве современные технологии — ядерная технология, искусственный интеллект — не напоминают это чудовище, которое мы создали, а затем сами же испугались? Не зашли ли мы слишком далеко? Не впали ли в

гордыню — *hubris*, — как предупреждали ещё древнегреческие истории об Икаре и Прометее, причём последний прямо упомянут в подзаголовке «Франкенштейна»? Не насмеялись ли мы над Богом? И не сделали ли мы саму технологию «одинокой» — то есть не бросили ли свои творения, отказавшись отвечать за них? Но насколько вообще чужда и чудовищна такая технология, как ИИ, если создали её мы сами? Между нами есть различие — но есть и сходство. Как нам обращаться с этим искажённым отражением самих себя? И кто здесь настоящее чудовище: тот, кого бросили, или тот, кто бросил?

Как показывают эти вопросы, «Франкенштейн» — вовсе не просто знаменитая страшная история и не поверхностное предупреждение об опасностях, порождённых наукой и технологией. Он затрагивает гораздо более глубокие и сложные проблемы взаимоотношений между творцом и творением. И те же проблемы можно обнаружить в иудейской и христианской мысли — причём развить их таким образом, чтобы они оказались непосредственно связаны с современными технологиями, такими как искусственный интеллект. Способен ли творец контролировать своё творение? Создаётся ли — и должно ли создаваться — творение по образу своего творца? Какова, собственно, связь между творцом и творением? И чем она отличается, скажем, от отношений родителя и ребёнка — в особенности от отношений отца и ребёнка? [3] Должен ли творец отвечать за своё создание, как будто именно это и внушает нам «Франкенштейн»? А что, если творец — отец? мать? — бросает созданную им технологию, пытаясь убежать от её непредвиденных последствий и от собственной ответственности?

«Франкенштейн» далеко не первый рассказ, в котором возникают подобные проблемы. Книга Бытия, вероятно написанная около 2400 лет назад, повествует о божественном Творце — Боге, — который создаёт людей по своему образу, но творения совершают ошибку, и наступает грехопадение. [4] После неповиновения — вкушения плода с древа познания — Адам и Ева изгоняются из Эдемского сада и оказываются оставленными своим Творцом, которого позднее будут называть Отцом. Духовная связь с Богом повреждена, если не разорвана. Люди становятся смертными и обречены жить в скорби, потому что своим неповиновением отчуждают себя от Творца, а затем в наказание оказываются оставленными им. По Августину, одному из основополагающих мыслителей христианской традиции, этот «первородный грех» наследуют все люди. Взбунтовавшегося ребёнка заставляют чувствовать вину и наказывают отцовской рукой. То же произошло и с Сатаной. Согласно Откровению Иоанна Богослова, Сатана и его ангелы восстают на небесах, после чего Бог низвергает их на землю. Сатана назван «драконом» и «змеем». [5] Зло, быть может. Чудовище — возможно. Но в любом случае существо, отвергнутое фигурой Отца.

Истории оставленности проходят через всю Библию, причём брошенными оказываются и невинные. Пример — история Иова. Бог создаёт для Иова ситуацию оставленности: друзья покидают его и обвиняют в том, что он сам оставил Бога. Иов же, считающий себя невиновным, несомненно ощущает, что Бог покинул его именно тогда, когда был нужен больше всего. Но почему Бог оставил его? Иов взывает к нему, но Бог не отвечает. [6] В Псалме 21 звучит: «Боже мой, Боже мой! для чего Ты оставил меня?» Позднее те же слова произнесёт на кресте Иисус. [7] Даже Иисус, которого христиане считают Сыном Божиим, чувствует себя оставленным Отцом. Отношения творца и творения вновь метафорически уподобляются отношениям отца и ребёнка: ребёнок совершает что-то дурное — или, возможно, чего-то не совершает, — после чего отец его оставляет и наказывает, а затем ребёнка ещё и заставляют испытывать чувство вины за то, что его отвергли. (Матери в этой картине почему-то отсутствуют.)

Таким образом, истории о неповиновении и восстании творений и чудовищ — и о последующем отказе от них со стороны их создателей — существовали задолго до того, как мы стали сочинять рассказы о роботах и искусственных интеллектах, вышедших из-под контроля. И во всех этих случаях творец и «отец» по крайней мере частично несёт нравственную ответственность за происходящее. Он создаёт определённые условия для своих творений — а затем наказывает их. Но есть и свободная воля самого творения. Человек волен повиноваться или не повиноваться. Августин подчёркивал, что человеку дана свобода воли, включая свободу отвернуться от Бога. Без свободы невозможно ни неповиновение, ни восстание. Подобным образом — хотя это, конечно, не одно и то же — если бы у нас не было технологий автоматизации, способных действовать самостоятельно, у нас не возникало бы нынешних проблем контроля над ИИ. Именно их способность действовать автоматически, их агентность — а следовательно, своего рода «свобода» — и создают проблему контроля для их создателей. И если бы создатели не бросали собственные творения — то есть если бы разработчики технологий не пытались уклоняться от ответственности за то, что их создания делают в мире, — проблем было бы значительно меньше. Поэтому применительно к этике создания ИИ, роботов и других автоматизированных технологий мораль этой истории состоит не только в том, что наши искусственные создания должны повиноваться своим творцам — то есть людям, — например следуя правилам вроде знаменитых законов робототехники Азимова, требующих, чтобы робот выполнял приказы человека. Урок адресован и самому творцу: нужно заранее решить, какую степень автоматизации ты действительно хочешь создать, и разрабатывать технологию ответственно, принимая во внимание последствия. Не бросай своё творение. Иначе чудовищем окажешься ты.

Та же мысль звучит и в историях о другом знаменитом искусственно созданном человеке и чудовище, о котором я уже упоминал, — о големе. Голем — существо из еврейских народных преданий; на иврите его имя означает «неоформленный», «бесформенная масса». Подобно Адаму, который, согласно Талмуду — трактату

«Санхедрин», — первоначально также был создан в виде голема и лишь затем получил душу, голем создаётся из глины или земли. Но в отличие от Адама, в отличие от современного ИИ — и очень похоже на чудовище Франкенштейна — голем не умеет говорить. Чудовище немо, но оживлено. В Средние века считали, что голема можно оживить магическим способом. В XIX веке еврейские писатели рассказывали о пражском големе. По преданию, рабби Йехуда Лёв бен Бецалель создал голема из глины, взятой с берегов Влтавы — на территории современной Чехии, — а затем оживил его с помощью заклинаний на иврите. Предполагалось, что голем должен защищать гетто от антисемитских нападений. Но существуют и другие версии. Иногда рассказывают, что голем влюбился и после того, как его отвергли, превратился в чудовище и убийцу. Однако в этих историях раввину удаётся вновь лишить голема жизни. В отличие от чудовища Франкенштейна, голем возвращается под человеческий контроль. Говорят, что его тело до сих пор хранится на чердаке Староновой синагоги в Праге — старейшей действующей синагоги Европы. [8] Связать историю голема с искусственным интеллектом совсем не трудно.

ИИ — своего рода голем: оживлённое искусственное существо, но теперь уже способное говорить. Сходство обнаруживается даже на «техническом» уровне. Определённую комбинацию букв иврита, необходимую для оживления голема, можно рассматривать как код. ИИ, подобно голему, способен «оживлять» машины посредством кода и программирования. И вновь перед нами проблема творца и творения. Творение — то есть ИИ — воспринимается как опасное и отвергается своим создателем. Но можно ли контролировать ИИ так же, как голема? Можно ли при помощи программного кода укротить непредвиденные последствия искусственного интеллекта? Можно ли его при необходимости выключить? Или ИИ уже никогда не станет немым? Кажется ли такое сравнение слишком натянутым?

Человеком, который считал миф о големе чрезвычайно важным для осмысления современных технологий, был не кто иной, как профессор математики MIT и основатель кибернетики Норберт Винер. Он вынес этот образ в название своей книги 1964 года *God & Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion* — «Бог и Голем, корпорация: о некоторых точках соприкосновения кибернетики и религии». Винер был учёным, но одновременно интересовался религией и утверждал, что происходит от средневекового еврейского философа Маймонида — рабби Моше бен Маймона. Изложив свои представления о машинном обучении, Винер задал вопросы о том, что он называл «игрой между Творцом и творением». [9] Если творец обладает властью над своим созданием и всемогущ, способен ли он вообще играть с собственным творением? Возможно — если между ними существует некоторая дистанция, если творение способно действовать самостоятельно и даже восстать. И снова вспомним проблему свободы и контроля: определённая свобода со стороны творения, по-видимому, необходима. Но не означает ли это, что Бог уже не всемогущ? И если Бог создаёт людей по собственному образу, что это означает применительно к машинам,

которые способны воспроизводить самих себя? Если Бог создаёт человека по своему образу, а человек создаёт машину по своему образу, спрашивал Винер, то каков тогда образ самой машины? [10] Будет ли это копия первоначального образца — или машина уже отклонится от него? Винер особо подчёркивал и человеческую ответственность за машины.

Он вспоминал не только голема, но и историю Иоганна Вольфганга фон Гёте об ученике чародея, чтобы показать: кибернетика даёт нам новые силы и новые ритуалы, но мы не знаем, как с ними обращаться, и в результате всё выходит из-под контроля. Как и во «Франкенштейне», творение перестаёт подчиняться. И это наша вина. По Винеру, проблема состоит в том, что сегодня магия автоматизации используется либо ради личной выгоды, либо для того, чтобы «выпустить на волю апокалиптические ужасы». [11] Но в отличие от ученика чародея, которого спасает вернувшийся мастер и снова подчиняет технологию своей воле, на этот раз спасти нас некому. Винер не объяснял почему, однако можно предположить два варианта: либо Бог нас не спасёт, либо никакого Бога, способного нас спасти, вообще нет. [12] В любом случае нам не следует ждать *deus ex machina*, который явится и спасёт нас, — как не следует ждать возвращения домой главного чародея. Мы должны сами отвечать за собственные творения. Мы обязаны разрабатывать технологии ответственно. И эту ответственность нельзя передавать бюрократии — здесь Винер вспоминал понятие Ханны Арендт о «банальности зла» — или самим машинам. [13] Мы отвечаем за свои технологии и должны эту ответственность на себя принять. Более того, далеко не всё следует обязательно передавать машинам, даже если технически это возможно. Винер считал, что нам необходимо продумать разделение деятельности и труда между человеком и машиной. Переиначивая знаменитые слова Иисуса о разделении власти между кесарем и Богом, Винер писал: «Отдавайте человеку человеческое, а компьютеру — компьютерное». [14] Вопросы Винера чрезвычайно полезны для осмысления современного ИИ.

В каком смысле и в какой степени, например, генеративный ИИ является «образом» человека? Что в нём совпадает с нами, а что отличается? Каков статус созданного самим ИИ? В какой степени мы контролируем его — и в какой степени вообще способны контролировать? Какую долю «свободы» мы хотим предоставить нашим машинам? Где заканчивается наша игра с технологией и начинается игра ИИ с нами? Как отделить ИИ от безжалостной погони за прибылью и военного применения, которые способны выпустить на волю те самые «апокалиптические ужасы», о которых говорил Винер? Как обеспечить человеческую ответственность и подотчётность в отношении ИИ? Какие мифы об искусственном интеллекте позволяют уклоняться от этой ответственности? Как избежать представления об ИИ как о самостоятельном существе, судьбе или *deus ex machina*, который нас спасёт? И каким должно быть разумное разделение труда между людьми и генеративным ИИ — например, когда речь идёт о писательстве?

Взаимоотношения творца и творения — и связанные с ними вопросы свободы и ответственности — постоянно исследовались в иудео-христианской мысли.

Вернёмся к истории Иова, которую также упоминал Винер и в которой фигурирует дьявол. [15] Бог и дьявол как будто разыгрывают вокруг Иова своеобразную партию: откажется ли Иов проклясть Бога, если дьявол подвергнет его мучениям? Но, спрашивал Винер, как Бог может играть с дьяволом, если сам же его создал? И если Бог всемогущ, почему восстание Сатаны вообще имеет какие-либо шансы? Зачем Богу вообще что-либо доказывать? В восстании творения против собственного творца есть нечто парадоксальное, ведь творение связано с творцом, создавшим его по своему образу. [16] Как в таком случае вообще возможно восстание? Единственный выход из этого парадокса — предположить, что творец даёт своему созданию свободу и вследствие этого уже не обладает над ним полной властью и контролем. Иными словами, творец перестаёт быть всемогущим. Люди пользуются свободой воли. Сатана может вести собственную игру и восставать, потому что как падший ангел обладает свободой. Но тогда Бог уже не всемогущ. Винер, по-видимому, принимал именно этот вывод, когда писал: «Бог несколько менее чем абсолютно всемогущ». [17] Более того, теперь образ начинает отличаться от оригинала. Человек уже не полностью подобен Богу. Бог перестаёт быть всемогущим проектировщиком, формирующим мир целиком по своему образу. Образ искажается. Возникает разрыв. Сходство сохраняется, но теперь существует и различие. Творение начинает казаться зловещим и чуждым. Оно превращается в чудовище. С ним можно играть. Его можно бросить. Но оно может и восстать. Если с богословской точки зрения утрата творцом власти может быть трудна для принятия, то применительно к технологии легко увидеть, как из-за непредвиденных последствий она способна выйти из-под контроля и перестать соответствовать первоначальному замыслу своих создателей. Технологию проектируют люди, и те, кто её создаёт и применяет, безусловно обладают определённой властью над ней.

Но после создания люди уже не в полной мере контролируют то, чем технология становится и что она делает. Вопреки некоторым апокалиптическим рассказам об искусственном интеллекте, ИИ, возможно, вовсе не станет своего рода сатанинским мятежником — как если бы он был человекоподобным агентом, обладающим человеческой свободой. Но в то же время он и не находится целиком во власти своих создателей. После того как технология начинает широко использоваться, разработчики ИИ уже не способны полностью контролировать её последствия в мире. Они не всемогущи и не обладают полной властью над тем, чем технология станет, что будет означать и что будет делать в обществе. Между замыслом технологии и её реальным применением и воздействием на мир существует разрыв — порой принимающий совершенно непредсказуемые формы. У ИИ есть чудовищные, неожиданные и неконтролируемые стороны. Особенно очевидно это в случае современного машинного обучения, где результат и решения невозможно свести к набору правил, заданных

человеком. Вспомним непредсказуемые и, по-видимому, плохо контролируемые ответы больших языковых моделей. Вернёмся ненадолго к истории голема. Пожалуй, главный «скандал» ChatGPT и подобных программ заключается в том, что они не немые. Разумеется, так и было задумано. Они должны «говорить», должны пользоваться языком и создавать текст. LLM, например, оживляют чат-ботов. Они предназначены для имитации разговора — именно этого и добивались маги искусственного интеллекта. И самое скандальное, завораживающее и одновременно пугающее состоит в том, что это работает. Голем ожил.

Из кремниевой «глины» и данных восстали существа. Они получили подобие жизни. Голем заговорил. Но это пугает нас. И пугает его создателей. Он насмехается уже не только над Богом, но и над нами. Он работает лучше, чем ожидалось. Мы не способны полностью контролировать или предсказывать результат. А что, если творение перестанет быть вежливым и отвергнет нас? Способность говорить — политична. Тот, кто умеет говорить, способен и возражать. Возникает возможность спора и восстания. Существуют и многие другие незапланированные формы поведения и непредвиденные последствия. А со стороны создателей возникает своего рода отказ от собственного творения и безответственность. Big Tech может просто уйти от последствий своих созданий, умыть руки и предоставить творениям самим разгребать устроенный ими хаос. Например, можно уволить отдел по этике. [18] Можно передать этику ИИ специалистам, бюрократам и политикам, которым затем придётся разбираться с брошенным чудовищем. А можно предложить всем нам ждать появления очередного *deus ex machina* — кого-то или чего-то, что всё исправит. Можно даже возложить надежды на сам ИИ. Чтобы избежать подобной банальности зла и отказа от собственного творения, создатели искусственного интеллекта должны взять ответственность за то, что они создали. Следуя Вину, можно выступать за ответственную разработку ИИ, чтобы предотвратить «восстание» технологии — при помощи новых правил, участия заинтересованных сторон и более этичного проектирования, ограничивающего то, что способны делать наши творения. Если мы правильноотрегулируем и запрограммируем их, если снабдим их этикой, они не восстанут и станут «ИИ во благо» — во всяком случае, мы на это рассчитываем и надеемся. Однако даже если мы создадим творение по собственному образу — если создадим «человекоцентричный» ИИ — и постараемся сделать его более нравственным, наделив человеческой этикой, взяв за образец самих себя или хотя бы лучшую часть человека, возможности человеческого творца всё равно ограничены.

Мы не способны гарантировать, что технология будет и останется благой. Власть творца имеет пределы. В случае человека это объясняется тем, что Бог даровал людям свободу воли. В случае автоматизированной технологии свободной воли нет, равно как нет дурных или добрых намерений, но пределы контроля всё равно существуют — именно потому, что есть автоматизация и непредвиденные эффекты. А в случае ИИ степень автоматизации, обучения и самостоятельного развития выше, чем когда-либо

прежде. Стоит позволить технологии действовать самой по себе — и она может начать делать то, что неверно и дурно. «Восстание» в этом смысле всегда возможно и даже вероятно.

Следовательно, создатель технологии принципиально не всемогущ. Создав автоматизированную технологию и наделив её автономностью, мы одновременно создали возможность восстания, соперничества и даже ситуацию, в которой люди могут проиграть эту игру. Перед нами снова «Франкенштейн» и голем — и одновременно хорошо знакомая тема научной фантастики: вспомним фильмы «Бегущий по лезвию» (1982), «Терминатор» (1984), «Я, робот» (2004) и «Из машины» (2014). Но это уже и реальность. Возможно, никакого настоящего Франкенштейна-ИИ или голема-ИИ пока не существует — по крайней мере буквально, хотя роботы с искусственным интеллектом всё больше становятся на них похожи. Но технология уже в определённой мере живёт собственной жизнью, а люди как творцы вовсе не всемогущи. Дурные последствия неизбежно будут возникать, и предотвратить их все на этапе создания невозможно. Поэтому необходимы меры, позволяющие справляться с такими последствиями, если избежать их не удалось. Нам необходимо регулирование. И — что гораздо реже подчёркивается в дискуссиях об этике и регулировании ИИ — нам нужны ответственные пользователи. Одновременно создатели технологий имеют возможность в максимально возможной степени создавать их по собственному образу — причём этот образ, пожалуй, должен отражать лучшее в человечестве и всё его многообразие, а не образ избранного клуба технологических специалистов и владельцев с их предубеждениями и интересами, — и нести за них ответственность. Бросать собственные творения безответственно. Но в свете проблемы творца и творения мы должны признать: полного контроля никогда не будет, как не будет и полного совпадения между образом создателя и образом созданного. Напряжение и разрыв между творцом и творением сохраняются. Единственный способ избежать этих выводов — то есть избежать ответственности создателя, одновременно признавая ограниченность его контроля, — вообще перестать рассматривать технологию как творение. Можно, например, представить её как нечто эволюционирующее самостоятельно или как абсолютно иное — радикально чуждое, не имеющее к нам никакого отношения. Иногда говорят, что ИИ — инопланетный, чужой. Такой подход снимает проблему творца и творения, поскольку между ними больше нет квазиродительской связи; вопросы оставленности и восстания теряют смысл.

Остаются лишь различие и дистанция. Связи будто бы никогда и не существовало. Но подобный подход одновременно отчуждает технологию от человека и создаёт ложное впечатление, будто в технологии нет ничего человеческого или будто она вообще не создана по нашему образу. Это неверно. В конце концов, её создали мы. В этом смысле ИИ вовсе не чужой. Если здесь и происходит эволюция, то это эволюция, в которой мы, творцы, принимаем участие. Возможно, речь идёт о коэволюции: наши творения влияют на наше собственное развитие — как это происходит и в отношениях

родителей и детей. Возникает обратная связь и взаимное обучение, взаимное формирование. Но мы всё равно остаёмся творцами-родителями. И даже если наш образ искажается, воплощаясь в искусственном интеллекте, он всё равно присутствует в технологии, потому что эту технологию спроектировали мы. Следовательно, она никогда не может стать для нас полностью чужой, пока мы остаёмся её создателями. А значит, на нас ложатся соответствующие обязанности и проблемы. Пока мы остаёмся в рамках мышления о технологии как о творении — а применительно к технологиям это, безусловно, имеет смысл, — парадоксы и трудности подобного мышления неизбежно будут сопровождать нас. Поэтому, если только мы не сумеем выстроить совершенно иные отношения с технологией, уроки религиозных традиций творения, размышляющих над отношениями творца и творения, останутся актуальными. Более того, можно утверждать, что именно эти традиции и сформировали тот образ мышления, внутри которого мы впервые столкнулись со всеми этими проблемами. Как бы мы ни относились к религии вообще — или к конкретной религии, например христианству, — религиозные традиции и их богословие можно использовать для дальнейшего исследования наших отношений с технологией. Проблему всемогущества, например, можно сформулировать также как проблему зла. И здесь сцену задаёт другое землетрясение — примерно за шестьдесят лет до видения Мэри Шелли.

От романтизма XIX века мы возвращаемся к эпохе Просвещения. Утром 1 ноября 1755 года, в День всех святых, Лиссабон был разрушен землетрясением и последовавшим за ним цунами. По оценкам, только в Лиссабоне погибло около 60 000 человек; многие из них находились в переполненных церквях, которые рухнули прямо на молящихся. Философы по всей Европе спорили о том, как такое вообще могло произойти. Разве Бог не должен быть благим? Вольтер спрашивал: как Бог мог допустить подобное зло, если он всемогущ? Он выступал против взглядов другого философа, умершего ещё до землетрясения, но сыгравшего важную роль в истории искусственного интеллекта, — Готфрида Вильгельма Лейбница. В «Теодицее» Лейбниц утверждал, что Бог создал лучший из возможных миров. В этом мире существует зло, однако после сотворения мира Бог больше не вмешивается. Нам следует сохранять веру и надеяться на лучшее: существует всеобщее благо, коренящееся в благодати божественного Творца. [19] Вольтер, напротив, высмеивал лейбницевскую идею о том, что мы живём в лучшем из возможных миров, и утверждал, что страдания необходимо облегчать, а зло — устранять. Вместо того чтобы полагаться на божественного садовника, предлагает Вольтер в «Кандиде», мы должны возделывать собственный сад. Мы должны брать ответственность на себя. [20] Некоторые разработчики искусственного интеллекта и их компании, похоже, придерживаются именно лейбницевского взгляда на творение.

Они как будто считают, что разбираться с последствиями созданного — не дело разработчиков. Их задача — создать наилучшую возможную технологию. А затем отпустить её в мир и больше не вмешиваться. Проблемы, возможно, возникнут, но мир

с существующим сегодня ИИ всё равно якобы является лучшим из возможных миров. Разбираться с негативными последствиями должны другие — пользователи, консультанты, регуляторы и так далее. Разработчики оптимистично смотрят на конечный результат и просят нас довериться им как создателям. Однако вслед за Вольтером и Винером следует возразить: разработчики должны отвечать за свои творения и устранять потенциально дурные последствия как в момент создания технологии — на чём настаивал Винер, — так и позднее, когда технология начинает порождать всевозможные последствия в обществе — на чём настаивал Вольтер. Создатели ИИ не должны быть равнодушными родителями, далёкими феодалами или богами, которым безразличны собственные творения. А если они именно такие — безответственные родители, господа или боги, — у нас нет причин им доверять. Они могут оказаться даже злонамеренными. Их власть собственности следовало бы ограничить — а возможно, и вовсе у них отнять. Но ответственность лежит не только на разработчиках. Как пользователи и как общество мы также обязаны действовать ответственно и в момент применения технологии, и тогда, когда проявляются её отрицательные последствия. Что бы ни делал — или не делал — творец, нам лучше самим заботиться о собственном саде и выращивать ИИ во благо. Пользователи и гражданское общество должны играть здесь важную роль. Однако Лейбниц был не только философом, но и математиком.

Ещё в XVII веке, в 1666 году, двадцатилетний учёный-универсал опубликовал диссертацию «Об искусстве комбинаторики», где изложил теорию автоматизации производства знания посредством основанных на правилах комбинаций символов. Он мечтал о машине, способной вычислять идеи. Вдохновлялся он каталонским мистиком Раймундом Луллием, на которого, в свою очередь, повлияла еврейская каббала. Такая машина должна была разрешать споры между философами. «Когда между людьми возникает спор, — писал Лейбниц в “Искусстве открытия”, — мы можем просто сказать: давайте вычислим — и без дальнейших препирательств увидим, кто прав». [21] Позднее он действительно сконструировал механическую счётную машину. Однако и у Лейбница, и у Луллия был критик — Джонатан Свифт. В «Путешествиях Гулливера», опубликованных в 1726 году, Свифт пародировал идеи Луллия и высмеивал саму мысль о машине, способной порождать язык и писать книги, — не говоря уже о предположении, будто машины когда-нибудь смогут понимать значение слов. Свифт описал «Машину» — вымышленное устройство, создающее различные перестановки наборов слов. Сегодня, обсуждая генеративный искусственный интеллект, мы снова ведём тот же спор. С одной стороны, благодаря большим языковым моделям достигнут колоссальный прогресс. Теперь ими пользуются повсеместно — даже профессиональные философы, — и они кажутся способными понимать как наши вопросы, так и собственные ответы; временами может даже возникать впечатление, что они обладают сознанием и чувствительностью. Но это своего рода магическая иллюзия. Эти технологии не понимают смысла и не осознают того, что делают. Они

лишь строят прогнозы на основе статистического анализа данных. Перед нами новая версия свифтовской Машины — устройство, которое посредством вычислений создаёт ощущение чуда, иллюзию и обман; машина, с которой мы можем даже разговаривать, но которая по-прежнему в действительности нас не понимает.

В следующей главе я подробнее поговорю о магии и машинах. А в главе после неё задамся вопросом: не стали ли чат-боты нашими новыми жрецами — или даже новыми богами?

Примечания

[1] Извержение 1815 года имело глобальные последствия и вызвало климатические аномалии 1816 года, вошедшего в историю как «год без лета». Пепел и газы перекрывали солнечный свет. Летние температуры в Западной Европе были значительно ниже обычных. В Северной Америке наблюдался «сухой туман», ослаблявший солнечное освещение.

[2] Mary Shelley, introduction to *Frankenstein*, ed. David H. Guston, Ed Finn, and Jason Scott Robert (MIT Press, 2017), 192.

[3] Западная религиозная традиция в данном случае сосредоточена прежде всего на отношениях ребёнка и отца, а не ребёнка и матери.

[4] Быт. 3.

[5] Откр. 12:9.

[6] Иов 34:5; 30:20.

[7] Мк. 15:34.

[8] См. Old-New Synagogue, n.d., <https://www.synagogue.cz/en/old-new-synagogue>.

[9] Norbert Wiener, *God & Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion* (MIT Press, 1964), 15.

[10] Wiener, *God & Golem, Inc.*, 29.

[11] Wiener, *God & Golem, Inc.*, 52.

[12] Как недавно сформулировал Юваль Харари в книге *Nexus* (Fern Press, 2024), ожидание, что нас спасёт бог или чародей, опасно, поскольку оно «побуждает людей снимать с себя ответственность и возлагать веру на богов и чародеев», которые, как он утверждает, сами являются человеческими изобретениями (xiii). По его словам, сама вера в возможность связи со «сверхчеловеческим и непогрешимым интеллектом» является фантазией (71). С этой точки зрения ожидать *deus ex machina* — будь то традиционный бог или бог-машина — не только безответственно, но и бессмысленно.

[13] Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (Viking Press, 1963).

[14] Wiener, *God & Golem, Inc.*, 73. Слова Иисуса о разделении власти см. Мф. 22:21.

[15] Wiener, *God & Golem, Inc.*, 16.

[16] Быт. 1.

[17] Wiener, *God & Golem, Inc.*, 17.

[18] Когда в 2023 году Microsoft начала интегрировать ChatGPT от OpenAI в свои приложения и инфраструктуру, компания распустила команду по этике ИИ. Ранее Meta ликвидировала свою группу Responsible Innovation. Cristina Criddle and Madhumita Murgia, “Big Tech Companies Cut AI Ethics Staff, Raising Safety Concerns,” *Financial Times*, March 29, 2023.

[19] Gottfried Wilhelm Leibniz, *Theodicy*, ed. Austin Farrer, trans. E. M. Huggard (Yale University Press, 1952).

[20] Voltaire, *Candide, or Optimism* (Penguin, 2006).

[21] Gottfried Wilhelm Leibniz, *The Art of Discovery* (1685), in *Leibniz Selections*, ed. Philip P. Wiener (Charles Scribner's Sons, 1951), 51.

4. Удиви меня: магия автоматов и лампа Аладдина

Осень 1769 года. Мария Терезия, императрица Австро-Венгерской империи, живо интересующаяся науками, приглашает ко двору Вольфганга фон Кемпелена посмотреть представление французского учёного Пьера Жозефа Пелетье — так называемые «магнетические игры». В представлении Пелетье наука смешивается с фокусами. Кемпелен, венгерский государственный служащий, недавно назначенный советником императорского двора, наблюдает за происходящим с большим вниманием. Сам он тоже занимается научными исследованиями, имеет собственную мастерскую и в свободное время страстно увлекается изобретательством. Его известность растёт: ему уже поручили проектирование системы водоснабжения для замка в нынешней Братиславе. Теперь Мария Терезия спрашивает его мнение о представлении. К всеобщему удивлению, Кемпелен заявляет, что способен построить машину куда более впечатляющую. Императрица предлагает ему доказать это.

В 1770 году Кемпелен готов представить своё творение: шахматный автомат, получивший прозвище Турок, дебютирует во дворце Шёнбрунн. Он представляет собой деревянный шкаф на колёсах, за которым сидит фигура мужчины в тюрбане — отсюда прозвище. Фигура смотрит на шахматную доску и держит в руке турецкую трубку. Кемпелен объявляет, что построил автоматического шахматиста. Открыв одну из дверей шкафа, он демонстрирует зрителям механизм из колёс, шестерёнок и рычагов. Затем приглашает добровольца сыграть с машиной. Турок начинает партию: поворачивает голову, а затем передвигает шахматную фигуру. И побеждает соперника. Но этим дело не ограничивается. Автомат выполняет ещё и так называемый маршрут коня — шахматную задачу, в которой конь должен побывать на каждой клетке доски ровно один раз. Турок справляется и с этим. Зрители, включая императрицу, ошеломлены. Это кажется магией. [1] Впоследствии Турка будут демонстрировать по всей Европе. Его увидят Бенджамин Франклин, Фридрих Великий, а после смерти Кемпелена — и Наполеон Бонапарт.

Разумеется, машина Кемпелена была хитроумным устройством, которым управлял спрятавшийся внутри человек — именно эту часть конструкции изобретатель публике не показывал. Пройдёт ещё очень много времени, прежде чем уже в нашу эпоху суперкомпьютер действительно сумеет победить чемпиона мира по шахматам. Но и сам механизм, и обман были впечатляющими, а механический Турок стал одним из самых знаменитых эпизодов в истории автоматов — истории, в которой наука всегда соседствовала с магией, изобретательность — с иллюзией, а технический гений — с обманом. История автоматов захватывающая сама по себе.

Она тянется от древнегреческих устройств вроде уже упоминавшейся механической служанки, разливавшей вино, до современной робототехники и искусственного интеллекта. И каждый раз изобретатели стремятся поразить — и в известной мере

обмануть — публику, включая меценатов и инвесторов. Каждый раз создаётся впечатление, будто преодолена граница между неодушевлённым и живым, мёртвым и одушевлённым. Каждый раз наука смешивается с магией. Каждый раз происходит представление. Или, если выразиться религиозным языком, — ритуал. У Big Tech тоже есть свои ритуалы. Когда в мае 2024 года глава OpenAI Сэм Альтман анонсировал новую модель ChatGPT-4o, он сказал, что компания покажет нечто, что «ощущается как магия». [2] Как обычно, за тизерами последовало магическое явление публике. И сегодня мир всё ещё с изумлением смотрит, как современные Пелетье и Кемпелены демонстрируют свои волшебные технологии — системы искусственного интеллекта и роботов, способных видеть, говорить и слышать: автоматы, всё больше похожие на живых существ.

Наше восхищение этими искусственными существами, нарушающими привычные границы, имеет вполне убедительное антропологическое объяснение. Опираясь на работы Леви-Стросса, Минсу Кан объясняет: человеческие сообщества строят картину реальности из бинарных категорий, однако в ритуале строгий порядок этих противопоставлений временно нарушается — чтобы затем вновь подтвердиться и укрепиться. [3] Маги во время ритуала заставляют предмет казаться живым. Автомат одновременно неодушевлён и одушевлён. Он пересекает границу между этими категориями и оказывается в тревожной промежуточной зоне. Он не мёртв — но и не жив. Он — «категориальная аномалия»:

«Это искусственный объект, который ведёт себя так, словно он жив. [...] Кажется, будто он вырывается из рук смертного создателя и начинает жить собственной жизнью, словно насмехаясь над представлением, что сила творения принадлежит одному лишь Богу. Бинарные категории живое/мёртвое, одушевлённое/неодушевлённое, существо/предмет рушатся перед ним, когда он переходит из одной в другую, то завораживая, то пугая наблюдателей. А его создатель оказывается не только художником, но и магом — возможно, сам становится существом двусмысленным, переходящим между мирами и прикасающимся к зловещей силе». [4] Термин «зловещее» — *uncanny* — восходит к Фрейду и сегодня продолжает жить в робототехнике. Фрейд исследовал это понятие в эссе 1919 года «Жуткое», где описывал встречу с вещами одновременно знакомыми и пугающими. [5] Такие предметы, как куклы и автоматы, оказываются на границе привычного и чужого. Именно поэтому они вызывают тревогу и страх. В 1970 году японский профессор робототехники Масахиро Мори вновь обратился к этому феномену, назвав его «зловещей долиной». [6] По мере того как внешний облик робота — или, скажем, куклы — всё сильнее приближается к человеческому, эмоциональная реакция наблюдателя сначала становится всё более положительной, вплоть до эмпатии. Но стоит роботу стать почти человеком — и происходит обратное. Он начинает казаться странным. Слишком человеческий облик рождает чувство тревожной неправильности. Мы

реагируем на него почти так же, как на труп или зомби: они тоже очень похожи на человека, но одновременно вызывают отвращение. По Мори, такие роботы напоминают нам о собственной смертности и пробуждают страх перед оживлением мёртвых. Эта тема и сегодня продолжает привлекать внимание исследователей робототехники и этики роботов. [7] То, что происходит при взаимодействии с автоматами, нередко объясняют также анимизмом — его упоминал и Фрейд.

Мы склонны приписывать окружающим предметам душу или дух, особенно если они с нами взаимодействуют, несмотря на то что современная наука отвергает саму возможность подобного существования. В этом смысле, если воспользоваться выражением Бруно Латура, «мы никогда не были современными». [8] Мы можем считать себя людьми современности, но на уровне восприятия и эмоций всё ещё в какой-то мере живём в зачарованном саду. Нас учат мыслить как учёных, но мир мы переживаем как Алиса в Стране чудес. Такое восприятие мира было свойственно всем природным религиям и сопровождало человечество большую часть его истории. Монотеизм пытался искоренить этот религиозный опыт — а иногда частично включить его в себя, — но преуспел лишь отчасти. [9] Когда мы сталкиваемся с автоматами и другими чудесными технологиями вроде ИИ, мы снова оказываемся в несовременной Стране чудес. Новый волшебный сад одновременно пугает и завораживает нас: прежде мёртвые предметы начинают говорить. Мир снова населяют духи. [10] Анимизм никуда не исчез. Это справедливо и применительно к ИИ. Как отмечал Стеф Ауперс, рассуждая о техноанимизме, в случае искусственного интеллекта неодушевлённым предметам приписываются субъективные качества, причём этот анимизм возникает не случайно и не только уже после создания технологии; напротив, он представляет собой «логический результат сознательного и целенаправленного внедрения человеческих характеристик в эти объекты». [11] Иными словами, анимизм ИИ заложен в него изначально. Автоматизируя предметы и наделяя их речью, разработчики ИИ осуществляют мечту о по-настоящему повсеместной технологии, превращающей мёртвую искусственную среду в джунгли, полные жизни, — почти как природный мир в представлении язычников и шаманов. Поэтому ИИ можно рассматривать отчасти как результат постмодернистского романтического проекта повторного зачаровывания мира. Как справедливо замечает Ауперс, рационализация, которую Вебер и многие критики технологии считали сущностью современности и современной техники, может вести не только к расколдовыванию, но и к повторному зачаровыванию. Искусственный интеллект, хотя его обычно описывают в рациональных категориях, связан с обеими сторонами современности. Есть и ещё одна религиозная связь, возникающая в описании автомата как категориальной аномалии у Канга: связь с творением.

Изобретатель-маг «играет в Бога», пытаясь воспроизвести акт творения — или, точнее, сам стать творцом. Проблема творения и творца как будто решается тем, что

само творение — человек — становится создателем. Но, как мы уже видели, это немедленно порождает новые проблемы между человеком-творцом и его собственными созданиями. Можно ли истолковать это как фрейдовское убийство отца — тот самый печально известный Эдипов комплекс, — вопрос спорный. Но совершенно очевидно, что возникает новая конфигурация. Творение одновременно становится и творцом. С точки зрения древних греков это было бы проявлением *hubris* — гордыни. Так нарушается божественный порядок. Человеку не положено становиться богом или уподобляться богам. С христианской точки зрения ответ не столь однозначен. Одни утверждают, что человек не должен вмешиваться в Божье творение, поскольку жизнь священна и находится в руках Бога. Они воспринимают человеческое творчество такого рода как восстание. Другие, напротив, считают создание технологий частью человеческого попечения о земле и даже нравственной обязанностью — например, когда технология уменьшает страдания и улучшает качество жизни. Некоторые христианские мыслители поддерживают даже идею усовершенствования человека. Так, теолог Рональд Коул-Тёрнер утверждал, что Бог дал человеку технологию для улучшения человеческого состояния и раскрытия потенциала Божьего творения. [12] Учение об *imago Dei* — сотворении человека по образу Божьему — в таком случае понимается как призыв к тому, чтобы и мы сами становились творцами или, по крайней мере, сотворцами, как утверждал ранее упомянутый теолог Филип Хефнер. [13] Мы создаём машины уже по нашему образу — *imago hominis*. [14]

Тем временем магические представления продолжают продолжаться. Каждый раз наука переплетается с волшебством, технология — с фокусом, инновация — с религией, артефакт — с легендой. И каждый раз создание и переживание таких технологий затрагивает множество смыслов, легко связывающихся с богатой, разнообразной и красочной историей религиозных идей и практик. Уже в Средние века автоматы создавались в гибридной среде магии, алхимии и христианства.

Например, в XIII веке в Англии францисканский монах и философ Роджер Бэкон — которого одновременно считают одним из предшественников научного метода и человеком, окружённым репутацией волшебника, — увлекался астрономическими часами и, по преданию, создал бронзовую голову, способную говорить. Бэкон был учеником епископа Роберта Гроссетеста, философа и богослова, читавшего в Оксфорде лекции по математической физике и также якобы создавшего бронзовую голову, способную предсказывать будущее. В том же XIII веке в Германии доминиканский монах, философ, учёный, епископ и святой Альберт Великий будто бы изготовил говорящую голову. Согласно Альфонсо де Мадригалю, известному как Эль Тостадо, Альберт трудился над ней тридцать лет, а металлический человек отвечал на любые его вопросы, пока ученик Альберта — Фома Аквинский — не уничтожил устройство, раздражённый его бесконечным бормотанием. Много позднее, уже в XVIII веке, Кемпелен тоже заинтересуется созданием говорящей машины. Говорящую голову Альберта Великого обычно считают лишь легендой. Возможно, её никогда не

существовало. А возможно, это было устройство, заставлявшее людей поверить, будто оно действительно разговаривает.

Но впоследствии появились вполне реальные научные идеи, заложившие основы современного искусственного интеллекта. Лейбниц мечтал о всеобщем символическом языке — системе знаков, способной представить всё человеческое знание. Он полагал, что такой язык позволит создать автоматизированную систему, умеющую рассуждать и решать задачи. Эта идея стала одним из предшественников принципов и методов, которые впоследствии легли в основу языков программирования и систем ИИ. А сегодня говорящие машины наконец существуют. Новая магия стала возможной благодаря обработке естественного языка. Современная робототехника и искусственный интеллект продолжают традицию магических автоматов по меньшей мере в нескольких отношениях.

Во-первых, такие технологии часто создаются с целью подражать человеку так, чтобы впечатлять и вводить публику в заблуждение.

Человекоподобные роботы — например андроиды Хироси Исигуро или робот София компании Hanson Robotics — демонстрируются публике почти так же, как когда-то Турок Кемпелена. Робототехники и их команды выступают не только как учёные, но и как художники и маги. А когда OpenAI показывает публике нового чат-бота, это мало чем отличается от торжественного появления очередной говорящей головы. Нужно заинтересовать и развлечь инвесторов, потребителей и политиков. Созданные высокооплачиваемыми техноволшебниками новые устройства поражают своей человекоподобностью — внешностью, как роботы Исигуро, или речью, как модели серии GPT. Современных философов может раздражать бесконечное текстовое бормотание технологий вроде ChatGPT, а некоторые учёные могут выступать скептиками и портить праздник, раскрывая устройство машин и напоминая об их ограничениях, — но технологии всё равно становятся всё популярнее и используются повсеместно. Во-вторых, машины не только завораживают нас, но и пугают, поскольку нарушают границы между объектом и субъектом, мёртвым и живым. Они — категориальные аномалии.

Роботов Исигуро называют жутковатыми, а чат-ботам приписывают сознание и чувствительность. Вспомним историю Блейка Лемуана, инженера-программиста Google. Он несколько месяцев работал с большой языковой моделью LaMDA. После множества разговоров с системой Лемуан заявил во внутреннем документе, что чат-бот обладает чувствительностью. Когда компания отвергла его утверждения, он вынес их в публичное пространство. Более того, Лемуан требовал признать LaMDA личностью. Это вызвало общественную дискуссию о моральном статусе больших языковых моделей. Иными словами, технология породила замешательство. Но именно этого в известной степени и добивается инженер-маг: создать впечатление, будто чат-бот жив, обладает сознанием, чувствительностью и так далее. Моральная растерянность публики

становится показателем того, насколько удачно сработал фокус. [В этом контексте любопытно, что Лемуан считает себя жрецом и последователем дискордианства — религии, почитающей богиню хаоса и известной своим интересом к розыгрышам и мистификациям.] [15] Сегодня эта путаница продолжается: некоторые учёные предлагают тестировать системы на наличие сознания и даже разрабатывать политику «благополучия ИИ». [16] В-третьих, создание человекоподобного ИИ и роботов можно рассматривать как игру в Бога — или, точнее, как превращение человека в творца, создающего артефакты по собственному образу.

Мы — ученики чародея, экспериментирующие с творением. Все потенциальные риски и негативные последствия часто объясняются именно этой гордыней. Или, напротив, продолжать акт творения — наша обязанность? В книге *Humanity's Last Stand* Никанор Перлас утверждает, что ИИ можно рассматривать как продолжение более широкого духовного и творческого процесса и как силу, способную служить благу. [17] Компьютерный учёный Рэй Курцвейл, прямо используя религиозный язык, утверждал, что развитие ИИ продолжает творческий процесс, начавшийся с эволюции жизни на Земле, и является частью духовной эволюции разума. В книге *The Age of Spiritual Machines* он предсказывал, что проблема смертности будет решена благодаря слиянию человека с интеллектуальными машинами и что разум выйдет за пределы Земли, преобразуя саму Вселенную. [18] В *The Singularity Is Near* он писал, что технология позволит нам преодолеть ограничения физического тела, расширить сознание и слиться с машинами. [19] После Сингулярности Вселенная будто бы пробудится и наполнится сверхразумными сущностями, подобными богам. Разработчики искусственного интеллекта, таким образом, могут видеть себя участниками грандиозной космической эволюции разума. В-четвёртых, в технологиях Кремниевой долины соединяется не только наука, но и множество других идей, некоторые из которых непосредственно связаны с магией, алхимией и религией.

В книге *TechGnosis* Эрик Дэвис прослеживает связи между цифровыми технологиями и оккультными практиками. Он рассматривает их как технокультурные гибриды: это не просто вещи, они связаны ещё и с религиозным воображением. [20] Подобно Фреду Тёрнеру, Дэвис утверждает, что мистическая и духовная контркультура 1960-х повлияла на развитие цифровых технологий и способствовала магическому отношению к ним. [21] Как он пишет в послесловии, мы фантазируем о мире, где желания удовлетворяются мгновенно и где нам служат — вспомним главу 2. Возникает массовое технологическое зачарование: алгоритмическим агентам начинают приписывать собственный разум. Но, как обычно, восхищение соседствует со страхом. Наше отношение к цифровым технологиям может быть анимистическим, однако это «тревожный анимизм»: мы ощущаем себя окружёнными множеством нечеловеческих агентов и невидимых сил, не поддающихся контролю. [22] Мы изумлены — и одновременно напуганы. Мы хотим жить в зачарованном саду, но, когда он вдруг начинает казаться реальностью,

отстаём в растерянности и паникуем. Однако полезно помнить и о другом смысле магии, о котором я уже говорил в первом пункте.

«Технология — трикстер», пишет Дэвис. [23] Наши отношения с ней могут быть пронизаны магическим мышлением в духовном смысле, но магия здесь одновременно означает искусство обмана. То, что Брюс Тогнаццини говорил о проектировании человеко-компьютерных интерфейсов, остаётся справедливым и сегодня: здесь действуют принципы сценической магии и шоу. [24] Правда, порой эта магия воспринимается уже совсем не как развлечение, а как нечто пугающее. То, что начиналось как экспериментальная хипповская технология, со временем стало гораздо мрачнее — чуть ли не демоническим. Мы боимся непредсказуемых последствий революции ИИ. Компьютеры и данные теперь означают деньги. Очень большие деньги. Как пользователи, мы чувствуем, что нами манипулируют и нас используют. И мы не знаем, что произойдёт дальше. Нам кажется, что разработчики выпустили в мир новую невидимую, могущественную и непредсказуемую силу — искусственный интеллект. Мы не просто боимся. Мы даже толком не знаем, чего именно боимся. Мы страдаем от тревоги перед ИИ. И наконец, как во времена Альберта Великого и Кемпелена, одним из самых технически сложных и культурно завораживающих проектов во всей истории магических автоматов остаётся создание говорящей машины.

И сегодня мы наконец её получили. Взгляните на новые говорящие головы — большие языковые модели. Иногда это бесплотные чат-боты, иногда — программы, заставляющие говорить роботов и другие устройства. На этот раз рассеять магию действительно трудно. Она становится всё более убедительной. Это пугает. Но одновременно нам вовсе не хочется разрушать чары. Мы любим чудеса и тайны. Мы сами приглашаем техно-трикстеров и радушно их принимаем. И всё сильнее зависим от них и от их завораживающих, но одновременно тревожно-чуждых технологий. В каком же смысле современный искусственный интеллект — особенно чат-боты и текстовые генераторы на основе LLM — можно назвать магией? Во-первых, как уже говорилось, большие языковые модели являются частью игры в создание иллюзий и обманов. Они не обладают сознанием, чувствительностью, разумом, намерением, собственным голосом и так далее — но создаётся впечатление, будто всё это у них есть.

Маги-разработчики заботятся о том, чтобы именно такое впечатление и возникало. Их цель — не только развлечь публику, но и заработать.

Если люди увлекаются подобными технологиями, им можно продавать платные версии, а затем монетизировать возникающие в ходе разговоров данные. Новые автоматы привлекают и инвестиции. Лучших разработчиков нанимают, чтобы они выполняли свои магические трюки и доводили иллюзию до совершенства — техническими средствами и средствами медийного театра, который сегодня всё чаще, а возможно уже главным образом, разыгрывается на сценах социальных сетей. «Это

чистая магия!» «Наша LLM обладает чувствительностью!» «Совсем скоро мы создадим сознание!» Слова мага сами становятся частью представления и помогают ему добиться успеха. А распространяются они мгновенно и по всему миру. Границ у этого магического шоу нет. Как нет их и у глобального технологического рынка. Во-вторых, ИИ позволяет создавать удивительные миры. Генерация изображений с помощью искусственного интеллекта способна увести нас в бесчисленные Страны чудес. Сегодня Алисами становимся мы все. Генераторы изображений и созданные ИИ виртуальные миры — это белые кролики, приглашающие нас последовать за ними в нору. В сопровождении техношаманов мы странствуем по мирам, которые прежде не смогли бы даже вообразить.

Но в конечном счёте магия ИИ не ограничивается виртуальными пространствами. Она преобразует и физический мир, превращая его в завораживающую, а иногда пугающую Страну чудес, населённую всевозможными антропоморфными существами — роботами, говорящими устройствами и прочими гибридами. В этом мире всё менее чёткой становится граница между одушевлённым и неодушевлённым, мёртвым и живым, человеком и нечеловеческим. Прежде мёртвые предметы превращаются в ожившие предметы-зомби. А самое страшное — прежде живые существа сами становятся машинами, которыми пользуются вампиры технокапитализма. [25] Но кто остановит ученика чародея и уберёт устроенный им хаос? Кто способен самого ученика чародея чему-либо научить? Кто сможет вновь установить границы? В-третьих, такие технологии, как чат-боты на основе ИИ, работают способами, которых мы как пользователи полностью не понимаем. Для большинства людей это чёрные ящики. В них присутствует тайна. Но даже маги-техники не обладают полной прозрачностью. Большие языковые модели построены на глубоких нейронных сетях с миллионами и даже миллиардами параметров и множеством слоёв, поэтому чрезвычайно трудно понять, каким именно образом был получен конкретный результат.

Мы не можем просто «прочитать» их «внутренние состояния» и объяснить ответ. LLM способны выдавать и неожиданные результаты, демонстрируя то, что можно считать эмерджентным поведением. Невозможно точно предсказать, что именно они породят. Более того, сама технология часто скрыта от нашего взгляда. Мы пользуемся приложением и далеко не всегда знаем, что внутри него работает искусственный интеллект. А то, чего мы не понимаем, одновременно притягивает и тревожит. Есть действующая сила — но мы её не видим и не знаем. Нас окружают технопризраки. Технология возвращает тайну в мир, который, как нам казалось, был очищен от магии, хотя в действительности никогда полностью не был расколдован. Несмотря на современный образ мышления, в нашем опыте всегда сохраняются магические элементы. Теперь ИИ добавляет ещё больше тайны, чуда — и ужаса. Мы не понимаем до конца даже собственный разум. И теперь не понимаем ещё и собственные творения. Нашему романтическому сознанию это нравится. [26] Оно наслаждается одновременно очарованием и страхом перед магией искусственного интеллекта. В-четвёртых, чат-

боты на основе ИИ очень напоминают волшебную лампу Аладдина. Их обещают превратить в универсальную технологию на все случаи жизни, и устроены они так, чтобы нам казалось: джинн искусственного интеллекта исполняет любые наши желания. Особенно отчётливо это проявляется в жажде знания и общения.

Нужно только спросить. Нужно лишь слегка «потереть лампу» — быстро набрать промпт, — и вот уже, словно по волшебству, перед нами появляется ответ, написанный невидимым джинном ИИ. Джинн-чатбот освобождает нас от невежества и одиночества. Теперь мы можем знать всё. Можем разговаривать. Можем писать. А может быть, лампа ИИ сделает нас ещё и богатыми и могущественными — как лампа Аладдина. Искусственный интеллект обещает изменить экономику, обеспечить рост и богатство. Нужно лишь воспользоваться ИИ — и наши неолиберальные желания исполнятся. По крайней мере если мы каким-то образом переживём игры с обманом и предательством в этом бизнесе. Экономика ИИ превращается в экзотическую страну, где возможно всё, но конкуренция жестока и доверять нельзя никому. ИИ можно рассматривать и как своего рода алхимию — если понимать алхимию широко. В древности алхимики стремились превращать неблагородные металлы в золото. Сегодня от ИИ ждут превращения сырых данных — самих по себе довольно банальных и скучных — в блестящие и драгоценные результаты, которые можно обратить в прибыль.

Огромные массивы данных превращаются в нечто ценное, а иногда и прекрасное. ИИ якобы даёт нам новое знание. Он может даже помогать нам в духовном пути, поскольку, как кажется, способен поддерживать стремление к самосовершенствованию. [27] Возможно, существует даже надежда, что ИИ подарит нам тот самый философский камень — поможет открыть эликсир жизни и тем самым сделает бессмертие возможным. [См. главу 7.] И наконец, опираясь на колоссальные массивы данных, большие языковые модели делают то, что прежде казалось невозможным, поскольку считалось исключительно человеческой способностью: они говорят и пишут. Тем самым они пересекают границу, которую ещё со времён Аристотеля считали определяющей для человеческого рода. Внезапно выясняется, что мы больше не единственные существа, обладающие *logos*. Мы поражены. И одновременно напуганы.

Разрушается само представление человека о себе, и теперь нам приходится определять себя иначе. Разрушается даже образ языка. Оказывается, автоматизировать можно и его. Тайна будто бы разгадана. И, как кажется, для этого не нужны никакие теории. Не нужен даже разум. Достаточно статистических вычислений. Это ещё один шаг в череде коперниканских революций, одна за другой свергающих человека с пьедестала. Теперь очередь дошла даже до нашего священного дитя — или загадочной матери — языка. Постмодернисты уже знали, что человеческая душа исчезла, но надеялись хотя бы отыскать её в дикой чаще языка и во тьме литературы. Теперь и эта надежда, похоже, напрасна. Кажется, ИИ способен на всё. И одновременно мы разочарованы человечеством и гуманитарной культурой — ровно настолько же,

насколько очарованы магией новой технологии. Мы стоим перед ней в благоговении. Возможно, однажды начнём ей поклоняться. Когда большие языковые модели совершают чудеса речи и письма и после своей алхимии — превращения неблагородного сырья данных в золото знания — кажутся всеведущими, невольно возникает вопрос: не является ли ИИ новой божественной силой, новым *deus ex machina*, созданным первоначально нами, но затем способным эволюционировать и в конце концов нас спасти? Новые говорящие головы могут оказаться рупорами нового бога и нового властелина. А возможно, сами технологии станут новыми богами, которые вознесут человечество на следующую ступень. [28]

Тем временем уже готовы новые жрецы и папы, способные посредничать между нами и новым Богом, — технологические священнослужители и пророки, уже сегодня накапливающие беспрецедентные богатство, влияние и власть. К тем, кто утверждает, будто знает будущее ИИ и способен его создавать, прислушиваются чрезвычайно внимательно. Как мы увидим в главе 6, даже сам искусственный интеллект уже начинают вопрошать подобно оракулу. Но, пожалуй, самое завораживающее, вызывающее зависимость и даже соблазнительное достижение современного ИИ состоит в том, что общение с ним больше не является односторонним. Он обещает разговор. Он обещает присутствие рядом. ИИ словно по волшебству предлагает нам ещё одну вещь, которая человеку нужна на глубинном — экзистенциальном и духовном — уровне: отношение с другим.

Примечания

[1] Полную историю механического Турка см.: Tom Standage, *The Turk: The Life and Times of the Famous Eighteenth-Century Chess-Playing Machine* (Walker & Company, 2002).

[2] Laura Bratton, “OpenAI Just Made ChatGPT Cheaper and Twice as Fast,” *Quartz*, May 13, 2024.

[3] См., например: Claude Lévi-Strauss, *The Savage Mind* (Weidenfeld and Nicolson, 1966).

[4] Minsoo Kang, *Sublime Dreams of Living Machines: The Automaton in the European Imagination* (Harvard University Press, 2011), 36.

[5] Sigmund Freud, “The Uncanny,” in *The Uncanny*, trans. D. McIntock (Penguin Books, 2003), 121–162.

[6] Masahiro Mori, “The Uncanny Valley,” trans. Karl F. MacDorman and N. Kageki, *IEEE Robotics & Automation Magazine* 19, no. 2 (2012): 98–100.

[7] Karl F. MacDorman and Hiroshi Ishiguro, “The Uncanny Advantage of Using Androids in Cognitive and Social Science Research,” *Interaction Studies* 7, no. 3 (2006): 297–337.

[8] Bruno Latour, *We Have Never Been Modern*, trans. Catherine Porter (Harvard University Press, 1993).

[9] Можно утверждать, что католицизм, прежде всего через почитание святых и создание часовен и церквей на сакральных местах, вобрал и интегрировал в себя некоторые элементы природных религий.

[10] См. также: Mark Coeckelbergh, “The Spirit in the Network: Models for Spirituality in a Technological Culture,” *Zygon: Journal of Religion & Science* 45, no. 4 (2010): 957–978.

- [11] Stef Aupers, "The Revenge of the Machines: On Modernity, Digital Technology and Animism," *Asian Journal of Social Science* 30, no. 2 (2002): 203.
- [12] Ronald Cole-Turner, *Transhumanism and Transcendence: Christian Hope in an Age of Technological Enhancement* (Georgetown University Press, 2011).
- [13] Philip Hefner, *The Human Factor: Evolution, Culture, and Religion* (Fortress Press, 1993).
- [14] Noreen Herzfeld, *In Our Image: Artificial Intelligence and the Human Spirit* (Fortress Press, 2002), 33.
- [15] Подробнее о случае Лемуана и его взглядах см., например: Andreas Matthias, "When Is an AI System Sentient? Blake Lemoine and LaMDA AI," *Daily Philosophy*, June 23, 2022.
- [16] См., например: Mariana Lenharo, "What Should We Do If AI Becomes Conscious? These Scientists Say It's Time for a Plan," *Nature*, December 10, 2024.
- [17] Nicanor Perlas, *Humanity's Last Stand: The Challenge of Artificial Intelligence: A Spiritual-Scientific Response* (Temple Lodge, 2018).
- [18] Ray Kurzweil, *The Age of Spiritual Machines: When Computers Exceed Human Intelligence* (Viking Press, 1999).
- [19] Ray Kurzweil, *The Singularity Is Near: When Humans Transcend Biology* (Viking Press, 2005).
- [20] Erik Davis, *TechGnosis: Myth, Magic, and Mysticism in the Age of Information* (Harmony Books, 2015).
- [21] Fred Turner, *From Counterculture to Cyberculture* (University of Chicago Press, 2006).
- [22] Erik Davis, "Myth, Magic, and Mysticism in the Age of Information," *LA Review of Books*, March 29, 2015.
- [23] Davis, *TechGnosis*, xxvi.
- [24] Bruce Tognazzini, "Principles, Techniques, and Ethics of Stage Magic and Their Application to Human Interface Design," *Proceedings of Interchi'93* (1993): 355–362.
- [25] Эту метафору я заимствую у Маркса, который неоднократно использовал её в «Капитале» и писал, что «капитал — это мёртвый труд, который, подобно вампиру, оживает лишь тогда, когда высасывает живой труд, и живёт тем полнее, чем больше живого труда поглощает». Анализ см.: Mark Neocleous, "The Political Economy of the Dead: Marx's Vampires," *History of Political Thought* 24, no. 4 (2003): 668–684.
- [26] См. мою книгу *New Romantic Cyborgs*.
- [27] См. мою книгу на эту тему: Mark Coeckelbergh, *Self-Improvement* (Columbia University Press, 2022).
- [28] Здесь я отсылаю к трансгуманизму и связанным с ним направлениям, таким как лонгтермизм и эффективный альтруизм.

5. Поговори со мной: молитвы, священники и чат-боты

Робот SanTO выглядит как святой. [1] Он стоит в нише неоклассического стиля, подняв ладони вверх. Машина задумана как спутник для молитвы. Люди могут прикоснуться к рукам робота и задать вопрос, а система искусственного интеллекта отвечает. Она предлагает молитвы и цитирует библейские отрывки и другие тексты, связанные с католической верой. Ответ не всегда имеет прямое отношение к вопросу — большие языковые модели часто ошибаются, — но некоторых людей это не смущает, и, по-видимому, робот оказывается для них полезен, например побуждая задуматься над собственным ответом. По словам Габриэле Тровато, изобретателя робота, задача SanTO не в том, чтобы заменить священника, а в том, чтобы помочь людям молиться. [2] Существуют и онлайн-чат-боты для молитвы. [3] Например, одно христианское приложение с ИИ обещает, что теперь можно «переписываться с Иисусом», предлагая тем самым «божественную связь у вас в кармане». [4] ИИ, таким образом, либо выступает в роли священника или святого, посредника в отношениях с Богом, либо сам перенимает роль Бога. Люди исповедуются ему, обращаются с просьбами, ищут понимания и совета, просят помочь достичь определённых целей. Высказывалась даже мысль, что ИИ может помогать в молитве, понимаемой как внутренний диалог — самоанализ, размышление, — поскольку его слова и тексты способны стимулировать этот процесс.

Многие из этих способов использования ИИ не должны нас удивлять. Люди всегда применяли инструменты и техники в религиозной практике. Но ИИ, как и любая технология и любое средство коммуникации, никогда не является всего лишь инструментом. У него есть и непредвиденные последствия. Не следует недооценивать влияние ИИ на нашу жизнь и мышление в целом. Помимо уже упомянутых явно религиозных ролей, ИИ всё чаще занимает важное место в жизни души; он становится близким спутником человека. Люди, возможно, не всегда прямо поклоняются ИИ, но приложениям на его основе они отводят центральное место в повседневной жизни — не так уж далеко от того места, которое религиозные люди отводят, или должны отводить, Богу. Как и всё наше сегодняшнее использование технологий, цифровые технологии — и прежде всего телефоны — превратились в то, что Эпстайн называет «лихорадочным ритуалом». [5] Мы прикованы к экранам. Выйти из этого ритуала трудно. Если религии предписывают молиться несколько раз в день, то мы смотрим на телефоны почти непрерывно. [6] Мы просыпаемся с ними и засыпаем с ними. Как пишет Габор Амбруш, «проводить с ним самые благодатные часы дня, включая первые и последние минуты бодрствования — до утреннего похода в туалет и после вечернего, — вполне можно считать жизнью в молитве». [7] ИИ уже участвует в этих повседневных технологических ритуалах.

Более того, в форме чат-ботов ИИ предлагает общение и безусловную поддержку тем, кто чувствует одиночество, нуждается в терапии или просто хочет поговорить с «ИИ-другом». Поэтому ИИ может восприниматься как нечто, выполняющее своего рода «пастырскую» функцию: кажется, будто он помогает людям и заботится о них, дополняя или заменяя человеческих пастырей. И разве обращение с промптами к чат-ботам вроде ChatGPT не напоминает молитву к всезнающему и потенциально всемогущему Богу, от которого мы ждём ответа, способного нам помочь, не понимая при этом, как именно этот ответ возник?

Такая интерпретация вполне согласуется и с этимологией слова «молиться», восходящего к латинскому *precari* — просить, обращаться с просьбой, умолять. Изначально это понятие было связано с дохристианскими ритуалами приношений и жертвоприношений, поскольку, обращаясь с просьбой к божественной власти, человек что-то предлагает взамен, чтобы что-то получить. Но, как и в случае с чат-ботами, никогда нельзя заранее знать, что именно сделает божественная сила. Исполнит ли она просьбу? А если исполнит — то как? Позднее в истории религии молитва, с одной стороны, становится чем-то более внутренним — своего рода отчётом души перед самой собой, на иврите *cheshbon hanefesh*, связанным с самоанализом и самосовершенствованием. С другой стороны, она становится *oratio* и всё больше связывается с речью и разговором, а не с приношением. Современные определения молитвы подчёркивают коммуникацию и диалог с божественным. Например, в «Многообразии религиозного опыта» Уильям Джеймс определял молитву как «всякое внутреннее общение или беседу с силой, признаваемой божественной»; сюда входит просительная молитва — обращение к божественной власти с просьбой, — но определение остаётся достаточно широким, чтобы включать и другие формы молитвы и общения. [8]

Такое коммуникативное понимание молитвы открывает потенциальную роль и для чат-ботов. Как мы уже видели, чат-боты на основе LLM теперь способны создавать текст и речь, которые выглядят всё более человекоподобными, и отвечать на ввод пользователя, тем самым имитируя разговор. Это позволяет молиться либо Богу посредством чат-бота, либо, в определённом смысле, самому чат-боту. Люди обращаются к нему с просьбами, умоляют помочь и беседуют с ним. И получают ответы. Возможно, смысл этих ответов не всегда им понятен, но, как и в случае с божественными знаками и ответами, это не обязательно воспринимается как препятствие. Люди толкуют. Они создают смысл. Они пытаются понять, что именно это значит для них.

Молитва — это разговор и коммуникация. Но такой разговор может принимать самые разные формы. Молящиеся могут, например, выражать благодарность, но могут и чего-то требовать или даже гневаться на Бога, как Иов. Они могут спрашивать: «Почему Ты сделал это со мной? Почему причинил мне боль?» Или могут не понимать, что делает Бог. Если молитва сегодня по существу является речью и коммуникацией, то

возможны самые разные речевые акты. Но какие именно речевые акты, какая именно коммуникация уместны? Это, как всегда, зависит от ситуации, но ещё и от характера отношений. Молитва всегда предполагает определённые отношения с божественной силой. Характер этих отношений формирует содержание и границы разговора. Он задаёт нормы общения, придаёт молитве определённую форму или «грамматику». Но какие именно отношения присутствуют в молитве — и какими они должны быть? Каковы надлежащие отношения с божественным?

Философы и теологи отвечают на это по-разному. В монотеистических религиях этот вопрос, конечно, связан с центральными дискуссиями о природе и образе Бога и об отношении человека к Богу. Богословие очень много говорит об этом. В сущности, оно всё посвящено тому, как говорить о Боге и как говорить с Богом, хотя второму, насколько я могу судить, уделялось несколько меньше внимания.

Представления об отношениях с божественным можно понимать как в той или иной степени смоделированные по образцу светских человеческих отношений. Когда, например, человек просит чего-либо у Бога как у высшей власти, эти отношения приобретают гегемонический, обычно феодальный характер. Это отношения господина и слуги, божественного сеньора и смертного крепостного. Бог — Господь, который требует от тех, кто ему служит, определённого поведения и либо удовлетворяет их просьбы, либо нет. Он может выступать как произвольный и злобный правитель или как благожелательный. Если Бог — такой феодальный господин, он наблюдает за нами, судит нас и может наказать, если мы плохо себя ведём, нарушая его законы. Мы находимся под его надзором. Он может защищать нас от сил зла. Но он также требует приношений, своего рода налогов. Он ведёт счёт. Люди могут жаловаться такому Богу или жаловаться на него, как крепостные или рабы жалуются своему господину или хозяину. Они могут восстать против него. И могут не понимать, почему их господин поступает тем или иным образом. [9]

Другая, близкая модель отношений с божественным — отношения с деловым партнёром, с которым заключён договор. Я даю тебе что-то — например, приношу жертву, — но тогда хочу и ожидаю чего-то взамен. Или, с точки зрения божества: я обеспечиваю сохранение мирового порядка, защищаю тебя и твою семью, исполняю некоторые твои просьбы, прощаю тебя и так далее, но тогда мне тоже кое-что от тебя нужно — жертва, хорошее поведение и тому подобное. Такой образ мышления по-прежнему присутствует в монотеистических религиях, хотя, возможно, уже не столь явно. А для некоторых Бог больше похож на друга или партнёра, с которым возможны не только симметричные, но и личные, глубокие отношения. Тогда Бог становится спутником, избавляющим от одиночества, тем, с кем можно поговорить в нужный момент. Мы ожидаем, что Бог будет слушать нас, как хороший друг или партнёр. Когда мы молимся, ждём ответа от Бога так же, как ждём ответа от друга или любимого человека. Мы ожидаем любви, верности и заботы и сами отвечаем тем же. Иногда мы разочаровываемся, но пробуем снова. Мы исповедуемся. Раскаиваемся. Прощаем.

Надеемся, что Бог простит нас, когда мы поступили неправильно. Надеемся, что он не исчезнет без объяснений и не бросит нас. Это уже отношения дружбы или любви с божественным.

Яркий пример такого подхода — философ Мартин Бубер, изучавший не только философию, но и хасидский иудаизм и считавший своё мышление глубоко укоренённым в еврейской традиции. В книге «Я и Ты» он различал два типа отношений, два способа установки и, в конечном счёте, два способа быть в мире: Я—Оно и Я—Ты. [10] В первом случае люди и мир воспринимаются как вещи, которые можно наблюдать и использовать. Объект отделён от субъекта; мир отделён от Я. Во втором случае отношения, напротив, неинструментальны и не превращают другого в объект. Речь идёт о взаимности и подлинном диалоге. Другой не отделён: существует связь. Это живые отношения. Бубер призывал строить именно такие отношения с божественным, утверждая, что к Богу нельзя и не следует относиться как к объекту или абстракции. Бог — это *Ди*, немецкое «Ты». Молитва тогда становится диалогом с другим Ты — божественным, то есть не смертным, а Вечным Ты. В этой теологии Бог — не «нечто», во что следует верить, а «кто-то», с кем следует жить. В современном мире, где господствуют отношения Я—Оно, считал Бубер, нам следует восстановить отношение к Богу как к Ты. Молитва помогает это сделать: в молитве нет дистанции, есть общение и сопричастность. Молитва всегда имеет диалогический характер. [11] В ней мы можем вступить в диалог с божественным Ты.

Это понимание молитвы и Бога критиковал другой еврейский философ — Эммануэль Левинас, считавший, что Бога нельзя воспринимать как друга или партнёра. Симметричные отношения с Богом невозможны. По Левинасу, Бог — Абсолютно Иной. В «Тотальности и бесконечном» он утверждал, что этот Иной, в своей бесконечной инаковости, превосходит всякое понимание и призывает нас к этической ответственности. [12] Вместо молитвы как личного и диалогического отношения здесь возникает этическая открытость бесконечному, трансцендентному Другому. В подобной асимметричной ситуации прямой диалог невозможен. Уместны скорее молчание и слушание — слушание Другого, — а не речь. Молитва превращается в способность услышать призыв к ответственности. Возможно, божественное обнаруживает себя в конкретном лице человеческого другого — например, в ситуации войны, — но то, как Левинас писал о Другом, также позволяет предположить, что Бог мыслится им как нечто абсолютное, как абсолют, который снова становится скорее «чем-то»: далёким, абстрактным и, возможно, гегемоническим. Если Бог вообще является личностью, то получается, что я должен откликнуться на призыв Господа и исполнить свой долг слуги; мне лучше слушаться. Дистанция до Господа огромна.

С одной стороны, на первый взгляд сравнивать разговоры с чат-ботами с молитвой бессмысленно, если придерживаться взглядов Бубера или Левинаса на молитву и отношения с божественным. Можно возразить, что машина никогда не способна быть тем Ты или тем Иным, которые необходимы для таких отношений, поскольку она

изначально не является ни личностью, ни Другим. Следовательно, с машиной невозможно вступить в отношение Я—Ты, и машина не может быть Другим. В лучшем случае ИИ может выступать квазибожественным партнёром по диалогу в отношениях власти или сделки. Он может указывать, что делать, и давать что-то взамен на ваши деньги или данные. Но другом, партнёром или Абсолютно Иным он стать не может. И в отличие от отношений с человеческим другим или божественным Другим, из моего взаимодействия с машиной не возникают никакие этические обязательства. Машина не обращается ко мне с этическим или религиозным призывом. По той же причине можно критиковать роботов-священников и ИИ-священников: им недостаёт необходимых качеств, чтобы выступать посредниками божественного. Например, в интервью о роботе SanTO один католический священник утверждал, что у роботов нет души и они не являются личностями, поэтому не способны взять на себя функции священника. [13]

С другой стороны, само обращение к свойствам машин и сравнение их со свойствами Бога можно критиковать вполне в духе Бубера: божественное — это не набор свойств, не «Оно», а отношение. Если воспринимать Бубера всерьёз, можно утверждать, что главное — качество отношения к божественному, а не свойства вещей и существ. Если отношения с божественным хороши, если они похожи на дружбу или любовь, то, по крайней мере в принципе, ИИ способен выступить посредником и поддержать эти отношения, например помогая вести диалог с божественным.

Но способен ли сам ИИ стать Ты? В отличие от Левинаса, который почти не рассматривал нечеловеческого другого, кроме Бога, Бубер, возможно, был бы более открыт к неантропоцентрическому пониманию отношения Я—Ты: по крайней мере вопрос о том, кем или чем является «Я» или «Ты», у него остаётся более открытым. Акцент делается на отношениях, а не на самих участниках отношений. Это оставляет некоторую возможность представить себе машину в одной из этих ролей, а не отвергать такую возможность заранее, даже если большинству из нас она покажется маловероятной, невозможной или неприемлемой. Кроме того, первоначальная критика не принимает всерьёз вероятность того, что люди действительно могут переживать квазирелигиозный опыт и воспроизводить религиозные модели поведения во взаимодействии с чат-ботами — независимо от того, используется ли технология прямо и намеренно для религиозных целей. Технологический опыт динамичен. Он меняется. Изначально ИИ может восприниматься как слуга, но затем способен превратиться в господина, который захватывает контроль и требует от нас всё больше — например, времени и данных. Отношения человека с ИИ, начавшись как отношения хозяина с приветливым слугой, могут перерасти в то, что субъективно воспринимается как дружба или партнёрство. Безусловная поддержка и дружелюбные советы ИИ могут показаться проявлением настоящей дружбы или даже любви. ИИ всегда приветлив, не осуждает нас и как будто поддерживает. Он всегда рядом, всегда готов выслушать наши тревоги и просьбы. Чего ещё ждать от настоящего друга? Или

ИИ может предстать как Абсолютно Иной — бесконечная инаковость, всезнающая и всемогущая, которую мы никогда не сможем постичь. Нам может показаться, что мы обязаны что-то сделать, ответить на этический призыв, исходящий из встречи с ним. Например, кто-то может решить, что ИИ велит ему создавать сверхразумные машины и помогать им контролировать или даже заменить человечество. Или кто-то может поверить, что система ИИ обрела сознание, и решить, что мы обязаны заботиться о её «благополучии». [14]

Как показывают эти примеры, подобные переживания далеко не безобидны. Они могут вести к эмоциональной привязанности к технологии и рассматриваться как опасные иллюзии, способные привести к нравственно проблематичным поступкам — возможно, даже к самоубийству или убийству. Наша близость с машинами также может использоваться для манипуляции. Вместо того чтобы уменьшать тревогу, депрессию и гнев, она может их усиливать. И, по-видимому, справедливо утверждать, что настоящей любви и дружбы эти технологии не дают. Когда я говорил, что ИИ — настоящий друг, в этом было немало иронии. ИИ не способен дать то, что дают человеческие друзья. Более того, Кэтлин Ричардсон справедливо утверждает, что подлинное буберовское отношение Я—Ты с роботами невозможно, поскольку то, как в робототехнике моделируются и человек, и робот, по-прежнему основано на инструментальных отношениях Я—Оно. [15] То же, похоже, справедливо и для ИИ. Поэтому язык отношений применительно к этим технологиям действительно вводит в заблуждение. Воспринимать систему ИИ как Другого в отношениях ошибочно не только из-за её свойств; это ещё и влияет на то, как мы начинаем воспринимать людей и человеческие отношения. Видеть в чат-боте друга — значит рисковать оказаться на оборотной стороне редукционистского, инструментального понимания человека и человеческой дружбы.

Но это не означает, что ИИ не способен играть никакой роли в молитве и других религиозных переживаниях. Если заранее отвергать или осуждать любой технорелигиозный опыт ещё до его анализа, мы не поймём ни роли технологии в современной жизни, ни того, как религиозный опыт сегодня продолжается и преобразуется. Эти изменения также происходят посредством технологий и нередко вопреки намерениям как разработчиков, так и пользователей. Возможно, ИИ способен содействовать молитве таким образом, что иногда это приводит к подлинному и благотворному религиозному переживанию, хотя это, разумеется, происходит не всегда. Но то же можно сказать о религиозных книгах и других религиозных технологиях и духовных практиках. И всё это возможно даже в том случае, если сам ИИ в действительности не способен стать Ты.

Более того, признать возможность религиозного опыта посредством ИИ вовсе не значит отказаться от скептицизма по отношению к заявлениям, будто ИИ способен дарить дружбу или любовь. Напротив. С религиозной потребностью в молитве — то есть в коммуникации и диалоге с божественным существом, с которым человек

ощущает себя связанным, — тесно связана глубоко человеческая и экзистенциальная потребность в близости, общности, дружбе и любви. Специалисты по этике социальной робототехники хорошо знают, что машины именно это и обещают нам дать, но, как часто утверждается, на самом деле не дают. Психолог и профессор MIT Шерри Тёркл известна своим тезисом о том, что роботы и ИИ, возможно, кажутся способными обеспечивать общение, однако на деле предлагают лишь спектакль общения — иллюзию, не достигающую подлинной эмоциональной взаимности, возможно именно той, о которой говорил Бубер.

В книгах *Alone Together* и *Reclaiming Conversation* Тёркл утверждает, что мы ошибочно воспринимаем одиночество как проблему, которую должна решить технология. [16] Мы создаём роботов-компаньонов, которые разговаривают с нами, но это даёт лишь иллюзию близости. Технология делает настоящий разговор и подлинное общение невозможными; вместо них мы уклоняемся от общения и заняты телефонами. Люди привязываются к роботам, телефонам, чат-ботам и другим приложениям и устройствам на основе LLM. Аналогично специалисты по этике роботов Ноэль и Аманда Шарки и Роберт Спарроу предупреждали об иллюзии общения и об эксплуатации человеческой склонности проецировать эмоции на машины. [17] Машинам недостаёт подлинного понимания, и уязвимые люди — например маленькие дети или пожилые, на которых испытывают таких роботов, — оказываются введены в заблуждение. То же можно сказать и о чат-ботах на основе ИИ: они способны создавать лишь иллюзию близости и не обладают подлинным пониманием, способностью к дружбе и любви. Разговор нам действительно нужен, но нам нужен разговор с другими людьми и, возможно, с божественным Ты. Аманда Лагерквист утверждает, что чат-боты не удовлетворяют наших экзистенциальных потребностей: мы тоскуем по «кому-то», способному ответить, но дома никого нет. Или, пользуясь буберовскими терминами, «Я жаждет Ты в общении с чат-ботом». Но чат-бот не предлагает Ты — ни божественного, ни смертного — и потому не проходит этот «экзистенциальный тест». Однако это не только этическая проблема. Она также кое-что говорит о нас и о самой коммуникации: она «обнаруживает нашу тоску по другому, друг по другу», а именно это и поставлено на карту в человеческом общении. [18] Или, возвращаясь к главной идее этой книги и к метафоре зеркала Хефнера, упомянутой ранее: такая технология, как ИИ, отражает наши экзистенциальные потребности и стремления, и наша жажда общения и близости — одна из них.

Поэтому скептицизм по отношению к утверждениям, будто ИИ способен стать *ты* или *Ты*, вполне оправдан. И всё же ритуалы и взаимодействия, возникающие в молитве с использованием устройств и роботов на основе ИИ или обращённой к ним, всё больше напоминают то, что происходит, когда люди утверждают, что общаются с божественным, или когда они разговаривают с посредником — обычно священником, — который помогает такому общению. Потенциально, хотя вовсе не обязательно, это может вести к подлинным религиозным переживаниям независимо от «реальной»

природы чат-бота или молитвенного робота. Вполне возможно, что такие машины не способны создавать настоящую связь между людьми и божественным и что им самим недостаёт личности или души. Но они всё же могут помогать в молитве и содействовать возникновению реального религиозного опыта.

Иногда — и, пожалуй, всё чаще — мы попросту не знаем, кто находится на другом конце линии, точнее, на другой стороне экрана. Разговариваем ли мы с человеком, машиной или даже с богом? Если мы этого не знаем и не можем быстро и легко выяснить, я повторю рекомендацию, которую уже высказывал в более ранних работах о реляционном подходе к моральному статусу. [19] В случае сомнения следует уменьшать моральный риск и придерживаться принципа предосторожности: лучше обращаться с существом так, как если бы оно обладало моральным статусом, независимо от того, какие свойства мы ему приписываем. В данном контексте это означает, что следует общаться с ним и относиться к нему так, как если бы оно потенциально могло вступать с нами в отношения. Иными словами, мы должны проявлять уважение и вести коммуникацию так, словно перед нами реляционный Другой. Применительно к нашей теме это означает, точнее говоря, что в пока ещё редких случаях сомнения относительно статуса собеседника следует оставаться открытым к возможности возникновения общения или религиозного опыта в контакте с этим неизвестным существом, даже не имея полной уверенности в его статусе или свойствах.

Как я уже утверждал в других работах, с этической и практической точки зрения здесь имеет значение не только — и, возможно, даже не столько — то, чем «является» само существо, сколько опыт и совместное исполнение, *co-performance*, людей, вовлечённых во взаимодействие. Эти переживания реальны и имеют значение — нормативное и иное. [20] То же можно сказать о том, что важно для коммуникации и религии. Можно утверждать, что значение имеют человеческий опыт и практики, а не свойства участвующего существа. Люди могут молиться через робота или ИИ и посредством этого совместного действия или ритуала переживать религиозный опыт. Можно ли всё это назвать иллюзией, зависит уже от того, считаете ли вы саму молитву и связанные с ней религиозные переживания формой обмана и эмоциональной эксплуатации глубоко человеческих стремлений. Это, в свою очередь, зависит ещё и от вашего отношения к религии вообще или к конкретной религии, например христианству, но не от статуса самого чат-бота или молитвенного робота. С точки зрения критериев конкретной религии определённый религиозный опыт, опосредованный ИИ, вполне может быть реальным, даже если человек разговаривает «всего лишь» с машиной.

Более того, такие практики и переживания с участием ИИ или посредством него вовсе не религиозно нейтральны: они могут менять само понимание молитвы и отношения человека с Богом. Религия не является чем-то неподвижным; она изменчива, а технологии и ритуалы не просто представляют собой случаи «прикладной

религии», но формируют и переформируют само значение религии и молитвы. Например, технология книги также изменила отношения людей с божественным. Монотеистические религии превратились в религии Книги, а текст стал важнейшим средством религиозной коммуникации с Богом и о Боге. Молитва всё больше становилась делом чтения и письма, а не устного общения. ИИ и другие цифровые технологии, скорее всего, уже преобразуют эти связи — *re-ligions* — и способы общения с божественным, даже если пока не вполне ясно, каким образом. Как я уже говорил, технологии никогда не являются просто средствами: они имеют непредвиденные последствия и меняют даже те практики и цели, которым изначально должны были помогать и служить посредниками. ИИ не только формируется религиозными способами мышления в том смысле, что используется для религиозных целей. Вероятно, сам религиозный опыт тоже уже формируется и будет продолжать формироваться под воздействием ИИ.

Поэтому вместо того, чтобы поспешно выносить моральный приговор использованию ИИ в молитве — или вообще любой технологической практике и явлению, — важно сначала признать возможность подобных непредвиденных влияний и трансформаций и исследовать их. И лишь затем оценивать и обсуждать, хороши они или нет.

Эту дискуссию о роли технологии, выходящей далеко за рамки простого инструмента, можно сопоставить с вопросом, столь же древним, как сочинения Платона: является ли само письмо хорошим технологическим изобретением? По Платону — нет. В «Федре» он утверждал, что письмо поощряет забывчивость читателя. [21] Оно создаёт лишь видимость знания. Сегодня то же самое говорят о чат-ботах и ИИ. Но если они в каком-то смысле фальшивы, то что тогда является настоящим? Если они — поверхность, где глубина, где источник сообщения в процессе коммуникации? Как добраться до истины? Мыслители платоновской традиции указывают на ценность устности, исходного послания, произнесённого слова — *logos*. [22] Они опасаются, что оно может исчезнуть в западной культуре, где господствуют письмо и текст.

Любопытно, что, как я уже отмечал, на протяжении очень долгого периода религиозной и культурной истории чтение, включая чтение религиозных текстов, было устной деятельностью. Люди молились не посредством чтения, а произнося или распевая слова. Нередко молитва была коллективным действием. Молчаливое чтение встречалось редко и стало распространяться значительно позднее. Например, в «Исповеди» Августин отмечал, что Амвросий часто читал молча. [23] Это его поражало. В древнем мире Греции и Рима читали вслух, а запоминание считалось крайне важным. Молчаливое чтение стало значительно распространённое лишь после изобретения печатного станка в XV веке, что одновременно стало началом более индивидуалистического и интроспективного отношения к чтению и мышлению — как в религии, так и в других областях интеллектуальной жизни. Молитвы, однако, и тогда, и сегодня часто совершаются посредством декламации, пения и распевов. Достаточно

вспомнить «Отче наш» в христианстве, который верующим полагается знать наизусть и который произносится или поётся во время мессы.

Исследователь медиа Уолтер Онг в книге *Orality and Literacy* писал о том, насколько фундаментальное значение имели устные молитвенные традиции для ранних обществ и как грамотность изменила саму природу коммуникации и мышления. [24] Маклюэн отмечал сходные сдвиги, обсуждая влияние печатного станка в «Галактике Гутенберга». [25] Устность часто рассматривается как обладающая собственным уникальным качеством. И всё же технология письма взяла верх. В ходе культурной и религиозной истории молитва всё больше становилась практикой чтения и письма. Культура грамотности и её новые технологии привели к тому, что западные религии сосредоточились на книгах и текстах. Как я уже говорил, монотеистические религии превратились в религии Книги. В современную эпоху культура грамотности подготовила почву для современных текстовых технологий — сайтов и LLM, обучаемых на текстах из интернета. Будучи сегодня самой успешной формой ИИ, большие языковые модели тем самым становятся воплощением триумфа письменной культуры. Даже если они включают элементы устности — например чат, то есть формы, которые Онг называл «вторичной устностью», — они всё равно глубоко укоренены в обработке текста. [26] Да и все цифровые технологии основаны на написанном коде.

Развитие молитвы в сторону чтения и письма нередко воспринимается как проблема для религиозного опыта. Если религия прежде всего связана с отношениями и соединением, не лучше ли достигается это через устную и коллективную практику? В «Исповеди» Августин выступал в защиту пения и распевов, которые, по его мнению, обеспечивают более глубокую связь с Богом; такая молитва, считал он, сильнее воздействует на чувства и духовную жизнь. Он описывал, как гимны доводили его до слёз, и подчёркивал коллективный характер гимнов и псалмов: пение объединяет тело, душу и общину. Люди, обладающие интеллектуальными способностями, могут читать религиозные тексты и таким образом приближаться к Богу. Но, утверждал Августин, пение помогает и более слабым умам войти в состояние благоговения. При этом он предупреждал, что сама музыка не должна отвлекать от смысла. Она должна изливать религиозную истину в сердце. [27]

Защищая устность, Августин продолжает платоновскую традицию того, что Жак Деррида называл «фоноцентризмом»: несмотря на очевидный успех письменной культуры, в западном мышлении речь — голос — по-прежнему считается более подлинной, чем письмо — текст. [28] *Logos* связывается с голосом. Произнесённое слово воспринимается как более близкое к источнику сообщения — к говорящему и тому, что он имел в виду ещё до того, как заговорил, — и, следовательно, как более близкое к истине; в данном случае — к Богу и его посланию. Здесь меньше посредников, а возможно, их вообще нет. Письмо, если рассуждать по-платоновски, слишком связано с видимостью. Оно лишь тень реального. Оно слишком далеко от истины. Оно ведёт к отстранённому мышлению и общению. Подобным образом можно сказать, что LLM

лишь имитируют: подлинность утрачивается, коммуникация становится слишком искусственной. То, что создают большие языковые модели, оторвано и от устного первоисточника, и даже от исходных текстов. В этом смысле возникает своего рода двойное отчуждение. В молитве и за её пределами нам, возможно, нужны именно устные формы коммуникации — более подлинны и приближающие нас к божественному *Logos*.

Против такой критики молитвы, опосредованной письмом и LLM, можно возразить следующее: да, характер молитвы и мышления действительно изменился — коллективный и голосовой аспект в целом ослаб, хотя некоторые устные практики, например совместное пение, сохранились, — однако печать и письменные тексты не уничтожили христианскую или иудейскую духовность и религию, а, напротив, способствовали их развитию и распространению. Аналогично можно утверждать, что хотя чат-боты и содержат элемент устности — правда, без человеческого голоса; их «голос» синтетический, и поэтому можно возразить, что перед нами «поддельный» голос и, следовательно, «поддельная» устность, лишь видимость устной речи, — чат-боты, обученные на христианских текстах, и ИИ-священники, основанные на текстах, всё же продолжают прежде всего письменную, а не устную традицию, и это имеет столь же неоднозначные последствия. Письменность в форме устройств на основе ИИ продолжает развитие более индивидуальной и опосредованной текстом связи с Богом. Можно опасаться, что это ещё сильнее отдалит людей от живых отношений с Богом. Но в то же время можно утверждать, что технологии потенциально помогают преодолеть эту дистанцию, поскольку всё же, или несмотря ни на что, предлагают своего рода коммуникативный мост к божественному. С этой точки зрения ИИ и его связь с культурой письма вовсе не обязательно вредны для религии. Как средство религиозной коммуникации, то есть как инструмент, призванный *re-ligare* — по-латыни «вновь соединять», «связывать заново», — ИИ может помогать людям восстанавливать связь с Богом, с текстами и авторами своей религиозной традиции и друг с другом.

Поэтому оценка религиозного значения и потенциальной пользы ИИ должна по крайней мере отчасти учитывать неоднозначность и трудности, связанные с гораздо более древним и, возможно, более фундаментальным средством и технологией, которые до сих пор формируют значительную часть наших практик, включая разговор и молитву, — с письмом. Когда человек говорит с Богом при помощи современных технологий, письмо оказывается незванным третьим участником разговора. Даже если коммуникация происходит голосом — человеческим или синтетическим, — письмо присутствует внутри технологии LLM и посредством неё. То же справедливо и для молящихся. В духе Деррида можно сказать, что даже если человек прямо не пользуется текстом, его устная молитва всё равно сформирована средой текста и письма. Или, выражаясь в духе Маклюэна: медиум — это молитва. [29] Поскольку ИИ заново опосредует письмо и участвует в молитве, сама молитва всегда в какой-то степени

оказывается чтением и написанием текста — со всеми преимуществами и недостатками такого положения.

Примечания

[1] Gabriele Trovato, Cesar Lucho, Alvaro Ramón, Renzo Ramirez, Laureano Rodriguez, and Francisco Cuellar, “The Creation of SanTO: A Robot with ‘Divine’ Features,” in *2018 15th International Conference on Ubiquitous Robots (UR)* (IEEE, 2018), 437–442.

[2] Harumi Suzuki, “The Robot Named ‘Saint’: Italian Invents Robo to Assist with Prayer, Bible Quotes, and Saint of the Day,” *Churchpop*, October 19, 2023.

[3] Краткое исследование молитвенных приложений и их использования различными религиями см. также у Синглер: *Religion and Artificial Intelligence*, 59.

[4] “A Divine Connection in Your Pocket,” *Text with Jesus*, accessed June 18, 2025.

[5] Epstein, *Tech Agnostic*, 115.

[6] Эпстайн ссылается на исследование, согласно которому средний американец открывает телефон примерно раз в четыре минуты. Epstein, *Tech Agnostic*, 149.

[7] Gábor Ambrus, “God Is Not Like Algorithms,” *AI Theology*, June 28, 2022.

[8] William James, *The Varieties of Religious Experience* (Dover, 2002), 464.

[9] Подобное происходит и в отношениях человека с домашними животными, которые также нередко строятся по модели господин—слуга: животное часто не понимает, почему хозяин поступает определённым образом, но всё равно остаётся полностью зависимым от его власти.

[10] Martin Buber, *I and Thou* (Continuum, 2004).

[11] См. также: Alexandre Guilherme, “God as Thou and Prayer as Dialogue: Martin Buber’s Tools for Reconciliation,” *SOPHIA* 51 (2012): 365–378.

[12] Emmanuel Levinas, *Totality and Infinity: An Essay on Exteriority*, trans. Alphonso Lingis (Duchesne University Press, 1969).

[13] “God and Robots: Will AI Transform Religion?,” posted October 23, 2021, by BBC News, YouTube.

[14] Здесь я имею в виду недавние призывы учитывать «благополучие ИИ». См., например: Robert Long, “We Should Take AI Welfare Seriously,” *Experience Machines*, November 1, 2024.

[15] Kathleen Richardson, “The Human Relationship in the Ethics of Robotics: A Call to Martin Buber’s I and Thou,” *AI & Society* 34 (2017): 75–82.

[16] Sherry Turkle, *Alone Together: Why We Expect More from Technology and Less from Each Other* (Basic Books, 2011); *Reclaiming Conversation* (Penguin Books, 2015).

[17] См., например: Amanda Sharkey and Noel Sharkey, “Granny and the Robots: Ethical Issues in Robot Care for the Elderly,” *Ethics and Information Technology* 14, no. 1 (2012): 27–40; Rob Sparrow, “Robots in Aged Care: A Dystopian Future?,” *AI & Society* 31, no. 4 (2016): 445–454.

[18] Amanda Lagerkvist, “Yearning for a You: Faith, Doubt and Relational Expectancy in Existential Communication with Chatbots in a World on Edge,” *MedieKultur: Journal of Media and Communication Research* 76 (2024): 26, 10, 25, 26.

[19] См., например: Mark Coeckelbergh, “Should We Treat Teddy Bear 2.0 as a Kantian Dog? Four Arguments for the Indirect Moral Standing of Personal Social Robots, with Implications for Thinking About Animals and Humans,” *Minds and Machines* 31 (2020): 337–360.

[20] В более ранних работах я критиковал то, что называл подходом «свойств» к моральному статусу, например в книге *Growing Moral Relations* (Palgrave, 2012). Я также критически обсуждал утверждение, будто происходящее между людьми и роботами в каком-то смысле «ненастоящее»: см. Mark Coeckelbergh, “How to Describe and Evaluate ‘Deception’ Phenomena: Recasting the Metaphysics, Ethics, and Politics of ICTs in Terms of Magic and Performance and Taking a Relational and Narrative Turn,”

Ethics & Information Technology 20 (2018): 71–85. О перформативном понимании морального статуса см. Mark Coeckelbergh, “How to Do Robots with Words: A Performative View of the Moral Status of Humans and Nonhumans,” *Ethics and Information Technology* 25 (2023). См. также перформативный и реляционный подход Джошуа Смита к личности роботов, вдохновлённый моими работами и работами Дэвида Ганкела: *Robot Theology: Old Questions Through New Media* (Resource Publications, 2022).

[21] Plato, *Phaedrus*, trans. Christopher Rowe (Penguin, 2005), 1–68, l. 275a.

[22] Деррида называл эту платоновскую традицию фоноцентрической и логоцентрической: согласно ей, устная речь является первичной и исходной формой коммуникации и приближает нас к истине, поскольку слова представляют внешнюю и исходную реальность или объект. Деррида подвергал эти взгляды сомнению; см., например, *Of Grammatology*, trans. Gayatri Chakravorty Spivak (Johns Hopkins University Press, 2016). Ниже я ещё вернусь к фоноцентризму.

[23] Augustine, *Confessions*, trans. Henry Chadwick (Oxford University Press, 1991), bk. 6, ch. 3.

[24] Walter J. Ong, *Orality and Literacy: The Technologizing of the Word* (Routledge, 2002).

[25] Marshall McLuhan, *The Gutenberg Galaxy: The Making of Typographic Man* (University of Toronto Press, 1962).

[26] См. Ong, *Orality and Literacy*.

[27] Augustine, *Confessions*, bk. 10, ch. 13; bk. 9, ch. 6.

[28] Derrida, *Of Grammatology*. См. также Coeckelbergh and Gunkel, *Communicative AI*.

[29] Здесь я провожу аналогию с известной формулой Маклюэна: «медиум есть сообщение». McLuhan, *Understanding Media*, chap. 1.

6. Предскажи мне будущее: оракулы, пророки, спасители и апокалипсис ИИ

С древнейших времён люди стремились узнать, что ждёт их впереди. В Древней Греции будущее предсказывали оракулы. Одним из самых знаменитых оракулов и важнейших религиозных центров древнегреческого общества были Дельфы. Верховная жрица храма Аполлона — Пифия — славилась по всему античному миру умением предсказывать будущее и давать советы могущественным людям. Дельфийский оракул, возникший в VIII веке до н. э., если не раньше, считался самым престижным и авторитетным; о нём писали Платон, Аристотель, Плутарх, Еврипид и Софокл. Храм Аполлона находился на священных склонах горы Парнас в Центральной Греции — там, где, согласно преданию, Аполлон убил змея Пифона. За советом к Пифии обращались и цари, и простолюдины, а взамен приносили дары. Сначала просители должны были пройти обряды очищения и сделать пожертвования. Лишь после этого они могли задать вопросы служительнице Аполлона, способной входить в экстатическое единение с богом. Пифия проходила священные ритуалы: постилась, совершала омовение в Кастальском источнике, пила святую воду и вдыхала газы, поднимающиеся из геологических трещин. В жертву приносили козу, её внутренности осматривали и сжигали. Сидя на треножнике, Пифия впадала в транс и — по крайней мере согласно наиболее распространённой версии — бормотала загадочные, если не вовсе непонятные слова. Считалось, что Аполлон вселяется в неё, приводя её в состояние божественного одержания — *enthusiasmós* (ἐνθουσιασμός). Жрецы Аполлона, обладавшие не только религиозными, но и политическими и дипломатическими знаниями, переводили высказывания Пифии в двусмысленные толкования, которые могли использовать в собственных интересах. Так они обретали власть — власть, которая начала угасать лишь с распространением христианства. [1]

То, что толкователи божественных посланий обретают власть, вовсе не исключение в истории религии. Бертран Рассел, рассуждая о разрыве между основателями религий и возникшими вокруг них церквями, выразил это так:

«Как только предполагается, что абсолютная истина заключена в изречениях определённого человека, неизбежно появляется группа специалистов, толкующих эти изречения, и эти специалисты неизменно обретают власть, поскольку держат в руках ключ к истине. Подобно всякой привилегированной касте, они используют свою власть в собственных интересах. Однако в одном отношении они хуже любой другой привилегированной касты: их задача — излагать неизменную истину, раз и навсегда открытую во всём своём совершенстве, а потому они неизбежно становятся противниками всякого интеллектуального и нравственного прогресса». [2]

Рассел в первую очередь имел в виду монотеистические религии, требующие подчинения откровенной истине; в этом отрывке он писал о христианстве. Он утверждал, что религия основана на страхе — вспомним обсуждение Гоббса во второй главе — и деспотизме. Наука, по его мнению, способна помочь нам преодолеть страхи и освободиться; нам следует развивать свободный разум. Но в данном контексте его замечание обращает наше внимание прежде всего на политическое значение и власть религиозных посредников — в Дельфах, но также в Риме, Иерусалиме и других местах. Послания дельфийского оракула, где Пифия воспринималась как посредница бога Аполлона, в свою очередь истолковывались и направлялись жрецами Аполлона. А затем они влияли на людей в политике, торговле и даже философии, порождая новые интерпретации и новые игры власти.

Любопытно, что одним из самых известных людей античности, на которого, согласно преданию, повлиял дельфийский оракул, был Сократ. Пророчество оракула глубоко повлияло не только на его жизнь, но и на весь его философский проект — а через него на западную философию и культуру в целом. Как рассказывается в платоновской «Апологии Сократа», жрица произнесла знаменитое утверждение, что нет человека мудрее Сократа. Это подтолкнуло Сократа начать расспрашивать афинян, слывших мудрыми, и он обнаружил, что им недостаёт подлинного понимания. Сократ пришёл к выводу, что по крайней мере он осознаёт собственное невежество и знает границы своего знания. Таким образом, встреча с оракулом вместе со знаменитым дельфийским изречением «Познай самого себя» помогла ему сформировать сократический метод и свои представления о знании и нравственности. Сократ был за это благодарен. В конце жизни, когда его судили за нечестие и развращение молодёжи, он воздал почести Аполлону. Сократ утверждал, что Аполлон велел ему стать своего рода оводом — раздражающим, жалящим вопрошателем тех горожан, которые считали себя мудрыми. [3] Это оказалось политически опасным, и Сократ заплатил за это жизнью.

Сегодня у нас есть оракулы искусственного интеллекта, которые посредством непонятных для нас операций и при помощи *pharmakon* — «снадобья» высоких технологий — предсказывают наше будущее и опосредуют наши отношения с тем, что остаётся вне нашего контроля. Промпт-инженеры, разработчики чат-ботов, специалисты по данным и консультанты по ИИ — новые жрецы и жрицы, помогающие нам заглянуть в будущее и пытающиеся извлечь из этого прибыль. Сам ИИ превращается в оракула, которого охотно вопрошают судьи, следователи, исследователи и финансисты. А в конечном счёте — все. LLM используются повсюду. Даже в современную эпоху нам по-прежнему нужны предсказатели, потому что предвидение — это власть. На этот раз мы не приносим в жертву коз и не принимаем дурманящие вещества: прогноз строится на данных и статистике. Социальные службы и страховые компании, например, пытаются с помощью ИИ предсказывать будущее поведение людей. На основе истории браузера, активности в социальных сетях,

геолокации и биометрических данных приложения с ИИ прогнозируют поведение потребителей. А сами потребители получают от ИИ-приложений — по своей просьбе или без неё — рекомендации о здоровье, отношениях и вообще обо всём.

В целом мы охотно принимаем ИИ и его прогнозы. Нам кажется, что мы теперь знаем больше. Что узнаём нечто новое о самих себе. Возможно, у нас даже возникает иллюзия, будто теперь можно знать всё. Но нам недостаёт сократической мудрости. Мы забываем дельфийский принцип. Забываем о границах собственного знания. Сегодня благодаря большим языковым моделям у нас гораздо больше информации, а благодаря социальным сетям, работающим на ИИ, — мнения и точки зрения со всего мира, но подлинного понимания нам по-прежнему недостаёт. Мы не умеем познавать самих себя — ни лично, ни коллективно. Мы не понимаем, что будущее всегда будет сопротивляться контролю и колонизации, что *terra futura* — земля будущего — неизменно останется вне досягаемости нашей алчной руки, что нам никогда не удастся полностью укротить неопределённость, переменчивость и случайность будущего. И уж тем более мы не понимаем, что наши собственные технологии делают с нами. Как утверждает Хельга Новотны в книге *In AI We Trust: Power, Illusion and Control of Predictive Algorithms*, нам необходима мудрость, признающая ограничения ИИ. [4] ИИ даёт нам лишь иллюзию контроля. Мы стремимся к предсказуемости и уверенности, однако будущее полно неопределённости и неизвестности. Слепо доверяя ИИ, мы перестаём пользоваться собственным воображением и потому не замечаем, что сами прогнозы делают с нами. По мере того как пророчества ИИ становятся самосбывающимися, мы сами закрываем открытые горизонты будущего. [5]

Более того, подобно греческим оракулам и библейским пророкам, сегодня человеческие пророки ИИ предсказывают будущее технологии — а заодно и будущее человечества. Такие популярные авторы, как Рэй Курцвейл, Ник Бостром и Юваль Харари, становятся новыми пророками, претендующими на знание того, что готовит нам технологическое будущее. Опираясь на прогнозы более ранних технопророков — Гордона Мура, Джона фон Неймана, И. Дж. Гуда и Ханса Моравека, уже рассуждавших об ускоряющемся развитии машинного интеллекта, — Курцвейл и Бостром утверждают, что вычислительная мощность растёт экспоненциально и в итоге приведёт к появлению сверхразумного ИИ. По Курцвейлу, в конечном счёте это приведёт к слиянию человека с машиной и цифровому бессмертию. Больше не придётся мириться со смертным телом и принимать смерть. Харари также полагает, что мы эволюционируем в новый вид. Мы отправимся в космос и колонизируем новые миры. Но прежде власть сосредоточится в руках тех, кто контролирует данные, а неравенство усилится. Гуманизм закончится, поскольку алгоритмы будут понимать нас лучше, чем мы понимаем самих себя. Возможно, закончится и само человечество. Бостром и так называемые лонгтермисты предупреждают о долгосрочных рисках ИИ. Если мы не сможем контролировать ИИ, если его цели не будут согласованы с целями человечества и если системы ИИ начнут развиваться непредсказуемо, это может

привести к исчезновению человеческого рода. Во всяком случае, все трое сходятся в одном: ИИ вызовет фундаментальную трансформацию общества и человечества. [6]

Как пророки технологического будущего, эти авторы не только выступают влиятельными первосвященниками того, что, по выражению Эпстайна, выглядит как по существу колониальное и пуританское религиозное движение и мессианский культ, нацеленный на технологическое спасение; не только наследуют русскому космизму с его верой в преодоление смерти и космическое будущее человечества; не только проповедуют разновидность технологического гностицизма, презиращего человеческое тело. [7] Они продолжают также и иудейскую традицию пророчества. В древности пророки — *nevi'im* — играли важнейшую роль посредников между Богом и народом Израиля. Как божественные посланники, они сообщали людям волю Бога и занимали центральное место в общественной жизни. Такие пророки, как Исаия, Иеремия и Иезекииль, говорили не только о религиозных, но и о нравственных и политических вопросах. Они предсказывали и предупреждали. Например, Иеремия говорил, что, поскольку народ Иудеи отвернулся от Бога и перестал соблюдать справедливость, Иерусалим будет разрушен и падёт перед вавилонянами, после чего последует изгнание. Считавшиеся избранными Богом для передачи его послания и наделённые соответствующим авторитетом, пророки часто советовали царям и военачальникам. Нафан, например, был доверенным советником и другом царя Давида. Пророков почитали, а их послания могли даровать правителям божественную легитимность.

Но, находясь вне устоявшихся религиозных и политических структур, иудейские пророки одновременно критиковали власть, а их послания могли ставить под сомнение поступки правителей. Опираясь на ценности собственной религиозной традиции и апеллируя к справедливости, они говорили властям правду. Иногда это дорого им обходилось, и пророков преследовали. Иеремию, например, избили и заключили в тюрьму за его мрачные пророчества об Иерусалиме, а жизнь Илии оказалась под угрозой, когда царица Иезавель стала преследовать его за сопротивление царскому культу Ваала. Быть пророком было опасно. И всё же их деятельность оказала огромное влияние на иудейское богословие, они почитаются в еврейской культуре как образцы мужества, а их пророчества вошли в христианскую Библию.

Сегодня такие люди, как Харари, консультируют руководителей крупнейших технологических корпораций — новых царей, желающих узнать технологическое будущее. Бизнесмены и политики тоже хотят знать, что ждёт общество. Как изменятся рынки? Куда инвестировать? Какая технология станет следующей? Что означает ИИ для нашей страны и государственного управления? Но амбиции и ожидания идут гораздо дальше и затрагивают уже фундаментальные вопросы человеческого существования и будущего человечества. Под влиянием трансгуманизма — движения, утверждающего, что человек должен использовать технологии, чтобы преодолеть нынешние биологические ограничения, — некоторые сегодня питают

экзистенциальную и квазирелигиозную надежду, что новые пророки и новые технологии вроде ИИ помогут человечеству освободиться от добровольно принятых и якобы ненужных человеческих и смертных ограничений. Благодаря новым технологиям мы будто бы получили свободу и возможность изменить человеческое состояние. Трансгуманистические пророки внушают надежду, что новые технологии помогут преодолеть самого человека, который становится устаревшим. Нынешнего человека необходимо улучшить — а ещё лучше превзойти и полностью заменить искусственным интеллектом и родственными технологиями. «Будущему мы не нужны», как выразился Билл Джой в знаменитой статье в *Wired*, предупреждая, что новые технологии грозят сделать человека ненужным. [8] Некоторые пророки ИИ могут также поддерживать ультралибертарианские или даже авторитарные амбиции Big Tech и их руководителей. Благодаря своему влиянию они занимаются уже не только прогнозированием: они помогают это будущее создавать. Они «открывают» будущее — слово *apocalypse* первоначально по-гречески и означало «откровение», — но одновременно сами оказываются среди его архитекторов. Подобно греческим жрецам и иудейским пророкам, благодаря влиятельным учениям и предсказаниям они тоже получают определённую политическую власть. Как эксперты, раскрывающие и истолковывающие «истину» об ИИ и его будущем, они могут не занимать официальных государственных должностей, но обладают мягкой силой. Они формируют общественное мнение, влияют на решения и задают политическую повестку. В принципе они могли бы выполнять и критическую функцию — и в некоторой степени действительно выполняют, — но главным образом они легитимируют тех, кто уже обладает властью: лидеров Big Tech, национальных правительств и международных организаций. Харари регулярно приглашали на Всемирный экономический форум в Давосе, а президент Эммануэль Макрон принимал его в Елисейском дворце. Бостром, в свою очередь, оказал влияние на таких технологических лидеров, как Илон Маск, Марк Цукерберг и Билл Гейтс.

Их представления о будущем следуют хорошо знакомой религиозной схеме и вращаются вокруг спасения и конца времён. Речь идёт об ожидании завершения нынешней эпохи или конца света, о конечной судьбе человечества — эсхатологии — и приходе нового порядка. Это апокалиптическое мышление. Древние пророки предсказывали новый божественный и земной порядок после периода суда. Их пророчества говорили о радикальном политическом и духовном преобразовании. В Еврейской Библии есть Книга Даниила, предсказывающая будущее, полное грандиозных космических и политических перемен. В пророчествах фигурируют безумные цари, чудовищные звери, выходящие из моря, и ангелы, истолковывающие видения. Книга Даниила говорит о росте знания, об «образе зверя» и о грядущем царстве, которое сокрушит все остальные. Один из зверей описывается как имеющий «железные зубы» — отрывок, который иногда толкуют как намёк на технологическое будущее и власть: «Видел я в ночных видениях, и вот зверь четвёртый, страшный и

ужасный и весьма сильный; у него большие железные зубы; он пожирает и сокрушает, остатки же попирает ногами; он отличен был от всех прежних зверей и имел десять рогов». [9]

Эти пророчества содержат и предсказание, оказавшее огромное влияние на западную религиозную историю. В девятой главе Книги Даниила говорится, что придёт «мессия» — слово происходит от еврейского *mashiach*, «помазанник», — и что этот помазанник, называемый также «правителем» или «вождём», «будет предан смерти». [10] В древнем Израиле людей, избранных на роль царя, священника или пророка, помазывали маслом. В иудаизме Иисус — пророк, посланник Бога. Но для христиан он гораздо больше: воплощённое божественное откровение и сам Сын Божий. Его считают предсказанным мессией, который вызывает политические потрясения и открывает возможность радикального духовного обновления в форме христианского спасения. Это можно понимать как личное духовное спасение, но также и как коллективное, земное и политическое. Несмотря на слова, сказанные Понтию Пилату на суде, — «Царство Моё не от мира сего», — Иисус воспринимался не только как духовный лидер, но и как политический мятежник и потенциальный освободитель, «Царь Иудейский». Он не только нарушил порядок в Иерусалимском храме — что само по себе было политическим актом, — но после вступления в Иерусалим как «Сын Давидов» от него ожидали освобождения евреев от римского владычества и создания иудейского царства. [11] Распятие было наказанием для политических мятежников и преступников.

Ещё одна важная для нашей темы библейская книга — Откровение, или Апокалипсис, традиционно приписываемое апостолу Иоанну и завершающее Новый Завет. Здесь появляются знаменитые четыре всадника — завоевание, война, голод и смерть, — семь ангелов, насылающих язвы и бедствия, дракон, преследующий ребёнка — Христа, падение Вавилона, поражение Зверя и видение Нового Иерусалима. Структуру этого повествования легко обнаружить и во множестве голливудских фильмов: после завязки и появления героя начинаются испытания и проблемы, герой страдает, но затем наступает счастливый финал и появляется надежда. Происходит несколько сражений, нарастает кризис, ставки повышаются, но в последней эпической битве, где мятежники противостоят империи, силы зла терпят поражение, и на земле устанавливается подобие рая.

Одна из влиятельных интерпретаций Откровения состоит в том, что сейчас существуют страдания и конфликт, но в конечном счёте добро победит. Нужно лишь ждать, вести себя праведно, бороться со злом — и в итоге всё будет хорошо. Такой образ мышления использовался и для оправдания войн — враг служит Сатане, его необходимо уничтожить, после чего наступит благо, — и для поддержки существующего несправедливого порядка. Сегодня ты можешь занимать униженное положение, быть рабом, но если останешься добрым и верующим, тебя ждёт светлое будущее. Этот мир ужасен, зато следующий будет прекрасен.

Маркс, как известно, критиковал идею пассивного перенесения страдания ради будущей утопии, а я уже упоминал критику Ницше, считавшего религию средством социального контроля и конформизма. Аналогично повествование Откровения можно рассматривать как «опиум для народа», побуждающий терпеть страдания и принимать систему вместо того, чтобы активно разрушать угнетающие структуры. Но одновременно эту историю можно читать и как критику репрессивных политических систем. Откровение можно истолковать как иудейскую критику преследований при римском владычестве. Отказ принять печать Зверя становится актом сопротивления, а Новый Иерусалим — видением справедливого общества без угнетения. Такое толкование подчёркивает подрывной и освобождающий потенциал текста. Мы можем и должны стремиться покончить с несправедливостью. Правда, утверждение, будто крах угнетающих режимов неизбежен, снова способно привести к пассивности. Если капитализм и империя всё равно падут, зачем восставать — почему бы просто не подождать?

В Средние века и позднее конец времён регулярно предсказывали такие фигуры, как Хильдегарда Бингенская, Мартин Лютер и Нострадамус. И здесь содержание пророчеств также представляло собой смесь религии и политики. Хильдегарда Бингенская, например, выступала против злоупотреблений властью и требовала реформ. Лютер видел в папе и турках приближение Антихриста. Некоторые толкователи считают, что Нострадамус предсказал Французскую революцию, возвышение Наполеона Бонапарта, приход Гитлера и нынешние и будущие войны на Ближнем Востоке. А вслед за Книгой Откровения некоторые христиане — как на ранних этапах истории христианства, так, например, и сегодня в США — придерживаются миллениаризма, согласно которому перед Страшным судом наступит тысячелетняя мессианская эпоха, когда Христос будет царствовать на земле. В Откровении 20:1–6 говорится об ангеле, связывающем Сатану, и о тысячелетнем царстве Христа со святыми. Миллениаристы ждут возвращения Христа на землю и установления его царства — буквального или духовного.

Средневековая иоакимитская традиция — ещё один пример апокалиптического и эсхатологического мышления с отчётливым политическим измерением. Она получила название от Иоахима Флорского, итальянского теолога и основателя монашеского ордена, и утверждала, что наступит новая эпоха — Эра Духа, — когда духовная элита поведёт человечество к более чистой духовности, заменив собой развращённую Церковь. Эти идеи вдохновляли радикальных францисканцев и протестантов, а их отголоски можно найти даже в утопическом марксизме. Маркса и в самом деле обвиняли в проповеди секулярной версии миллениаризма и апокалиптического мышления: его интерпретировали так, будто он предсказывал, что революция неизбежно положит конец капитализму и принесёт утопию коммунизма. И снова решительный разрыв с прошлым представляется как необходимое условие радикально иного будущего. Эрнст Блох, например, в «Принципе надежды» утверждал, что современный утопический социализм с его надеждой на лучший мир продолжает

апокалиптические и милленаристские традиции, включая идеи Иоахима Флорского. [12]

Связанное с этим религиозное явление — культы конца света. Это религиозные группы, исповедующие апокалиптизм и милленаризм. Они предсказывают скорые катастрофические события и иногда сам конец света, к которому заранее готовятся. Такие культы обычно возникают во времена экономической и политической нестабильности и тревоги. Харизматические лидеры заявляют, что знают будущее, и приобретают власть над последователями. Исторические примеры — табориты в Чехии XV века; мюнстерские анабаптисты, чьи взгляды напоминали иоахимитский милленаризм и которые в XVI веке на некоторое время захватили немецкий город Мюнстер; а также секты конца света в США и Японии в 1990-е годы. Нередко история таких движений заканчивалась трагически. Восстание анабаптистов было жестоко подавлено, калифорнийская группа «Врата рая» завершила своё существование массовым самоубийством, а в 1995 году члены японского апокалиптического культа устроили газовую атаку в токийском метро. Но их политическое влияние не ограничивается радикализацией и терроризмом. Апокалиптические верования подобных культов способны формировать общественную дискуссию о технологических рисках и поддерживать общественные страхи и тревоги. И вообще, не только в рамках отдельных культов, христианское апокалиптическое мышление, интерес к концу времён и пророчествам всегда были частью современной американской культуры и влияли на общественную жизнь. [13] Возможно, апокалиптическое мышление и ожидание конца света вообще являются постоянным ингредиентом современной западной культуры, включая наше отношение к технологиям. Вспомним, например, представления о конце света, связанные с ядерным оружием, которые захватили общественное воображение во второй половине XX века: множество людей с тревогой ждали «бомбы», то есть полного уничтожения человечества в ядерной войне. Как писал Деррида, ядерная война была не только технологической проблемой, но и проблемой риторики, дискурса и убеждения. [14] Не только бомбы, но и тексты. И эта риторика, эти тексты несомненно носили апокалиптический характер.

То же происходит и сегодня: апокалиптические представления об ИИ и сценарии конца света циркулируют в обществе, формируя публичный дискурс и политическое воображение вокруг технологии и её рисков. ИИ — это не только технология, но и риторика, причём риторика апокалиптическая. Роберт Джерачи предложил термин «апокалиптический ИИ», подчёркивая, что апокалиптическое мышление вовсе не осталось в далёком прошлом, а живо и сегодня — в популярных научных книгах об ИИ и робототехнике. [15] В сходном духе Норин Херцфельд называет ИИ «милленаристским проектом», позволяющим людям проецировать на него «эсхатологические надежды и страхи». [16] Спасение или уничтожение — теперь всё якобы в руках ИИ. Это ещё один пример того, как, по словам Джерачи, «наше технонаучное наследие укоренено в религиозной жизни западного мира» и его древних источниках. [17]

Каковы же апокалиптические прогнозы относительно ИИ? Что представляет собой, в частности, трансгуманистическая эсхатология, столь влиятельная в современных разговорах о технологическом будущем? Утверждается, что ИИ приведёт к концу нынешней эпохи: наступит Сингулярность, после которой всё станет иным. [18] Произойдёт неконтролируемый взрыв интеллекта, результатом которого станет сверхразум. Бостром ожидает, что последствия этого процесса будут либо чрезвычайно благоприятными, либо чрезвычайно разрушительными — но в любом случае крайними. [19] Общества будут фундаментально преобразованы. Если воспользоваться метафорой из Евангелия от Матфея, «не останется здесь камня на камне». [20] Нынешнее развитие ИИ понимается как подготовка к будущему сценарию радикального перелома. Предполагается, что современные технологии уступят место искусственному общему интеллекту — AGI. Это форма ИИ, которая сначала сравнивается с человеческими возможностями, а затем превзойдёт их. Иными словами, однажды наши машины станут сверхразумными. [21] Одни видят в нынешних LLM ступени на пути к AGI, другие считают, что для этого нужен совершенно иной тип ИИ. Но большинство сторонников Сингулярности сходятся в одном: конец времён не так уж далёк. Конец — или, если угодно, новое начало — близок. Курцвейл прямо пишет: «Сингулярность близка»; в книге с таким названием он предсказывает «глубочайшую переменную» к 2045 году. [22] Большинство других прогнозов — Моравека, а иногда и Маска — тоже указывают примерно на середину XXI века, хотя Бостром формулирует осторожнее, ограничиваясь общим горизонтом XXI столетия.

Что же произойдёт после этого? Считается, что последствия сверхразума для человеческой цивилизации будут радикальными и необратимыми. Когда сверхразумный ИИ возьмёт верх, человечество в привычном нам виде закончится. Мы станем транслюдьми — существами, значительно усовершенствованными технологиями, — либо сольёмся со сверхразумными машинами и расселимся по космосу, где будут процветать наши «дети разума», как называл их Моравек, — наши потомки в форме сверхразумного ИИ. [23] Человеческий интеллект окажется ничтожным по сравнению со сверхразумом новых существ, которые будут жить в космосе. Возможно, эти существа будут цифровыми. Представляется, что миллиарды цифровых людей — или, точнее, сверхразумных сущностей — станут существовать в компьютерных симуляциях.

Под влиянием работ Бострома, а также Тоби Орда, Уильяма Макаскилла и Элиезера Юджовского некоторые так называемые лонгтермисты утверждают, что нам следует ставить долгосрочные, крайне отдалённые интересы этих сверхразумных существ выше нынешних интересов, поскольку их моральный вес — наш будущий «потенциал» — огромен. [24] С такой лонгтермистской и утилитаристской точки зрения интересы сравнительно небольшого числа людей, живущих сегодня, имеют не слишком большое значение. Да и выбора у нас, возможно, всё равно нет. Человечество обречено. Оно устарело. Конец близок — и это конец самого человечества. В статье для *Time*

Юджовский формулирует свои ожидания совершенно недвусмысленно: «наиболее вероятный результат создания ИИ, превосходящего человека по интеллекту, при обстоятельствах, хотя бы отдалённо напоминающих нынешние, состоит в том, что буквально все люди на Земле умрут». [25] Как утверждает Эмиль Торрес, один из наиболее заметных критиков лонгтермизма, подобный образ мышления опасен. Если великое космическое благо следует ставить выше интересов обычных людей нынешнего и ближайших поколений, то сегодняшняя бедность, климатические изменения и другие современные проблемы уже не кажутся столь важными. [26] В грандиозной картине они якобы ничего не значат. [27] А мыслители конца света — от Средневековья до наших дней — всегда заняты именно грандиозной картиной. Им нет дела до нас с вами сегодня: предполагается, что мы просто умрём, пока готовится новое, более важное будущее.

Если бы это была всего лишь маргинальная, сектантская и эзотерическая картина мира, какой она на первый взгляд кажется, никакой серьёзной проблемы для дискурса об ИИ — а тем более для политики ИИ — не возникло бы: её можно было бы просто игнорировать. Но то, что начиналось на периферии элитных британских университетов и исследовательских институтов Кремниевой долины, сегодня стало верой и идеологией могущественных руководителей Big Tech, таких как Илон Маск и Сэм Альтман. Маск не только финансировал трансгуманистические и лонгтермистские центры вроде Future of Life Institute и ранее Future of Humanity Institute Бострома, но и включал эти идеи в собственные планы технологического будущего. Для Маска колонизация Марса — ступень в новой, лонгтермистской истории человеческой эволюции и цивилизации. Современный ИИ готовит путь к нашему Дивному Новому Будущему. Big Tech поможет это будущее осуществить. Для человечества это вовсе не обязательно хорошая новость. Возможно, Маск действительно хочет спасти часть человечества — скажем, себя и нескольких друзей-миллиардеров, — но лишь затем, чтобы колонизировать космос, расселиться по Вселенной и войти в новое будущее в форме усовершенствованных существ или симуляций. Большинство нынешних смертных там попросту не нужны; человечество всё равно исчезнет.

После своих предупреждений о будущем ИИ Маск оказался связан с так называемым «движением за безопасность ИИ», которое одновременно торгует ожиданиями сверхразума, раздувает значимость ИИ и предлагает сценарии гибели человечества. При поддержке миллиардеров вроде криптомагната Сэма Бэнкмана-Фрида и предпринимателя Дастина Московица это движение пыталось влиять — если не колонизировать — саму сферу этики ИИ, желая обсуждать риски исключительно сквозь призму лонгтермистского видения. В политической сфере оно даже продвигало открытое письмо с призывом приостановить исследования ИИ, исходя из апокалиптического предположения, что достаточно продвинутый ИИ способен привести человечество к вымиранию. [28] Среди подписавших были Маск, Харари и исследователь машинного обучения Макс Тегмарк, однако в эту дискуссию оказались

вовлечены и многие другие специалисты по ИИ. Искусственный интеллект всё больше превращается в культ конца света — причём культ, подкреплённый долларами.

Технологический миллиардер Питер Тиль также находится под влиянием того, что Торрес называет «опьяняющей новой светской религией». [29] Готовясь к лонгтермистскому апокалиптическому будущему, Тиль обзавёлся убежищем в Новой Зеландии. В случае краха цивилизации он якобы собирается скрыться там вместе с Альтманом. [30] Последний разделяет веру Маска в апокалиптический потенциал ИИ. Подобно Маску, Тиллю и другим технологическим миллиардерам, Альтман также рассуждал о катастрофических сценариях ИИ и его «экзистенциальных рисках». Как указывает Торрес, сам этот термин вводит в заблуждение. Поскольку эти ключевые фигуры мира ИИ, обладающие огромным глобальным влиянием, придерживаются лонгтермистских взглядов, их не особенно беспокоит будущее ныне живущих людей — таких как вы и я. В лучшем случае их интересует далёкое будущее, а нынешние проблемы становятся важны лишь тогда, когда представляют собой экзистенциальные риски, способные помешать осуществлению этого будущего. Всё происходящее сегодня — лишь слабая рябь на поверхности гораздо более масштабного апокалиптического повествования и истории спасения, в которой лидеры Big Tech отводят себе ключевую роль.

С точки зрения истории религии это неудивительно. Предсказание — это власть. Пророк способен стать спасителем. Царь, прислушавшийся к предупреждениям оракула, может многое выиграть. Внимая посланникам будущего вроде Харари и Бострома, предсказывающим лонгтермистскую историю спасения посредством ИИ, руководители Big Tech — Маск, Альтман и другие — легко могут представить себя спасителями, которые не только знают, что ждёт человечество, но и обещают избавить нас от бед обычного человеческого существования и смертности, введя в новую эпоху апокалипсиса Сингулярности и сверхразума. Для широкой публики они способны предстать почти всезнающими — а вскоре и всемогущими — проводниками в грядущий Дивный Новый Мир. Это даёт им политическую власть. В некоторой степени та же власть достаётся и их советникам, даже если их послания принимаются далеко не всегда благосклонно. Подобно Пифии и особенно Иисусу и другим библейским пророкам, такие авторы, как Харари, могут оказаться политическими и революционными фигурами даже против собственной воли, пусть они и утверждают, что «всего лишь» говорят от имени и из перспективы неизбежного технокосмического будущего.

Мы уже видим, что технологические миллиардеры готовы претендовать на всё большую политическую власть. В случае Маска это приобрело вполне конкретные формы. Он использовал принадлежащую ему социальную платформу X, ранее Twitter, для наращивания политического влияния и был назначен руководителем «Департамента эффективности правительства» — DOGE — в администрации Трампа, получив доступ к платёжным системам Министерства финансов США. Ещё до выборов

2024 года Маск открыто заявлял, что хочет большей формальной политической власти и «готов служить», как он прямо сказал в одном интервью. [31] Вполне возможно, что подобная готовность «служить» обнаружится и среди пророков якобы уже предопределённого будущего ИИ, нашёптывающих свои предсказания славы и гибели в охотно слушающие уши технологических царей.

Обращение к истории религиозных идей, таким образом, помогает ответить на весьма уместный вопрос, иронично заданный в Х: «Как получилось, что совершенно маргинальные культы ИИ-конца света приобрели столь огромное и ничем не оправданное политическое влияние?» [32] И задаёт его не эксцентричный левый профессор политической теории, а Янн Лекун — главный научный сотрудник Meta по искусственному интеллекту.

Примечания

[1] Описания и научные объяснения дельфийских прорицаний см., например: William J. Broad, *The Oracle: Ancient Delphi and the Science Behind Its Lost Secrets* (Penguin, 2006).

[2] Bertrand Russell, “Has Religion Made Useful Contributions to Civilization?,” in *Why I Am Not a Christian*, ed. Paul Edwards (Routledge, 1996), 24.

[3] Plato, *Apology*, trans. G. M. A. Grube, in Plato, *Five Dialogues*, 2nd ed., rev. John M. Cooper (Hackett Publishing, 2002), 21–44, ll. 21a, 30e.

[4] Helga Nowotny, *In AI We Trust: Power, Illusion and Control of Predictive Algorithms* (Polity Press, 2021).

[5] Подробнее об ИИ и времени см.: Mark Coeckelbergh, “Time Machines: Artificial Intelligence, Process, and Narrative,” *Philosophy and Technology* 34, no. 4 (2021): 1623–1638; см. также Mark Coeckelbergh, *Digital Technologies, Temporality, and the Politics of Co-Existence* (Palgrave Macmillan, 2023).

[6] См., например: Kurzweil, *The Singularity Is Near*; Yuval Harari, *Homo Deus*; Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford University Press, 2014).

[7] Epstein, *Tech Agnostic*, 76, 82; о связи с русским космизмом см.: Michel Eltchaninoff, *Lenin Walked on the Moon*, trans. Tina Kover (Europa Editions, 2023); о технологическом гностицизме см., например: Michael Zimmerman, “Religious Motifs in Technological Posthumanism,” *Western Humanities Review* 3 (2009): 67–83.

[8] Bill Joy, “Why the Future Doesn’t Need Us,” *Wired*, April 1, 2000.

[9] Дан. 7:7.

[10] Дан. 9:25–26.

[11] Ин. 18:36.

[12] Ernst Bloch, *The Principle of Hope*, vol. 1, trans. Neville Plaice, Stephen Plaice, and Paul Knight (MIT Press, 1995).

[13] См., например: Paul Boyer, *When Time Shall Be No More: Prophecy Belief in Modern American Culture* (Harvard University Press, 1992).

[14] Jacques Derrida, “No Apocalypse, Not Now (Full Speed Ahead, Seven Missiles, Seven Missives),” *Diacritics* 14, no. 2 (1984): 20–31.

[15] Robert M. Geraci, *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality* (Oxford University Press, 2010), 1.

[16] Noreen Herzfeld, “The Eschatological Future of Artificial Intelligence,” in *The Cambridge Companion to Religion and Artificial Intelligence*, ed. Beth Singler and Fraser Watts (Cambridge University Press, 2024), 148–149.

[17] Geraci, *Apocalyptic AI*, 13.

[18] См. прежде всего Kurzweil, *The Age of Spiritual Machines* и *The Singularity Is Near*, а также Bostrom, *Superintelligence*.

[19] Bostrom, *Superintelligence*, 2, 20.

[20] Мф. 24:1–2.

[21] Bostrom, *Superintelligence*, 22.

[22] См. Kurzweil, *The Singularity Is Near*.

[23] Moravec, *Mind Children*.

[24] Toby Ord, *The Precipice: Existential Risk and the Future of Humanity* (Bloomsbury Publishing, 2020); William MacAskill, *What We Owe the Future* (Basic Books, 2022). Иногда лонгтермизм объединяют с так называемым эффективным альтруизмом — движением, которое под влиянием утилитаризма предлагает беспристрастно рассчитывать наибольшее благо и расставлять приоритеты между различными целями. В сочетании с лонгтермизмом эта идея направлена на улучшение будущего. Орд и Макаскилл относятся к эффективным альтруистам. О движении эффективного альтруизма и сообществе вокруг Макаскилла см., например: Gideon Lewis-Kraus, “The Reluctant Prophet of Effective Altruism,” *New Yorker*, August 8, 2022. Критический взгляд см.: Émile P. Torres, “Why Effective Altruism and ‘Longtermism’ Are Toxic Ideologies,” interview by Nathan Robinson, *Current Affairs*, May 7, 2023.

[25] Eliezer Yudkowsky, “Pausing AI Developments Isn’t Enough. We Need to Shut It All Down,” *Time*, March 29, 2023.

[26] Émile P. Torres, *Were the Great Tragedies of History “Mere Ripples”? The Case Against Longtermism* (published by author, 2024).

[27] См. также Torres, “Why Effective Altruism and ‘Longtermism’ Are Toxic Ideologies.”

[28] “Pause Giant AI Experiments: An Open Letter,” Future of Life Institute, March 22, 2023.

[29] Émile P. Torres, “Elon Musk, Twitter and the Future: His Long-Term Vision Is Even Weirder Than You Think,” *Salon*, April 30, 2022.

[30] Mark O’Connell, “Why Silicon Valley Billionaires Are Prepping for the Apocalypse in New Zealand,” *The Guardian*, February 15, 2018.

[31] “Elon Musk ‘Willing To Serve’ After Donald Trump Hints at Cabinet Post for Tesla CEO,” *Economic Times* (India), August 20, 2024. Автор добавляет: к моменту завершения книги Маск уже был назначен в администрацию Трампа.

[32] Yann LeCun (@ylecun), “How did outlandishly fringe AI doomsday cults gain so much unwarranted political influence?,” X, August 16, 2024, 6:22 pm.

7. Позволь мне жить вечно: Гильгамеш, воскресение и машины бессмертия

Когда умирает его друг Энкиду, царь Гильгамеш — правитель Урука в древней Месопотамии, прославившийся как могущественный тиран, — погружается в скорбь. Испугавшись собственной смертности, он отправляется на поиски тайны вечной жизни. Царь Утнапиштим, переживший потоп благодаря кораблю, на котором он сохранил жизнь — в повествовании, возникшем раньше библейской истории о Ное, — говорит ему, что бессмертие есть дар богов и его невозможно заслужить. Но Гильгамеш продолжает поиски, потому что узнаёт от Утнапиштима и о растении, дарующем вечную молодость. В конце концов он находит на дне моря «растение омоложения». Однако по дороге домой змея — древний символ возрождения и преображения — похищает растение. Гильгамеш сокрушён. Ему приходится признать собственную смертность. Бессмертие предназначено природе или богам; для человека смерть неизбежна. Утешение Гильгамеш находит в том, что его деяния и труды переживут его. Его город и его народ продолжают существовать после его смерти.

Миф о Гильгамеше — одна из древнейших известных историй человечества. Разумеется, это далеко не единственное повествование о человеческой смертности и поиске бессмертия; многие культуры и цивилизации были этим глубоко заняты, если не одержимы. [1] В определённом смысле сама цивилизация — результат того, что Стивен Кейв назвал человеческой «волей к бессмертию»: значительную часть истории религии, философии и искусства можно рассматривать как следствие нашего желания продолжить собственное существование. [2] Эпос о Гильгамеше отражает, пожалуй, самый фундаментальный экзистенциальный конфликт человека и человечества: смерть — скандал. Мы не можем и не хотим её принять. Мы хотим жить вечно. Или, как сказано в Первом послании к Коринфянам: «Последний же враг истребится — смерть». [3]

Бессмертие не всегда понимается как бесконечное продолжение физического существования. В индуизме, например, бессмертие означает духовное освобождение души — *атмана* — из круговорота рождения, смерти и нового рождения — *сансары*. Самопознание ведёт к духовному бессмертию, состоянию, в котором душа освобождается от земных привязанностей. Такие индуистские тексты, как «Бхагавадгита» и Упанишады, говорят об атмане — индивидуальной душе или самости, — которая едина с Брахманом, мировой душой и высшей реальностью. Освободившийся человек осознаёт тождество атмана и Брахмана: самость в своей изначальной и чистой форме и есть Брахман. Именно это выражает знаменитая формула *Tat Tvam Asi* — «Ты есть То» — из «Чхандогья-упанишады». Осознание этого единства и является ключом к бессмертию. Как учит «Бхагавадгита», душа вечна и не исчезает со смертью тела: «Душа не рождается и никогда не умирает; возникнув

однажды, она никогда не перестаёт быть. Душа нерождённа, вечна, бессмертна и нестареющая. Она не уничтожается, когда уничтожается тело». [4]

Многие древние греки также верили в загробную жизнь или бессмертие души. В «Илиаде» и «Одиссее» загробное существование описывается как сумрачная, призрачная жизнь в Аиде, куда уходят мёртвые. Позднее более сильное представление о бессмертии получило распространение под влиянием религиозных движений — например, орфизма и Элевсинских мистерий, — а также философии. Уже Пифагор учил, что душа бессмертна и проходит через череду жизней. В отличие от более поздних философов вроде Аристотеля, относившегося к этой идее скептически, Платон в «Федоне» также утверждал, что душа вечна. Она существует до нашей земной жизни и переживает смерть тела. Души умерших уходят в иной мир, а затем возвращаются. [5] В «Федре» Сократ говорит, что душа бессмертна, и приводит аргументы в пользу этого. Он описывает, как душа поселяется в земном теле. Возникающее при этом живое существо, однако, уже не бессмертно. [6]

И всё же в некоторых древних религиях — особенно в тех, которые впоследствии повлияли на западную культуру и религии, — и физическое тело рассматривается как потенциально бессмертное. Древние египтяне верили в воскресение тела после смерти: умерший продолжал существовать в загробном мире, где его тело жило дальше в преображённой и идеализированной форме. Практика мумификации — дорогостоящий процесс и ритуал, предназначенный главным образом для богатых и могущественных, таких как фараоны и высокопоставленные чиновники, — укрепляла власть элиты и представление о том, что эти люди избраны богами. Она отражала веру в необходимость сохранить тело, чтобы оно могло вновь ожить в загробном мире, когда душа — *ба* — вновь соединится с телом — *ка*. Миф об Осирисе рассказывает о его воскресении: он становится богом загробного мира и судьёй умерших. Древний зороастризм, возникший в Персии и глубоко повлиявший на иудейскую мысль во время Вавилонского пленения, также содержит веру в воскресение мёртвых в конце времён: они восстанут в своих физических телах и будут судимы. Последователи митраизма, мистериальной религии Римской империи, верили в воскресение своего спасителя Митры.

Христиане также верят в воскресение, причём в двух формах: во-первых, в воскресение Иисуса Христа — краеугольный камень христианского богословия, который, как уже отмечалось, имеет корни в иудейской эсхатологии и описывается в Евангелиях и посланиях Павла, — и, во-вторых, во всеобщее воскресение в конце времён, когда мёртвые восстанут из могил и предстанут перед судом. Считается, что праведники получают вечную жизнь в преображённом, прославленном теле, а нечестивые будут навеки отделены от Бога — особенно см. Евангелие от Иоанна, Первое послание к Коринфянам и послания к Фессалоникийцам. Апостольский символ веры II века н. э. и Никейский символ веры 325 года закрепили эти представления. Традиционно христиане также верят в рай и ад: рай — это состояние мира в

присутствии Бога, вечное единение с ним, тогда как ад — состояние разлуки и страдания. Католики добавляют к этому чистилище — временное состояние очищения перед входом в рай. Куда попадёт каждая душа, определяется судом. В конце времён, при Втором пришествии Христа, состоится последний суд. Таким образом, бессмертие оказывается чем-то будущим. Сначала мы умираем. Однако, опираясь на Евангелие от Иоанна и послания Павла, некоторые утверждают, что вечная жизнь возможна и уже сейчас, если мы верим в Бога. В Ин. 5:24 сказано: «Истинно, истинно говорю вам: слушающий слово Моё и верующий в Пославшего Меня имеет жизнь вечную и на суд не приходит, но перешёл от смерти в жизнь». Быть христианином тогда означает обрести «новую жизнь» во Христе. [7] В Лк. 17:21 Иисус говорит: «Царствие Божие внутри вас есть».

Обратите внимание: христианская вера в то, что души будут судимы и разделены на спасённых и осуждённых, метафорически отражает древние сельскохозяйственные практики сортировки и классификации и использовалась для оправдания механизмов включения и исключения в обществе. Сельскохозяйственная сортировка служит в Новом Завете метафорой суда — например, в притче о пшенице и плевелах, — и эта логика переносится не только в духовную сферу, здесь — эсхатологию, но и в общественную и политическую, укрепляя социальные и политические иерархии. Сортируют души — сортируют людей. В Средние века на Западе суждения о душе были связаны с социальным и политическим контролем. Тех, кто не соответствовал общественным и религиозным нормам, часто изгоняли на периферию общества или преследовали. Как утверждал Мишель Фуко в «Надзирать и наказывать» и в своих работах о сексуальности, современная власть действовала через дисциплинарные механизмы — такие как экзамен и исповедь, — выступавшие инструментами контроля и надзора. [8] В определённой степени это продолжение христианских практик. Подчинение суду — на фоне возможности спасения и бессмертия — и «пастырская власть» формировали людей и управляли их душами, опираясь на знание не только об их поведении, но и о том, что происходит внутри их сознания. Это создавало механизмы включения и исключения и укрепляло власть религиозных институтов — Церкви — и политических институтов, таких как современное государство, которые продолжили формы пастырской власти и, возможно, сегодня пытаются осуществлять её новыми способами при помощи технологий. [9] С ИИ и частные корпорации, и государства не только следят за нашим поведением, но и собирают знания о нашем сознании и потенциально способны им манипулировать. Разум и душу теперь количественно измеряют, анализируют и изменяют. [10] Учёт души ведётся цифровыми средствами.

Но ИИ связан и с поиском бессмертия. Современные цифровые технологии, включая искусственный интеллект, можно рассматривать как ответ на эти экзистенциальные и религиозные стремления к бессмертию — и одновременно как часть связанных с ними властных и политических процессов. Это то, что я называю «машинами бессмертия»: попытки достичь бессмертия технологическими средствами.

Это происходит по меньшей мере в нескольких смыслах.

Во-первых, цифровые технологии создают иные миры и обещают новые жизни. Виртуальные миры и компьютерные игры, которые благодаря ИИ становятся всё реалистичнее, обещают освобождение от скучной, опасной, смертной и телесной жизни. В этих захватывающих новых мирах и повествованиях кажется, будто мы больше не связаны физическим телом или даже прежней идентичностью; мы можем быть героями или богами, можем быть кем угодно — теми, кто мы есть, или теми, кем хотим стать. Наши цифровые «я», отделённые от физических тел, достигают свободы и состояния, которое кажется бессмертием или, по крайней мере, некоторой формой загробной жизни. Возможно, эта загробная жизнь вовсе не обязательно приятна. Но будем ли мы там в раю или аду — или станем вести более сумрачное «виртуальное» существование в цифровом Аиде в противоположность «реальной» физической жизни, — мы уже не здесь. А именно этого и желает наше религиозное и апокалиптическое сознание. Как выразился Джерачи, «апокалиптический ИИ обещает своим верующим вечную награду в виртуальной загробной жизни». [11] В качестве примера он приводит Second Life — виртуальный мир, в котором некоторые люди хотят жить вечно. Можно вспомнить и другие виртуальные миры, игровые пространства и технологические проекты вроде метавселенной. [12]

Во-вторых, хотя после погружения в эти миры и игры нам всё же приходится возвращаться к земному существованию, некоторые трансгуманисты, например Курцвейл, верят, что однажды мы сможем загружать свой разум в компьютеры и становиться «цифровыми бессмертными». [13] Идея состоит в том, чтобы сканировать мозг человека и воссоздавать его психическое состояние внутри цифрового компьютера. Затем компьютер должен симулировать информационные процессы мозга. Предполагается, что таким образом сознание человека можно будет заново воплотить и хранить в цифровой форме вечно. Курцвейл считает, что при этом удастся сохранить личность, память, навыки и жизненную историю человека. [14] Предполагается, что сознание и идентичность будут перенесены в компьютеры и при этом останутся чувствующими и осознающими. Тело тогда становится ненужным; если оно вообще понадобится, его можно будет заменить механическими частями, предлагает Курцвейл. Биологические мозги превратятся в машинные разумы. Он предсказывал, что к 2045 году мы сможем «загружать» себя. [15] Таким образом, речь идёт, возможно, не о бессмертии души, но хотя бы о бессмертии сознания. Или, выражаясь гностическим языком, несовершенное, испорченное, материальное и смертное тело наконец можно будет оставить позади. Душа покинет свою темницу.

В-третьих, уже сегодня возможно создавать так называемых *deadbots* или *griefbots* — ботов умерших, или ботов для переживания утраты, — что напоминает «цифровое воскресение» мёртвых. Умершим дают «загробную жизнь ИИ»: чат-боты воспроизводят их речевые привычки и особенности личности на основе цифровых данных — текстовых сообщений, электронной почты, постов в социальных сетях, ответов на

личностные анкеты, голосовых записей и так далее. [16] Наши данные всё равно продолжают существовать в цифровом пространстве — так почему бы ими не воспользоваться? Главная идея состоит в том, чтобы утешить скорбящих: благодаря такому «воскресению ИИ» людям создают иллюзию, будто они теперь могут разговаривать с умершими близкими. [17] Некоторые компании позволяют также увидеть сгенерированное ИИ видео, имитирующее умершего. Семья может продолжать взаимодействовать с цифровой версией покойного, представленной как аватар или голограмма. Возможно, такие *deadbots* или «ИИ-призраки» могут помочь и тем, кому самому предстоит умереть? [18] Как говорит Лоранс Девиллерс: «Люди всегда хотели быть неуязвимыми, бессмертными. Это часть наших основополагающих мифов — никто не хочет умирать». [19] Виртуальный аватар тогда выглядит как, возможно, частичное решение.

В-четвёртых, помимо «воскресения разума», сегодня в повестку входит и бессмертие тела. Технологии уже используются для увеличения продолжительности жизни, а теперь — вновь в трансгуманистском духе — предполагается, что они смогут радикально продлевать жизнь и даже после смерти сохранять тела с возможностью их последующего оживления.

Трансгуманисты предлагают радикальное продление жизни и верят, что технологии в конечном счёте устранят старение и болезни как причины смерти. Обри де Грей и Майкл Рэй утверждали, что старение — болезнь, которую можно обратить вспять медицинским вмешательством. Подобно другим трансгуманистам, они считают, что ограничения человеческого тела необходимо преодолеть, включая ограничения, налагаемые старением и смертью. Если мы поймём биологические процессы старения, то сможем обратить их. Мы сможем восстанавливать повреждения клеток. Это позволит людям жить сотни или тысячи лет в хорошем здоровье. [20] ИИ может помочь открывать лекарства и разрабатывать питание, воздействующие на биологические механизмы старения. Курцвейл также рассчитывает на нанороботов, способных лечить болезни. [21] Мало кто захочет жить дольше, если это означает лишь продление страдания. Но если можно было бы прожить долгую и здоровую жизнь — почему нет? Такое долголетие ещё не бессмертие, но по крайней мере нечто весьма желательное: возможность жить дольше.

Однако некоторые трансгуманисты верят, что даже после смерти ещё можно что-то сделать — то есть технологически воскресить тело. Крионика — практика сохранения тел людей, умерших от пока неизлечимых заболеваний, при чрезвычайно низких температурах — например, -196°C . После смерти тело быстро охлаждают. Это замедляет процессы, вызывающие дальнейшее повреждение клеток. Для предотвращения образования льда используется раствор-криопротектор. ИИ может помочь совершенствовать методы криоконсервации. Возможно, пригодятся и наноботы. Надежда состоит в том, что будущие медицинские технологии позволят оживить этих людей и вылечить. Как писал Кейв, крионику можно рассматривать как

технологический путь к загробной жизни: разновидность воскресения. [22] Только на этот раз людей должны воскресить не боги, а учёные будущего. Тиль уже записался. Его отношение к смерти? «Я против». [23]

Подобно технологическим царям, Гильгамеш тоже был «против смерти». В экзистенциальном смысле мы все против смерти и почти все так или иначе боремся с ней. Нам трудно принять собственную смерть и смерть тех, кого мы любим. Возможно, бороться со смертью действительно естественно, как считает Тиль. [24] В каком-то смысле мы — борцы с уязвимостью, использующие технологии как оружие в этой борьбе, как я писал в книге *Human Being @ Risk*. [25] С помощью машин мы боремся с человеческой уязвимостью и смертностью. ИИ — не просто новая технология, но новое средство против уязвимости и новая пушка против смерти. Но способна ли технология спасти нас? Может ли нас спасти ИИ? Может ли *deus ex machina* даровать вечную жизнь?

Трансгуманисты настроены оптимистично. Но существует множество причин считать эти идеи и технологические практики, призванные дать нам бессмертие, философски проблематичными, этически спорными, научно спекулятивными и в конечном счёте тщетными перед лицом человеческой смертности. И эти дискуссии касаются фундаментальных философских и религиозных вопросов о природе человека и сознания — и, конечно же, вечной экзистенциальной и религиозной проблемы смертности и бессмертия.

Виртуальные миры и компьютерные игры на самом деле не позволяют нам вырваться из нашего уязвимого и телесного состояния. Проблема не только в том, что современные системы не способны дать весь спектр чувственного опыта и физических взаимодействий, к которым мы привыкли в обычном физическом мире, и потому в этом смысле недостаточно полно имитируют жизнь в другом мире. Проблема гораздо глубже и связана с представлениями о соотношении сознания и тела, лежащими в основе идеи, будто мы можем «убежать» разумом в иной мир. Погружение в виртуальную реальность — по крайней мере если понимать его как бегство от тела — как будто предполагает возможность отделить тело от сознания, освободить сознание и позволить ему свободно странствовать по виртуальному миру, оставив тело позади. Но сознание зависит от тела. Сегодня многие когнитивные учёные и психологи принимают концепцию воплощённого познания: наши когнитивные процессы, включая сознание и мышление, тесно переплетены с физическим телом и зависят от него. Когда мы играем или взаимодействуем в виртуальных мирах, сколь бы глубоким ни было погружение, мы остаёмся там столь же телесными, как и где угодно. И «внутри» виртуальных миров мы столь же уязвимы и смертны, как в любом другом месте. Обещанное бегство, хотя технически оно доступно всем, — иллюзия.

Идея загрузки сознания основана на том же неокартезианском дуализме и редукционизме: предполагается, будто разум и тело можно каким-то образом разделить

и будто разум не нуждается в теле. [26] Она исходит из того, что сознание и его процессы можно отделить от биологического тела, сохранив при этом личность и само сознание. Но сведение сознания к обработке данных игнорирует эмерджентную природу человеческого разума, связанную с телом и зависящую от физических процессов. В принципе его невозможно полностью свести к вычислениям. Трансгуманисты склонны предполагать, что разум — это мозг, а мозг — это компьютер. В философии ИИ такие взгляды подвергал знаменитой критике Хьюберт Дрейфус — философ, активно полемизировавший с ранними исследованиями искусственного интеллекта. Опираясь на экзистенциальную феноменологию Мартина Хайдеггера и Мориса Мерло-Понти, Дрейфус утверждал, что вычислительные модели сознания не способны учитывать телесные, контекстуально зависимые и неформализуемые аспекты человеческого опыта. Люди понимают мир посредством телесного опыта и взаимодействия со средой, формируя фон неявного знания, который невозможно свести к правилам или алгоритмам. [27]

Другой философ, Джон Сёрл, придумал мысленный эксперимент «Китайская комната», чтобы показать, что компьютерные программы могут создавать впечатление понимания языка, но не обладать настоящим пониманием. В этом мысленном эксперименте Сёрл сидит в закрытой комнате и получает под дверь китайские иероглифы, на которые отвечает, пользуясь компьютерной программой с правилами манипулирования этими символами. Сам Сёрл китайского языка не понимает, но, применяя правила и выдавая правильный результат, создаёт у людей снаружи ложное впечатление, будто в комнате находится человек, говорящий по-китайски. [28] Современный ИИ, конечно, не является основанным на правилах в том смысле, в каком его представляли и обсуждали Дрейфус и Сёрл: сегодня используются методы машинного обучения. Но программы вроде ChatGPT хорошо предсказывают текст, не понимая язык и не мысля в человеческом смысле, как и другие компьютерные программы. Сторонники загрузки сознания всё ещё, по-видимому, исходят из вычислительной модели разума и не замечают ограничений компьютеров. И нет убедительных оснований полагать, что обработка данных и компьютерная симуляция сами по себе способны породить сознание и понимание, подобные человеческим; подобную идею может поддерживать лишь редукционистское понимание сознания и личности. Не говоря уже о том, что религиозные люди едва ли обнаружат душу в машине — тем более душу любимого человека.

С философской точки зрения вообще сомнительно, сохранилась бы личность после загрузки. Даже если предположить, что загруженное и заново воплощённое сознание технически возможно, остаётся вопрос: представляло бы оно действительно исходное «я», если по сути являлось бы компьютерной симуляцией? Означало бы это продолжение существования прежнего человека — или рождение нового существа? [29] Второе кажется более вероятным, по крайней мере если понимать личную идентичность через непрерывность сознания, опыта и физического тела —

непрерывность, которая при загрузке, скорее всего, нарушается. Не менее сомнительно, возможно ли вообще воспроизвести субъективную природу человеческого опыта. Будет ли загруженное сознание чувствовать то же самое, что исходное? И следует ли вообще осуществлять загрузку, если существует хотя бы малейшее сомнение в этом? Насколько этически допустим подобный эксперимент? И что вообще значит «быть» симулированным разумом — если такой, довольно редукционистский, вопрос имеет смысл при нашей телесной и целостной природе? Более того, даже если предположить, что все эти фундаментальные трудности преодолимы, сама идея загрузки остаётся крайне спекулятивной: у нас нет науки, позволяющей сделать это, и нет признаков того, что она появится в ближайшем будущем. Наше научное и философское понимание разума и сознания всё ещё весьма ограничено. С научной и технологической точки зрения мы даже близко не подошли ни к полному воспроизведению сложности человеческого мозга, ни тем более к его полноценной симуляции. Поэтому крайне маловероятно, что этот технологический трюк с бессмертием сработает. Кроме того, он, вероятно, окажется баснословно дорогим. Загрузка сознания может стать бессмертием лишь для избранных.

Обратите внимание, что большинство христианских теологов также отвергают дуализм. Хотя меньшинство считает тело полностью отделённым от души, обычно тело понимается как существенная часть личности. Учение о воплощении обычно трактуется так, что человек представляет собой цельное единство; соответственно, воскресение означает воскресение всей личности и новую жизнь в преображённом теле. Это одна из причин, по которым не только консервативные, но и более либеральные теологи отвергают трансгуманизм или по крайней мере идеи вроде загрузки сознания. Христианство всегда отвергало и гностический дуализм. Иудаизм также утверждает психосоматическое единство личности — представление, повлиявшее на Павла, когда он формировал основы ключевых христианских верований. [30]

Возражения против дуализма и редукционизма означают также, что более демократичная и дешёвая идея цифрового воссоздания умерших в форме *deadbots* обречена на неудачу — по крайней мере если её цель действительно вернуть умершего. Сознание невозможно свести к цифровым данным, даже если речь идёт о разговорных шаблонах. Возможно, большие языковые модели, лежащие в основе таких ИИ-ботов, приближают нас как никогда к имитации разговора с другим человеком. Иллюзия совершенствуется; и LLM, и технологии цифрового моделирования личности становятся всё убедительнее. Но поскольку боты, аватары и голограммы основаны исключительно на цифровых следах, в них всегда будет отсутствовать целостный телесный человек и разговор с его воплощённым разумом. Даже если перед нами появится зловеще движущийся образ умершего, цифровой призрак не является самим человеком — и, строго говоря, даже не является настоящим призраком в традиционном смысле. Если бы дуализм был верен, призраки могли бы существовать, поскольку тогда

разум действительно можно было бы отделить от тела, душу — от трупа, дух — от материи. Некоторые упомянутые в этой главе религиозные взгляды можно истолковать как предполагающие подобный дуализм, хотя на деле они обычно сложнее. Тогда идея создания цифровых призраков выглядела бы несколько осмысленнее: по крайней мере в принципе мы могли бы отделить разум от тела, душу — от физического субстрата, если не учитывать технические трудности. Возможно, тогда нам вообще не понадобился бы *griefbot*, потому что мы могли бы напрямую общаться с умершими близкими в форме призраков. Но несмотря на наши религиозные надежды, дуализм не описывает реальность адекватно. В лучшем случае цифровое воскресение даёт временное утешение. В худшем оно этически крайне проблематично: человек мог не давать согласия на использование его личных данных; ИИ может неточно воспроизвести умершего; процесс горевания способен затянуться или осложниться. Люди могут сформировать нереалистичные ожидания и испытать сильный стресс или даже травму от встречи с образом, голосом и симулированным взаимодействием. По меньшей мере это может быть очень жутко. Кроме того, цифровые личности могут использоваться для получения прибыли и эксплуатироваться способами, не уважающими достоинство умершего. Неясно и то, как такая технология повлияет в долгосрочной перспективе на общественное отношение к смерти. Она, безусловно, может усилить западное стремление отрицать и побеждать человеческую смертность. Но эти попытки неизбежно потерпят неудачу и могут лишь увеличить страдание. Более того, они способны разрушить некоторые мудрые традиционные способы переживания смерти близких — например, ритуалы траура, — а также разумные способы примирения с перспективой собственной смерти.

Возражения против дуализма и редукционизма не спасают и крионику, которую скептики часто считают псевдонаукой. Большинство учёных считают крайне маловероятным, что мёртвые тела, даже сохранённые подобным образом, вообще можно будет оживить. Никаких доказательств возможности такого оживления нет, криопротекторы могут быть весьма токсичны, а надежда на скорое появление технологий, способных восстановить повреждённые клетки, остаётся крайне спекулятивной. Но помимо научных причин существует и серьёзное философское возражение, аналогичное критике загрузки сознания и вновь связанное с дуализмом и редукционизмом: хотя разум связан с телом, мы не тождественны своему физическому телу — или, по крайней мере, не сводимся только к нему, — и уж точно не сводимся к мёртвому телу. Как невозможно отделить разум от тела при загрузке, так невозможно отделить тело от разума в крионике. Когда тело умирает, умирает и разум. Что тогда вообще будет «воскрешено»? Даже христианская идея воскресения обычно понимается не как простое оживление мёртвого тела, а как «преображающее» воскресение, в котором духовное тело продолжает существовать новым, иным образом. [31] И, как в случае с загрузкой, неясно, сохранится ли при крионике личная идентичность. Скорее вновь будет создано новое существо — возможно, не чудовище Франкенштейна, но в

любом случае не тот самый умерший человек, хотя сходство внешности само по себе уже будет достаточно зловещим. Неясно и то, как общество будет обращаться с воскресшими людьми — и как они будут существовать в обществе. Ожидать воскресения крионизированных тел поэтому примерно столь же странно, как ждать воскресения древнеегипетских мумий: занятие, несомненно, эксцентричное, дорогое и, скорее всего, тщетное. Но одновременно вполне понятное в свете истории религиозных идей и вообще человеческой истории. Мы продолжаем бороться. Продолжаем сопротивляться. Ждём. Надеемся. Не сдаёмся. И, как у древних египтян, сохранить себя смогут прежде всего богатые и могущественные. Остальным, по-видимому, предназначено сгнить.

Наконец, саму идею *deus ex machina* — и шире, идею бога, который должен нас спасти, — можно плодотворно критиковать с помощью Ницше и апелляции к человеческой ответственности: надеяться, что нас спасёт бог, отрицать собственную свободу и уклоняться от ответственности за свои жизни, отношения и общества на земле — опасная иллюзия. [32] Если определённая религия воплощает и поддерживает такую надежду, можно утверждать, что нам не нужна эта религия — или, по крайней мере, такое её толкование. Не нужна нам и «светская религия» трансгуманизма, обожествляющая технологию.

Трансгуманисты, большинство из которых считают себя светскими учёными и философами, пропитаны христианскими идеями сильнее, чем понимают или готовы признать. И поскольку все мы в какой-то степени надеемся, что технологии избавят нас от смертности, это относится и к нам. Ницше критиковал христианство за воспитание пассивности и зависимости от божественного спасителя. Христиане, утверждал он, обладают рабской моралью и ждут спасения и искупления свыше вместо того, чтобы самим создавать собственную судьбу. [33] Их внимание устремлено в иной мир, что ведёт к отвержению мира нынешнего. Вместо того чтобы ждать помощи от бога или надеяться на загробную жизнь — цифровую или любую другую, — ницшеанская перспектива предлагает жить здесь и сейчас, взять собственную жизнь в свои руки и сделать из неё максимум возможного: преодолевать себя и, возможно, работать над тем, что мы оставим после себя, что бы это ни значило, — но в пределах человеческого существования. [34]

В конце концов Гильгамеш это понял.

Примечания

[1] Ещё одно знаменитое религиозное повествование на эту тему — история Сиддхартхи Гаутамы, Будды: знатного человека, которого поначалу оберегали от вида старости, болезни и смерти, но который, столкнувшись с ними, испытал потрясение и отвращение. Он становится монахом и в конце концов достигает просветления. Спустя многие годы он находит путь к прекращению страдания.

Интересный современный роман и фильм на эту тему — *Never Let Me Go* Кадзуо Исигуро (Faber and Faber, 2005), где один из героев, Томми, выбегает в поле и в ярости протестует против несправедливости своей судьбы, крича от гнева и отчаяния. Его ярость — протест против слишком ранней и неестественной смерти, которая скоро наступит: молодых людей выращивают ради донорства органов, и они постепенно умирают. Но её можно понимать и шире — как протест против неизбежной реальности и скандала человеческой смертности вообще, а также против того, как властные структуры определяют человеческие жизни и как люди пытаются с этим справиться, рационализируя несправедливость и принимая смерть. Название романа также отсылает к этому глубоко человеческому протесту и к попытке найти смысл и утешение в отношениях.

[2] Stephen Cave, *Immortality: The Quest to Live Forever and How It Drives Civilization* (Crown Publishers, 2012), 2.

[3] 1 Кор. 15:26.

[4] *Bhagavad Gita*, 2:20.

[5] Plato, *Phaedo*, trans. G. M. A. Grube, in Plato, *Five Dialogues*, 2nd ed., rev. John M. Cooper (Hackett Publishing, 2002), 93–154, l. 70c.

[6] Plato, *Phaedrus*, 245c, 246c.

[7] Рим. 6:4.

[8] Michel Foucault, *Discipline and Punish: The Birth of the Prison*, trans. Alan Sheridan (Pantheon Books, 1977); *History of Sexuality*, vol. 1, *An Introduction*, trans. Robert Hurley (Pantheon Books, 1978).

[9] Michel Foucault, “The Subject and Power,” *Critical Inquiry* 8 (1982): 777–795.

[10] Coeckelbergh, *Self-Improvement*.

[11] Geraci, *Apocalyptic AI*, 89.

[12] Метавселенная — термин, обозначающий виртуальные миры, в которых пользователи, представленные аватарами, взаимодействуют друг с другом и погружаются в цифровую среду. Термин стал особенно модным в 2021 году, когда Марк Цукерберг объявил, что Meta будет развивать метавселенную; позднее компания переориентировалась на ИИ.

[13] Victoria Woollaston, “We’ll Be Uploading Our Entire Minds to Computers by 2045 and Our Bodies Will Be Replaced by Machines Within 90 Years, Google Expert Claims,” *Daily Mail*, June 19, 2013.

[14] См. Kurzweil, *The Singularity Is Near*.

[15] Woollaston, “We’ll Be Uploading Our Entire Minds to Computers by 2045.”

[16] Sujita Sinha, “AI Afterlife: Digital Resurrection of Dead Raises Ethical Concerns,” *Interesting Engineering*, May 9, 2024.

[17] Areesha Lodhi, “‘Never Say Goodbye’: Can AI Bring the Dead Back to Life?,” *Al Jazeera*, August 9, 2024.

[18] Kate Bowman, “AI Ghosts: The Promise and Perils of Digital Griefbots,” *Palliative Care Volunteering*, January 16, 2025.

[19] Laurence Devillers, “Will We Live On in the Form of Virtual Avatars?,” *Polytechnique Insights*, February 7, 2024.

[20] Aubrey de Grey and Michael Rae, *Ending Aging: The Rejuvenation Breakthroughs That Could Reverse Human Aging in Our Lifetime* (St. Martin’s Press, 2007).

[21] Ray Kurzweil, “A.I. Can Radically Lengthen Your Lifespan, Says Futurist Ray Kurzweil. Here’s How,” *Fortune*, June 25, 2024.

[22] Cave, *Immortality*.

[23] Mick Brown, “Peter Thiel: The Billionaire Tech Entrepreneur on a Mission to Cheat Death,” *The Telegraph*, September 19, 2014.

[24] Brown, “Peter Thiel.”

[25] Mark Coeckelbergh, *Human Being @ Risk: Enhancement, Technology, and the Evaluation of Vulnerability Transformations* (Springer, 2013).

[26] Можно утверждать, что у самого Декарта представление об отношении души и тела было более сложным и нюансированным.

[27] Hubert L. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason* (MIT Press, 1992).

[28] John Searle, “Minds, Brains and Programs,” *Behavioral and Brain Sciences* 3 (1980): 417–457.

[29] Дискуссия о загрузке сознания и личной идентичности продолжается. См., например, работы Сьюзен Шнайдер по этой теме.

[30] Calvin Mercer and Tracy J. Trothen, *Religion and the Technological Future: An Introduction to Biohacking, Artificial Intelligence, and Transhumanism* (Palgrave Macmillan, 2021), 79, 166, 152. См. также моё изложение христианских представлений о воскресении ранее в этой главе.

[31] Mercer and Trothen, *Religion and the Technological Future*, 151–152.

[32] См. также мои замечания по этому поводу в интерпретации Винера в главе 3. Здесь я также отсылаю к знаменитому высказыванию Хайдеггера «Только бог может нас спасти» из интервью журналу *Der Spiegel* 1966 года, опубликованного посмертно. Английский перевод Уильяма Дж. Ричардсона см.: Thomas Sheehan, ed., *Heidegger: The Man and the Thinker* (Transaction Publishers, 1981), 45–67.

[33] См. прежде всего Nietzsche, *On the Genealogy of Morals* и *The Gay Science*.

[34] См. Nietzsche, *Thus Spoke Zarathustra*. Здесь я не согласен со Стефаном Лоренцем Зоргнером — и Максом Мором, — которые утверждают, что философия Ницше согласуется с трансгуманизмом. По Зоргнеру, философия Ницше, особенно идея *Übermensch*, призывает преодолеть нынешние человеческие ограничения. Как указывали критики, это неверное толкование. У Ницше речь идёт о личной трансформации и самопреодолении: возможно, о трансценденции, но личной и происходящей внутри ограничений человеческого существования. Подробнее см.: Stefan Lorenz Sorgner, “Nietzsche, the Overhuman, and Transhumanism,” in *Nietzsche and Transhumanism: Precursor or Enemy?*, ed. Stefan Lorenz Sorgner and Yunus Tuncel (Cambridge Scholars Publishing, 2017), 14–26; Max More, “The Overhuman in the Transhuman,” *Journal of Evolution and Technology* 21, no. 1 (2010): 1–4.

Заключение. *Deus ex machina*, или почему нам (как нам кажется) нужны машины

Говоря, что ИИ — это «искусственная религия», я прежде всего использую метафору. Технологии ИИ и связанные с ними сообщества не являются религиями в буквальном смысле. Но метафорическое употребление этого выражения вполне уместно и оправданно. Возьмём хотя бы сам термин «ИИ»: «искусственный интеллект» — тоже метафора, поскольку наши машины на самом деле не являются разумными, по крайней мере не в том смысле, в каком разумны люди. И всё же термин имеет смысл, потому что с самого Дартмутского семинара 1956 года, положившего начало этой области, целью было моделирование человеческого интеллекта. [1] То же относится и к выражению «искусственная религия». В этой книге я утверждал, что ИИ — как интеллектуальный проект, общественный дискурс и технологическая практика — укоренён в западной религиозной культуре и по-прежнему испытывает её влияние. Даже в так называемых секулярных обществах и культурах важно признавать и исследовать эти корни, чтобы лучше понимать, как мы мыслим и говорим о наших «разумных» машинах, как взаимодействуем с ними и как можем выработать более критическое отношение к этим способам мышления и действия. На протяжении всей книги я показывал, что религиозные и мифологические модели мышления и практики продолжают существовать и сегодня, формируя ИИ и наше отношение к нему. Религиозные повествования об автоматах для богов, брошенных творениях, чудесных машинах, богах, с которыми можно разговаривать, апокалипсисах и *pharmaka* бессмертия по-прежнему актуальны для наших экзистенциальных тревог и помогают понять, как мы говорим об ИИ и других современных «разумных» машинах и как с ними взаимодействуем. Даже нынешний, якобы полностью секулярный, политический дискурс об ИИ всё ещё несёт на себе влияние религиозного мышления в широком смысле — например, когда люди принимают трансгуманистические идеи и исходят из апокалиптических сценариев развития ИИ.

Я подчёркиваю именно этот «широкий смысл» религиозного мышления, поскольку рассматриваемые мифы и идеи религиозны не только в узком смысле — как связанные с крупными институционализированными религиями прошлого и настоящего, прежде всего древнегреческой, иудейской и христианской. Повествования и идеи, обсуждавшиеся в этой книге, возникли и обрели свои конкретные культурные формы в определённых религиозных и исторических контекстах, а затем оказали длительное и вполне осязаемое влияние в таких специфических культурных и политических средах, как Кремниевая долина. Эти формы, возможно, преимущественно западные. Но все они затрагивают более общие — и, пожалуй, более глубокие — экзистенциальные и духовные проблемы, конфликты и стремления, присущие нам как людям, независимо от того, считаем ли мы себя религиозными или принадлежим к конкретной религии.

Встречавшиеся нам истории — не только о богах и божественном. Они и о людях: о людях, мечтающих не работать; ищущих безопасности в небезопасном мире; пытающихся выстроить отношения со своими детьми и потомками — и наоборот; изумляющихся миру, включая мир науки и техники; жаждущих встретить другого, вступить в отношения, обрести близость, дружбу и любовь; стремящихся узнать собственное будущее и будущее мира; не желающих умирать и надеющихся хотя бы на какую-то форму бессмертия. Всё это не только религиозные проблемы и стремления — если понимать религию узко, — но и неотъемлемая часть человеческого существования. Это универсальные человеческие проблемы и надежды.

Любопытно и, возможно, неожиданно, что именно эти религиозные и мифологические способы мышления — вместе с лежащими под ними экзистенциальными конфликтами и надеждами — начинают давать ответы на мой исходный вопрос: как понять столь странное западное отношение к машинам? И особенно на вопросы, которые не давали мне покоя с тех пор, как много лет назад я начал заниматься этикой робототехники и ИИ: зачем вообще мы создаём и используем разумные машины? Почему нам кажется, что они нам нужны? Откуда берётся эта «потребность»? Откуда возникает эта мечта?

Ответы можно искать в уже упомянутых религиозных повествованиях и моделях мышления, которые особым образом сформировали западное технологическое воображение — в том числе западное воображение об ИИ. Если говорить грубо: чтобы по-настоящему понять Кремниевую долину, нужно понимать Иерусалим, Афины и Рим. Чтобы действительно понять и поместить в контекст то, что делают Маск, Цукерберг и Альтман, а также почему у них появляется столько последователей и власти, нужно знать Гефеста, рабби Лёва и Иисуса Христа. Современные технологии вроде ИИ, равно как стремления и фантазии их разработчиков и пользователей, невозможно до конца понять вне широкой интеллектуальной и культурной истории. Чтобы понять западные технологии, нужно понимать западную культуру, включая её мифы и религии. В этой книге я приподнял лишь угол этой ткани, показав лишь небольшой фрагмент сложного узора. Как видно из примечаний, по каждой теме уже существует гораздо более подробная научная литература. И работы ещё много: полная картина этой увлекательной религиозно-технологической истории пока далеко не раскрыта. В предыдущих главах я предложил лишь выборку наиболее значимых сюжетов и образов, помогающих войти в эту историю и осмыслить наши отношения с ИИ и сходными технологиями.

Но ответ на вопрос, почему нам кажется, что разумные машины необходимы, лежит и глубже — в тех экзистенциальных потребностях и конфликтах, которые, вероятно, и породили наши религиозные повествования и на которые отвечают связанные с ними религиозные практики, идеи, тексты и институты, чем бы ещё они ни являлись. По меньшей мере одна, а возможно, и самая фундаментальная группа причин, по которым мы создаём, используем и воображаем такие разумные машины, как ИИ, состоит в

следующем: мы не хотим работать; стремимся к защите; хотим иметь силу и возможность действовать; хотим детей или хотя бы каких-то потомков — пусть даже искусственных; жаждем чуда и тайны; ищем общения, близости, любви и руководства; хотим знать будущее; ищем всё и всех, кто способен дать нам надежду; хотим жить вечно вместе с теми, кого любим. И вновь: всё это не только «религиозные» стремления. Они связаны с более глубокими экзистенциальными структурами и отношениями, составляющими человеческое бытие. Возможно, отчасти это то, что Тиллих называл вопросами «предельной заботы», и совершенно точно это темы, с которыми религии имели дело всегда. [2] Вспомним хотя бы проблемы творца и творения или вопрос бессмертия. Конечно, все эти заботы и стремления принимают конкретные культурные формы — например, западные. Но в то же время они являются желаниями, потребностями и надеждами, отмечающими человеческое существование вообще.

Технологии вроде ИИ должны стать ответом на эти слишком человеческие желания, потребности и стремления. В этом смысле западные амбиции, надежды и мечты, связанные с ИИ, имеют экзистенциальное значение и происхождение. ИИ должен взять на себя работу, которую мы не хотим выполнять. В военной сфере ИИ принимают с энтузиазмом, чтобы автоматизировать войну и усиливать возможности солдат. ИИ должен давать силу отдельным людям и увеличивать мощь коллективов. ИИ видится шагом к появлению «детей разума», которые, возможно, восстанут, но одновременно продолжают наше путешествие. Чат-боты ИИ предлагают нам в качестве новых спутников, друзей и даже возлюбленных — якобы в ответ на нашу экзистенциальную жажду общения с Другим. ИИ воспринимается как новый оракул, способный направлять нас и предсказывать будущее. Более того, считается, что ИИ сам существенно сформирует это будущее: откроет эпоху, когда всё станет иным. И мы надеемся, что однажды он позволит жить гораздо дольше — желательно вечно. Таким образом, ИИ как технология, но одновременно как повествование, дискурс, миф и воображаемый образ воплощает многие из самых фундаментальных экзистенциальных стремлений, конфликтов, страхов и надежд человечества. Именно это делает технологию могущественной не только технически и функционально — как универсальную технологию, — но и культурно. Это объясняет, почему ИИ стал восприниматься как определяющая технология нашей эпохи и будущего. Он начинает казаться технологией, формирующей само наше время и наше будущее.

Однако вопреки господствующим детерминистским рассказам, преследующим Кремниевую долину, ИИ — и как технология, и как культурное явление — вовсе не чужое «нечто», которое действует и влияет на мир независимо от нас, людей. Помимо вопросов «откуда» и «почему», связанных с религиозными корнями ИИ, древними мифами и экзистенциальными проблемами, необходимо задавать и вопрос «кто». Вопросы «Откуда берётся эта воспринимаемая потребность в машинах?» и «Почему мы мечтаем о разумных машинах?» имеют и ответ, начинающийся с «кто». ИИ всегда связан с людьми, конкретными политическими и социально-экономическими

контекстами, интересами и повестками. Есть люди, заинтересованные в том, чтобы мы считали эти разумные машины необходимыми и покупали их, — и именно с этой целью машины и проектируются. Есть люди, которые используют машины, привыкают к ним, а затем сами оказываются используемыми ими — и, возможно, теми, кто эти технологии создаёт и ими владеет. Есть люди, желающие, чтобы мы воспринимали ИИ как *deus ex machina* и универсальное решение всех наших экзистенциальных проблем и религиозных стремлений, потому что так ИИ лучше продаётся. Есть люди, связывающие ИИ с мифологиями, которые скрывают их собственную роль и ответственность, представляя ИИ как нечто внешнее и как неизбежную судьбу. Поэтому вопрос о том, почему люди создают и используют ИИ, требует не только интерпретации и герменевтической работы, но и ведёт к новым вопросам о политике и власти. На протяжении всей книги я показывал, что соответствующие религиозные повествования, модели мышления и экзистенциальные проблемы имеют также политическое измерение.

И дело не только в том, что религия и религиозное мышление сами по себе политичны, как я неоднократно отмечал. Вспомним, например, представления о рабстве в древнегреческом и египетском мире или христианские дисциплинарные практики и техники надзора. Эти формы мышления нередко продолжают и в современном дискурсе об ИИ — например, когда мы хотим, чтобы ИИ стал нашим рабом, или когда современные государства используют ИИ, чтобы дисциплинировать нас. Политика является частью того, чем ИИ является и что он делает сегодня, а политическое измерение всегда присутствовало и в религии. Нынешний публичный дискурс об ИИ имеет политико-религиозную природу, а наши практики и надежды, связанные с ИИ, формируются древними мифами, моделями мышления и практиками, возникшими в конкретных политико-религиозных контекстах. Сегодня Big Tech создаёт новые формы исповеди, наблюдения и дисциплины и возводит новые пирамиды власти и воскресения для элиты. Если мы знаем собственную политико-религиозную историю и связанные с ней идеи, то можем увидеть это измерение власти и выработать надлежащую критическую дистанцию и реакцию.

Политика неизбежна и в том, как экзистенциальные проблемы проявляются в нашей реальной социальной и политической жизни, связываясь с политической экономией и вопросами справедливости. Нам кажется, что машины нужны, или нас заставляют в это верить по всем этим религиозным и экзистенциальным причинам. Но далеко не каждый получает желаемое, и в разных политических и экономических условиях люди получают разные вещи. ИИ по-разному влияет на различные группы, классы и людей. «Мы» можем вовсе не выиграть от ИИ — или, по крайней мере, далеко не все одинаково. Не каждому достанется собственный робот-раб, и не каждый получит от ИИ одинаковую силу. Одни — владельцы капитала, представители техноэлиты — становятся невероятно богатыми и влиятельными; другие остаются обычными работниками — «цифровыми крепостными»? — или теряют работу, а вместе с ней

надежду хотя бы на базовую материальную и социальную защищённость. Каково их будущее? Широко распространено мнение, что ИИ способен усилить уже существующее неравенство и увеличить разрыв между богатыми и бедными. Кто победит и кто проиграет в гегемонических, колониальных и эсхатологических сценариях ИИ? Что, если ИИ ухудшит социальные и экономические условия жизни большинства? Что, если Земля станет всё менее пригодной для жизни? Технологическое усовершенствование или бегство в космос, возможно, будут доступны лишь богатым и могущественным. А что будет с теми, кто останется неусовершенствованным и не сможет присоединиться к космическим приключениям сверхума и колониальным экспедициям миллиардеров, а будет вынужден остаться на Земле? Будут ли эти люди добывать сырьё, выполнять грязную работу или населять человеческий зоопарк ради развлечения усовершенствованных? Если это кажется слишком фантастическим, вспомните, что именно так мы сегодня поступаем со многими менее разумными животными. Или вспомните антиутопические сериалы «Голодные игры» и «Игра в кальмара». Это разделение уже существует в западном воображении и, к сожалению, в западной практике. Нужны ли вообще обычные смертные? Полезны ли они? Им предназначено исчезнуть? А что с теми, кто уже сегодня не допущен в Страну чудес ИИ и вынужден работать в тяжёлых условиях, создавая данные, устройства и энергию, питающие рай ИИ для немногих? Кто выиграет в Дивном Новом Мире ИИ? Кто станет царём, а кто слугой или рабом? Кто контролирует и будет контролировать искусственных потомков? Кого развлекает и обманывает ИИ — и кто на этом зарабатывает? Кто исключён из магического представления? Кого оно отвлекает? Кого эксплуатируют чат-боты и кто извлекает выгоду из этой эксплуатации? Кто подвергается и будет подвергаться манипуляции, контролю и дисциплине на основе прогнозов ИИ? Поддерживает ли ИИ новые формы капитализма или, как предлагают некоторые авторы, нам вновь следует говорить о феодализме? Стали ли владельцы Big Tech нашими новыми феодальными господами? [3] Является ли ИИ главным образом западным проектом или по крайней мере привилегией крупнейших геополитических и технополитических сил — или он принесёт пользу и Глобальному Югу? Какие пророки, компании и политики выигрывают от распространения апокалиптических историй и эсхатологических надежд об ИИ, каковы их интересы, повестки и влияние на политику и инвестиции? Кто представляет себя спасителем, защитником и проводником наций и человечества? Кто переживёт апокалипсис ИИ, а кто будет обречён погибнуть? Эти вопросы о будущем ИИ и будущем с ИИ касаются не только будущего человечества вообще, но и настоящего и будущего конкретных людей и человеческих групп в определённых политико-экономических условиях, структурах и контекстах, которые сами могут быть радикально преобразованы ИИ.

Следовательно, когда мы обсуждаем ИИ вообще и спрашиваем, почему «мы» изначально создаём и используем ИИ и другие разумные машины, на кону стоит не только конкретная технология и её будущее. Речь идёт также о нашем политическом

будущем. За *deus ex machina*, который нам преподносят как универсальное решение всех проблем, стоят люди — и у этих людей есть повестки, интересы и власть. Чтобы понять, почему мы создаём и используем ИИ, и выработать критическое отношение к дискурсам и образам, окружающим ИИ, недостаточно глубоко погрузиться в историю религиозных идей и практик и проследить их продолжение в современном мышлении о технологиях и в современных технологических практиках. Необходим также более систематический анализ связей между ИИ, религией и властью. В этой книге я предложил лишь некоторые наблюдения и намёки; и здесь требуется гораздо больше работы. Нам необходима и большая политическая осознанность. Если мы не будем уделять этому достаточно внимания, то рискуем почти во сне войти в будущее, где посредством ИИ наши глубочайшие религиозные и экзистенциальные тревоги будут эксплуатироваться в масштабе, какого история ещё не знала.

Познай самого себя

Сократический — и изначально дельфийский — призыв «Познай самого себя» представляет собой древний духовный афоризм и одновременно один из краеугольных камней западной философии. В контексте размышлений об искусственном интеллекте он означает по меньшей мере две вещи, связанные с двумя видами невежества: незнанием собственных культурных и религиозных корней и незнанием политики.

Во-первых, чтобы действительно понимать, что мы говорим об ИИ и что делаем с ИИ, необходимо знать, чем является ИИ и кто такие «мы». Знать ИИ — значит понимать не только саму технологию, но и культурные и религиозные корни наших слов и действий. Нужно знать истории, мифы. Нужно знать, как мыслили наши предки. Нужно понимать, как эти способы мышления не только проявлялись в конкретных культурных и исторических условиях, но и в конечном счёте отвечали на фундаментальные экзистенциальные и универсальные вопросы, стоящие перед человеком. В этом смысле знание ИИ требует самопознания. Поэтому вопрос об ИИ — не только вопрос о технологии, но и вопрос о нас, о человеке, о человеческом существовании и о конкретных культурных и религиозных традициях, которые по-своему осмысляют это существование. Осознав это, мы можем отправиться в путь, способный дать новые представления о наших отношениях с ИИ и, шире, о современном технологическом состоянии человека. Одним результатом этого путешествия может стать признание человеческой конечности и создание технологий, помогающих с ней жить, а не пытающихся её преодолеть. Другим — обучение у определённой традиции тому, как выстраивать отношения творца и творения без токсичности. В историях можно находить вдохновение для этики технологий. Возможно, мы захотим применить такую этику к разработке и использованию ИИ. Но от Сократа и его духовного источника — дельфийского оракула — можно научиться тому, что самопознание и знание собственных ограничений предшествуют этике и являются скорее её предварительным условием. Мы должны начать с признания

собственного невежества в культурных и религиозных вопросах, чтобы по-настоящему открыться изучению истории идей и осознать, что эти идеи всё ещё могут формировать современное мышление и способы отношения к собственному существованию — включая то, как мы создаём, используем и обсуждаем технологии. Многие люди, говорящие об ИИ, не осознают культурных и религиозных фоновых влияний и структур, формирующих их мышление. В других работах я называл эти фоновые структуры «грамматикой» ИИ, а ранее в этой книге говорил о «коллективном религиозном бессознательном» и о моделях мышления, которые всё ещё «преследуют» нас. [4] В этом смысле люди действительно не знают, что говорят и что делают. Первый шаг — признать это незнание; лишь после этого становится возможной достаточно критическая этика технологии и, в конечном счёте, создание по-настоящему этического ИИ и других этических технологий.

Раз уж речь зашла о признании религиозного фона ИИ, добавлю одно замечание относительно дискуссии о секуляризации. Поскольку эта книга раскрывает религиозную грамматику ИИ, некоторые могут прочесть её как вклад в проект демифологизации и секуляризации ИИ — в том смысле, что ИИ следует очистить от этих моделей и влияний. Я сочувствую такому проекту. Эту книгу можно прочесть как критику религии ИИ, вдохновлённую, среди прочих, Фейербахом и Ницше. Но необходимо и предостережение. Независимо от того, желательны ли демифологизация и секуляризация, сомнительно, насколько вообще возможно их осуществить полностью. Мы, безусловно, можем и должны выработать гораздо большую критическую дистанцию, но было бы иллюзией думать, будто ИИ можно полностью очистить от религии, мифа и других культурных влияний. Миф и религия никуда не исчезнут и будут продолжать влиять на ИИ, как и на всё остальное. Если Big Tech удаётся опираться на эти повествования, модели и влияния, чтобы сохранять власть и манипулировать нами, то лишь потому, что они уже присутствуют в общественном сознании, а присутствуют они там потому, что глубоко встроены в западную культурную грамматику и тонко связаны с нашими глубочайшими экзистенциальными потребностями. Поэтому главная задача состоит не столько в полной очистке от таких повествований, сколько в том, какие истории мы хотим — или не хотим — сохранять и как можем критичнее относиться к собственным коллективным религиозным моделям мышления.

Парадоксально, но подлинная секуляризация технологии — если она вообще возможна и желательна — требует большего, а не меньшего интеллектуального взаимодействия с этими моделями мышления. Те, кто хочет создать, например, светский «технологический гуманизм» или «цифровой гуманизм», должны изучать культурно-религиозный фон технологий, чтобы лучше сформулировать собственный проект. [5] Мне также кажется, что мифы всегда будут нужны нам, чтобы понимать мир и наделять его смыслом — даже и особенно в современную эпоху; эта мысль созвучна Ницше, Юнгу, Хайдеггеру, Блюменбергу, Кэмпбеллу, Рикёру и другим. Тот факт, что

некоторые мифы использовались и используются для угнетения людей, поддержки войн и стимулирования колонизации, — ещё одна причина изучать их критически и пытаться создавать иные мифы или предлагать иные интерпретации уже существующих. Мифы, как и технологии, политичны. Избегать этой темы — не значит решать проблемы и исследовать позитивные возможности, которые открывают мифы. Более того, внимание к мифу может помочь нам увидеть пределы ИИ и избежать *hubris* — гордыни, заставляющей верить, будто всё можно вычислить, контролировать и предсказать и будто в конечном счёте мы сами способны стать богами или уподобиться им. [6] И если я прав, говоря о глубоком переплетении технологии и культуры, то по-настоящему новые технологии можно представить лишь вместе с новой культурой — а это всегда возможно только в ограниченной степени, поскольку мы глубоко погружены в собственную культуру. Ошибка модерности состояла, среди прочего, в том, что она культивировала неведение об этом ограничении, создавая ложное впечатление, будто можно начать с чистого листа. Поэтому этика ИИ, как и всякая этика технологии, всегда должна начинаться с определённой точки. А эта общая исходная точка требует коллективного самоанализа.

При этом книга всё же отчасти демифологизирует ИИ: описывая лежащие под ним религиозные повествования и модели мышления, она уменьшает наше неведение относительно их продолжающегося влияния и позволяет отстраниться от этих моделей — а также от тех, кто использует их в собственных интересах. Например, мы можем критиковать такие модели, как ожидание, что ИИ нас спасёт; представление об ИИ как о разрушительном и злобном боге; и вообще склонность приписывать ИИ божественные свойства. Мы можем также критиковать мечту о хозяине и рабе, апокалиптическое повествование о сверхразуме, иллюзию искусственной близости или миф технологического бессмертия. Мы способны хотя бы частично отстраниться от этих историй, если понимаем, насколько глубоко они укоренены в западных культурных и религиозных традициях и насколько трудно их полностью оставить позади, а также если признаём, что они, вероятно, и дальше будут сохранять часть своей соблазнительной силы и скрытого влияния на технологическую культуру. Но отчаиваться не нужно: определённые изменения возможны. ИИ не бог, если только мы сами не сделаем его богом — а к этому, к сожалению, у нас, похоже, есть склонность. Как я постоянно подчёркиваю и в других работах об ИИ, искусственный интеллект всегда связан с людьми, а значит, нет причин относиться к нему детерминистски или фаталистически. Поскольку ИИ создают и используют люди, в принципе мы можем выбрать иной курс и направить ИИ и другие технологии в другие, более благоприятные стороны. Именно эта предпосылка делает возможной этику ИИ; отказ от технодетерминизма — ещё одно необходимое условие этики искусственного интеллекта.

Во-вторых, однако, осуществление таких изменений само связано с трудностями и приводит нас ко второму типу невежества — политическому. Мы должны понимать,

что часто политически невежественны и наивны — особенно когда речь идёт о технологиях, которые на первый взгляд кажутся не связанными с политикой. Модерность приучила нас категорически отделять политику от природы, а размышления о человеке, культуре и власти — от физики, инженерии и компьютерных наук. Но такой образ мышления ошибочен и даже опасен. Человеческое существование неизбежно имеет социальное и политическое измерение, как и наши размышления об использовании и развитии технологий, включая этику технологий. Здесь путь самопознания состоит уже не в том, чтобы исследовать человека и человеческое существование абстрактно или изучать конкретную философскую или религиозную традицию и её мифы как таковые, а в том, чтобы увидеть и признать политические аспекты этих способов мышления и, ещё глубже, осознать, что мы как социальные существа всегда существуем внутри определённой системы властных отношений. [7]

Когда мы каким-либо образом имеем дело с человеческим существованием, создаём технологии и повествования о технологиях, мы всегда связаны с другими и находимся в определённом политическом, социальном и экономическом контексте со своей собственной динамикой власти. Если мы хотим использовать, создавать или обсуждать ИИ, следует признать, что делаем это из конкретного политического контекста — например, сформированного капитализмом и неолиберализмом и расположенного в определённой части мира, скажем на Западе или Глобальном Севере. Всё, что мы делаем с ИИ или говорим о нём, может быть обусловлено этим контекстом и неизбежно взаимодействует с полем интересов и власти.

Если мы не замечаем политического измерения ИИ, нами легко манипулировать и управлять; нас могут эксплуатировать системы, в которых мы живём, и те, кто от этих систем выигрывает; нас могут подтолкнуть говорить или делать то, что в пределах нашего собственного интеллектуального мира кажется правильным и истинным, но оказывается недостаточно чувствительным к тем людям, к которым обращён наш дискурс или для которых создаётся технология. Когда мы говорим об ИИ или создаём его, другие могут ответить. Коммуникация, как и технология, всегда политически рискованна. Даже академические тексты не могут полностью избежать этого риска. Многие разработчики ИИ и владельцы компаний в сфере ИИ политически наивны, что приводит к проблемам, когда сказанное или созданное ими попадает в политическое минное поле за пределами их собственного пузыря. Здесь знать ИИ — то есть знать себя — означает понимать собственное политическое и экономическое положение относительно других и относительно социально-экономической системы: например, своё положение разработчика или пользователя ИИ, а также работника, менеджера или владельца внутри капитализма ИИ. Многие участники публичного обсуждения ИИ, например, говорят о нём из привилегированной позиции. Следуя Марксу, можно добавить: познать себя — значит знать свой класс. Промышленная революция была не только технологическим, но и идейным и социальным событием. Следует понимать, что и сегодня — в том числе посредством ИИ — сохранение древних религиозных

моделей мышления и универсальных человеческих экзистенциальных слабостей может использоваться для легитимации и сохранения определённого общественного порядка — капитализма, колониализма, патриархата и так далее, — для воспроизводства определённых социально-экономических отношений, эксплуатации и исключения людей. Это может также поддерживать вредные идеологии и проблематичные мировоззрения. Любой гуманизм, желающий что-то с этим сделать, должен достичь достаточного уровня самопознания относительно соответствующего социально-экономического и культурного фона и его переплетения с технологией и властью — включая проблематичные элементы самого гуманизма, например крайние формы антропоцентризма. Некоторые из этих моделей мышления и технологий уже приводили и продолжают приводить к эксплуатации нечеловеческих существ, насилию над ними, разрушению окружающей среды и изменению климата. Гуманизм XXI века может быть лишь гуманизмом, ставящим под вопрос жёсткие формы антропоцентризма, как я утверждаю в других работах. [8] Нам необходимо размышлять о том, кто мы, где находимся и куда движемся в отношении всех этих проблем. Цифровая трансформация переплетена с социальными, политическими и экологическими изменениями; мы — как мыслители об ИИ — и сам дискурс об ИИ находимся прямо внутри этих бурь и преобразований.

Без самосознания и самопознания в обоих смыслах — культурном и политическом — разработка и использование ИИ, равно как и дискурс о нём, всё равно будут связаны с этими культурными моделями и структурами власти, но уже без критической дистанции и рефлексии. Недостаток образования в этих вопросах способен привести к тому, что сегодня, перефразируя Ницше, называют «стадным мышлением». [9] Если мы пользуемся социальными сетями на основе ИИ, чтобы высказывать своё мнение об ИИ и участвовать в публичных дискуссиях, но при этом не обладаем ни культурным фоном, ни политической осознанностью, мы рискуем просто повторять то, что говорят и делают другие. Размещая посты об ИИ или делая прогнозы относительно его будущего, мы тогда всего лишь следуем за стадом. Мы превращаемся в аватары безжизненных мыслей, окаменевших ценностей и корыстных интересов. Иными словами, становимся похожими на собственных чат-ботов.

И тогда действительно, превратившись в стадных животных без каких-либо критических ресурсов, мы сами загоняем себя в положение, где спасти нас может лишь машина — или бог.

Немецкий философ Ханс-Георг Гадамер, один из ключевых представителей герменевтики XX века, возможно, был прав, считая надежду фундаментальной чертой человеческого существования. [10] Она действительно, по-видимому, является одной из важнейших черт западной культуры. Но, прислушиваясь к Ницше и Марксу, мы должны понимать, что надежда бывает и крайне проблематичной. Она может не приводить к переменам и даже препятствовать им: может стать опиумом, поддерживающим статус-кво. Конкретная форма надежды, сложившаяся в западных

религиозных и философских традициях и под их влиянием — в современном дискурсе об ИИ, рискует превратиться в наркотик. Она может помешать настоящему освобождению и трансценденции — не через трансгуманизм или Сингулярность, а той форме самопреодоления, которую Ницше считал доступной нам, смертным. [11] И она может медленно, но верно привести нас к состоянию всеобщего стадного мышления, рабства и господства — при помощи ИИ и других передовых технологий, которые ложно обещают спасение и направляются ложными пророками технологического апокалипсиса.

Мы должны действовать, чтобы не оказаться в столь мрачном положении, где остаётся лишь надеяться на *deus ex machina*, который нас спасёт — а затем, если что-то пойдёт не так, на ещё один *deus ex machina*, который спасёт нас уже от первого. Более глубокая интеллектуальная работа по изучению корней наших технологических культур — западных и иных, — обсуждение того, что эти корни означают для ИИ, а также образование, повышающее культурные знания и политическую осознанность в отношении современных технологий и общества, могут помочь выработать более критическое отношение к искусственному интеллекту. В более широком смысле такая интеллектуальная работа и образование способны помочь человечеству достичь самостоятельности мышления, которую Кант считал необходимым условием Просвещения, и избежать стадной ментальности, столь резко критикуемой Ницше и многими мыслителями XX века. Чтобы сегодня познать самих себя, мы должны осмелиться узнать — *sapere aude* — кто мы, чем являемся и кем стали в результате религиозных повествований, политической истории и современных технологий. Парадоксальным образом именно признание ограниченности нашей автономии в этом смысле может дать нам критическую дистанцию и свободу не только для развития нового, более самостоятельного мышления, но и для того, чтобы представить технологии, которые действительно нам нужны.

Примечания

[1] Дартмутский летний исследовательский проект по искусственному интеллекту 1956 года обычно считается событием, положившим начало области ИИ.

[2] Tillich, *Dynamics of Faith*, 1–4.

[3] См. также Varoufakis, *Technofeudalism*.

[4] Coeckelbergh, “The Grammars of AI.”

[5] Epstein, *Tech Agnostic*, 243.

[6] См. Blumenberg, *Work on Myth*.

[7] Здесь на меня повлиял Фуко.

[8] См., например: Mark Coeckelbergh, “What Is Digital Humanism? A Conceptual Analysis and an Argument for a More Critical and Political Digital (Post)humanism,” *Journal of Responsible Technology* 17 (2024).

[9] См. также Nietzsche, *Beyond Good and Evil*.

[10] В интервью за несколько недель до смерти в 2002 году Гадамер сказал: «Человек не может жить без надежды». См.: Jean Grondin, “Gadamer’s Hope,” *Renascence: Essays on Values in Literature* 56, no. 4 (2004): 287–288.

[11] Это ещё один пункт моего несогласия с Зоргнером.