СОДЕРЖАНИЕ

ЧАСТЬ III **2012–2019 годы**

ЧАСТЬ IV **2019–2024 годы**

ПРОЛОГ

Теплым ноябрьским вечером 2023 года знаменитый венчурный инвестор Питер Тиль устроил праздничный ужин в честь дня рождения Мэтта Данзайзена. Гости собрались в авангардном японском ресторане YESS, разместившемся в историческом здании банка начала XX века в Арт-квартале Лос-Анджелеса. В просторном зале, напоминавшем древний храм, рядом с Тилем сидел его друг Сэм Альтман[1]. Тиль вложился в первый венчурный фонд Альтмана более 10 лет назад и с тех пор оставался наставником молодого инвестора, ставшего генеральным директором OpenAI и лицом революции в области искусственного интеллекта. Запущенный годом ранее ChatGPT не только вывел акции высокотехнологичных компаний из затяжного пике, но и обеспечил им один из самых успешных периодов за последние десятилетия. И тем не менее Тиль был встревожен.

Задолго до знакомства с Альтманом Тиль взял под крыло другого юного гения, одержимого искусственным интеллектом, — Элиезера Юдковского. Тиль финансировал его исследовательский институт, работавший над тем, чтобы сделать ИИ дружественным к людям прежде, чем тот создаст искусственный разум, превосходящий человеческий. Но, по мнению Тиля, Юдковский превратился в «законченного пессимиста

и технофоба», сведя все свои мрачные прогнозы к простой формуле: «Единственное, что нам остается, — ездить на фестиваль Burning Man, ожидая, когда искусственный интеллект явится нас убить». В марте Юдковский опубликовал колонку в *Time*, где утверждал, что, если не остановить нынешние исследования в области генеративного ИИ, «все население Земли будет в буквальном смысле обречено на гибель»[2].

«Ты даже не представляешь, до какой степени Элиезер промыл мозги половине твоих сотрудников, — предостерег Тиль Альтмана. — Тебе стоит отнестись к этому серьезнее».

Альтман рассеянно ковырял свое вегетарианское блюдо, сдерживая желание закатить глаза. Не в первый раз за ужином Тиль предупреждал его о том, что компанию захватили «эффективные альтруисты» — последователи философского течения, родственного утилитаризму. В последнее время эти ребята переключились с борьбы против глобальной нищеты на предотвращение геноцида человечества взбесившимся ИИ. Тиль не уставал повторять, что «борцы за безопасный ИИ» погубят OpenAI. Он поддерживал компанию с первого дня — сначала субсидировав ее из личных средств в 2015 году, когда она была крошечной некоммерческой лабораторией, а затем, в начале 2023 года, через свой Founders Fund, после того как OpenAI обзавелась коммерческой дочерней компанией для привлечения миллиардных инвестиций от Microsoft и других организаций. Впрочем, Тиль славился своими апокалиптическими прогнозами — как шутили в Кремниевой долине, он верно предсказал 17 из последних двух финансовых кризисов.

«Ну, отчасти так было с Илоном, но мы от него избавились», — ответил Альтман, намекая на неприятное расставание в 2018 году с сооснователем компании Илоном Маском, который как-то сказал, что, создавая искусственный интеллект, мы

«призываем демона»[3]. «Потом была история с Anthropic, — продолжил он, имея в виду более десятка сотрудников OpenAI, которые в конце 2020 года, разочаровавшись в Альтмане, ушли создавать собственную конкурирующую лабораторию. — Но и с этим мы разобрались». Теперь в компании работало более 700 человек, и они, как на ракете, неслись вверх, навстречу возможности прикупить себе домик на побережье — в ближайшее время готовилось предложение по выкупу акций у сотрудников, по которому OpenAI оценивалась более чем в 80 миллиардов долларов. Паниковать не было причин.

Альтман давно превратил оптимизм в свою фирменную черту. Впрочем, на его месте кто угодно испытывал бы такое же воодушевление. Похожий на эльфа 38-летний руководитель завершал лучший в своей блестящей карьере год — год, когда его имя гремело повсюду, сенаторы ловили каждое его слово, а сам он встречался с президентами и премьер-министрами по всему миру. И — что в Кремниевой долине ценится превыше всего — создал технологию, которая, похоже, действительно могла изменить все.

В ноябре 2022 года OpenAI запустила ChatGPT — чат-бота с пугающе человекоподобным интеллектом (название расшифровывается как generative pre-trained transformer — «генеративный предварительно обученный трансформер»), который мгновенно стал хитом: сервис привлек 100 миллионов пользователей менее чем за квартал, побив все рекорды скорости роста[4]. А спустя всего несколько месяцев OpenAI представила еще более мощную версию — GPT-4, способную сдать квалификационный экзамен на адвоката и получить высший балл на сложнейшем тесте по биологии. Столь умопомрачительная динамика наводила на мысль, что амбициозная цель компании по безопасному созданию первого в мире

общего искусственного интеллекта (AGI) может быть вполне достижима.

Даже самые убежденные скептики — включая профессора информатики из Стэнфорда, который в разговоре со мной пренебрежительно назвал первую версию ChatGPT дрессированной собачкой, — начали смягчать свою позицию. В течение нескольких месяцев, когда американские компании спешно создавали рабочие группы по ИИ, пытаясь оценить потенциальный рост производительности, мы словно шагнули в научно-фантастический рассказ, автором которого был Альтман.

Сам Альтман код не писал. Он был провидцем, вдохновителем новых идей и мастером заключать сделки. Его уникальный дар, отточенный годами консультирования, а затем руководства престижным стартап-акселератором Y Combinator, заключался в умении взяться за практически неосуществимую идею, заразить других верой в возможность успеха, а затем привлечь столько денег, что эта идея становилась осуществимой. «Он единственный среди всех моих знакомых, кто берется исключительно за проекты, способные перевернуть мир, даже если шансы на успех — всего 1%», — говорил Али Рогани, управлявший инвестициями в Y Combinator, когда Альтман был президентом компании.

Пожалуй, никто лучше Альтмана не воплощал главный принцип Кремниевой долины — «добавь ноль справа». Этот образ мышления он перенял у своего наставника Пола Грэма — блестящего программиста, предпринимателя и эссеиста, одного из сооснователей Y Combinator. Грэм почти всегда помогал стартапам отказаться от мелочного мышления и задуматься, как вывести бизнес-модель на более высокий уровень — чтобы в прогнозах выручки строчка «миллионы» превратилась в «миллиарды». В 2019 году, встав у руля OpenAI, Альтман изложил

в блоге свою философию успеха: «Полезно стремиться добавить еще один ноль справа к любому показателю своих достижений — будь то деньги, статус, влияние на мир или что угодно еще»[5].

Тиль признавался: Альтман привлек его тем, что, как никто другой, воплощал дух Кремниевой долины той эпохи — миллениал 1985 года рождения, попавший в мир технологий в золотой момент между крахом доткомов и финансовым кризисом, когда вера в стартапы возродилась, но технологическая сфера еще не скатилась в то, что Тиль язвительно называл заезженной колеей. В середине 2010-х, когда все громче звучали разговоры о технологическом пузыре, Тиль подавил сомнения, доверившись инвестиционному чутью Альтмана, — и не прогадал. «Сэм был неисправимым оптимистом — качество, необходимое для инвестирования в эти компании, ведь все они казались переоцененными», — говорил он. Как выяснилось, у «единорогов», взращенных в Y Combinator, таких как Stripe и Airbnb, еще оставался колоссальный потенциал для роста. На протяжении почти всей взрослой жизни Альтмана, если не считать короткой паузы после финансового кризиса 2008 года и начавшейся в 2020 году пандемии, технологические рынки знали только один путь — вверх.

Но на этот раз мрачные предчувствия Тиля оказались пророческими. Пока два инвестиционных партнера поднимали бокалы в самом модном новом ресторане Лос-Анджелеса, четверо из шести членов совета директоров OpenAI (в их числе двое напрямую связанных с движением эффективных альтруистов) тайно обсуждали по видеосвязи увольнение Альтмана. И хотя сам Юдковский не имел к этому отношения, его влиятельный блог LessWrong сыграл ключевую роль в том, что страх перед экзистенциальной угрозой со стороны ИИ стал центральной темой движения эффективных альтруистов.

Отголоски этого страха прослеживались и в принципах самой OpenAI, чьей заявленной миссией было «развивать цифровой интеллект так, чтобы это принесло максимальную пользу всему человечеству, не ограничиваясь необходимостью получать финансовую выгоду». Еще отчетливее эти опасения проступили в необычном уставе OpenAI 2018 года. В нем говорилось, что компания «обеспокоена тем, что разработка AGI на поздних стадиях может превратиться в гонку, где не останется времени на необходимые меры безопасности», поэтому обязуется «прекратить конкуренцию и начать помогать» любым проектам, разделяющим ее ценности, если они первыми создадут AGI. Самым причудливым образом этот страх воплотился в структуре управления компании: коммерческим подразделением руководил некоммерческий совет директоров, который был обязан действовать в интересах не инвесторов, а «человечества»[6]. Инвесторов же предупреждали, что они могут потерять все вложенные средства, если совет сочтет это необходимым для выполнения высшей миссии.

Устав во многом опирался на «Принципы ИИ», разработанные на конференции 2017 года. Ее организовал Институт будущего человечества — некоммерческая организация, занимавшаяся предотвращением экзистенциальных рисков от ИИ, которую финансировали такие миллиардеры, как Маск и основатель Skype Яан Таллинн. Альтман присутствовал на конференции и поддержал эти принципы. Еще раньше, в 2015 году, став одним из основателей OpenAI, Альтман писал в своем блоге, что AGI — это «вероятно, величайшая угроза существованию человечества», и рекомендовал книгу «Искусственный интеллект. Этапы. Угрозы. Стратегии» Ника Бострома*, философа

---

* Бостром Н. Искусственный интеллект. Этапы. Угрозы. Стратегии. — М.: МИФ, 2016.

из Оксфордского университета, который на протяжении многих лет был частым гостем конференций MIRI, организованных исследовательским институтом Юдковского[7].

Опасения по поводу безопасности ИИ, о которых говорил Бостром, — например, история об искусственном интеллекте, запрограммированном делать канцелярские скрепки, который уничтожает человечество не из злого умысла, а потому, что люди мешают его стремлению превратить всю материю во Вселенной в скрепки (эту страшилку он позаимствовал, видоизменив, у Юдковского), — сыграли решающую роль в привлечении в OpenAI лучших в мире исследователей искусственного интеллекта. Не в последнюю очередь потому, что Маск, разделявший эти опасения, вложил в проект свои средства. И даже празднуя триумф ChatGPT, Альтман не забывал осторожно вплетать в свои восторженные прогнозы предостережения о потенциальных угрозах — во время слушаний в сенате он просил сенаторов заняться регулированием ИИ, потому что «если с этой технологией что-то пойдет не так, последствия могут быть катастрофическими»[8].

Многие в индустрии считали это не более чем искусным маркетинговым ходом. И действительно, в 2023 году Альтман в основном занимался тем, что у него получалось лучше всего: заключал сделки с новыми инвесторами, подогревал интерес прессы и выступал в роли глобального пророка будущего невиданного процветания. Возможно, его и правда тревожило, куда все движется, но педаль газа он выжимал до упора.

Однако вопреки мрачным предположениям Тиля члены совета директоров OpenAI совещались в тот вечер не из-за страха, что компания слишком быстро приближается к созданию AGI. На самом деле их желание сместить Альтмана почти не было связано ни с эффективным альтруизмом, ни с экзистенциальными

рисками. Причина крылась в том самом качестве, за которое Тиль так превозносил Альтмана, — в том, что тот был живым воплощением духа Кремниевой долины.

А дух Кремниевой долины начала XXI века, который стал визитной карточкой Y Combinator и который так ярко воплощал Альтман, заключался в том, что основатели компаний — короли, императоры, боги. Венчурные инвесторы, финансировавшие стартапы, из кожи вон лезли, доказывая свое дружеское расположение к основателям, что на практике означало: не увольнять основателя с поста генерального директора и не вставлять ему палки в колеса. Тиль довел эту идею до логического завершения, назвав свою венчурную фирму Founders Fund («Фонд основателей») и заявив, что она никогда не уволит основателя компании[9]. Отношения между венчурным инвестором и основателем считаются важнее даже успешности самого стартапа. В конце концов, если этот бизнес не выстрелит, основатель всегда может привлечь новый пакет инвестиций и начать что-то другое — и в этот раз достичь десятизначной оценки.

Альтман возглавил Y Combinator в 2014 году, и к этому времени основатели компаний настолько укрепили свои позиции, что он даже опубликовал в блоге пост с мягкой критикой практики, при которой основатели стартапов брали деньги инвесторов, но не давали им мест в совете директоров. Альтман отмечал, что на самом деле некоторые опытные венчурные инвесторы могли бы помочь компаниям своими знаниями. Впрочем, опасаясь, что стартапы Y Combinator решат, будто он зашел слишком далеко, Альтман завершил пост словами: «Хорошая идея — сохранять достаточный контроль, чтобы инвесторы не могли вас уволить»[10].

Компания OpenAI должна была стать исключением из этого правила. У Альтмана не только не было суперпривилегированных

акций, которые гарантировали бы ему пожизненную власть, как у Марка Цукерберга, — у него вообще практически не было акций. Альтман принял такое беспрецедентное решение еще на заре существования OpenAI: сначала потому, что компания была некоммерческой, а затем — чтобы остаться в совете директоров, не нарушая требование устава о том, что большинство директоров должны быть независимыми, то есть не владеть акциями компании. Он утверждал, что отсутствие у него власти — важнейший способ обеспечить подотчетность.

Однако совет директоров обнаружил: Альтман, привыкший к быстрым маневрам в непрозрачном мире венчурного капитала, сосредоточил в своих руках такое реальное влияние, что они не могли нормально работать. Через пять дней после ужина в японском ресторане, организованного Тилем, Альтман был уволен из компании, которую сам же и основал. Официальная причина: он «не всегда был искренен с советом директоров».

Я познакомилась с Альтманом за восемь месяцев до этого, в самый разгар первой волны ИИ-безумия. Редакция *The Wall Street Journal* командировала меня в Сан-Франциско взять интервью у основателя OpenAI в его штаб-квартире, расположенной в районе Мишн — этом причудливом, колоритном уголке города, где я когда-то, еще студенткой, подрабатывала летом в кофейне. Тогда, во времена первого пузыря доткомов, стены района пестрели самодельными листовками «Смерть яппи!», призывавшими айтишников убираться прочь и оставить панков и фриков в покое. За прошедшие годы, после того как Мишн захлестнула новая технологическая лихорадка, этот район превратился в подобие Бруклина, только с более вкусными буррито.

В воздухе витало что-то особенное. Стояла середина марта — то унылое время года, когда в Нью-Йорке уже почти перестаешь верить в существование солнца, но Сан-Франциско сиял. После

пяти дождливых дней воздух стал кристально чистым, выглянуло солнце. Вдоль шоссе 101 каждый билборд рекламировал какую-нибудь новинку в области ИИ. Вечерние новости начинались с сенсационных репортажей о том, как GPT-4 успешно сдал LSAT — вступительный экзамен в юридические вузы. Изящные цветочные композиции, кофе с нотками черники и бездомные на каждом углу — таким меня встретил Мишн.

Офис OpenAI располагался в невзрачном здании бывшей майонезной фабрики на некогда промышленной окраине Мишн, на границе с районом Потреро-Хилл. Никаких вывесок, намекающих на то, что скрывается внутри, — решение, казавшееся тогда нелепым, но обретшее смысл месяц спустя, когда Юдковский призвал бомбить дата-центры, чтобы остановить вышедший из-под контроля ИИ. Подойдя к зданию, я наткнулась на растерянного инвестора с бейджем на шее, который занимался тем же, что и я: метался между безымянной дверью и таким же безымянным грузовым въездом за углом, не в силах поверить, что через один из них можно попасть в самую захватывающую и внушающую трепет компанию Кремниевой долины. Ситуация не улучшилась, когда мы спросили охранника у грузового въезда, там ли мы находимся, — он наотрез отказался отвечать.

В итоге нам все-таки удалось попасть внутрь — в холл, больше всего напоминавший гибрид оранжереи и спа-салона. Со всех сторон нас окружали суккуленты и папоротники. Журчание каменных фонтанчиков сливалось с приглушенным гулом разговоров венчурных капиталистов, собравшихся на каком-то нетворкинг-мероприятии, посвященном инвестиционной экосистеме в сфере ИИ.

Через некоторое время в комнату вихрем влетел Альтман — в ослепительно-белых кроссовках и с располагающей улыбкой. Он выглядел моложе своих 37 лет благодаря ямочкам

на щеках. Первое, что бросается в глаза при встрече с Альтманом — и на чем заостряли внимание все ранние материалы о нем, — это его комплекция: тощий как жердь и невысокий, всего метр семьдесят с небольшим. Второе — его пронзительные зеленые глаза, которыми он смотрит прямо на собеседника, словно говорит с самым важным человеком на планете. Он извинился за задержку — предыдущая встреча затянулась. На экране в конференц-зале она до сих пор значилась под названием «Манхэттенский проект ИИ».

«Мы обсуждали, как можем еще больше внимания уделять элайнменту — то есть приведению целей ИИ в соответствие с человеческими ценностями, — пояснил Альтман, когда его спросили об этом зловещем названии. — Можем ли мы наладить более слаженную работу с другими группами по мере развития возможностей систем, чтобы решить проблему безопасности AGI? У нас есть несколько любопытных идей на этот счет».

Наш разговор состоялся спустя два дня после выхода GPT-4. Альтман стоял во главе организации, занимающейся таким эпохальным делом, что обычные продажи и прибыли казались мелочью. Он явно наслаждался необычностью сложившейся ситуации.

«Мне невероятно повезло — в начале карьеры я заработал больше денег, чем мне когда-либо понадобится, — сказал он. — Я хочу работать над тем, что считаю интересным, важным, полезным, заметным и критически важным, — над тем, что необходимо сделать правильно. Но больше денег мне не нужно. К тому же со временем мы, думаю, примем кое-какие решения, которые будут… — Альтман замялся, подбирая слово, — странными».

Он нарисовал картину будущего, в котором население планеты голосует за то, каким должен стать искусственный интеллект.

«Мы искренне хотели бы, чтобы эта технология управлялась всеми и чтобы ее преимуществами пользовались все, — сказал он. — Если нельзя реализовать это как государственный проект — а я думаю, есть масса причин, почему это может оказаться плохой или попросту неосуществимой идеей, — то разумным подходом кажется некоммерческая организация». Его определение безопасного AGI получилось довольно размытым. Когда Альтмана спросили, что для него означает безопасность ИИ, он заговорил о будущем, где «подавляющее большинство людей в мире живут намного лучше, чем жили бы в мире без AGI». Для многих это означало бы смену профессии. «Я не думаю, что большинство людей любят свою работу», — сказал он.

Лишь раз за все наше двухчасовое интервью альтруистическая маска сползла с его лица, обнажив лик беспощадного конкурента. В течение предыдущего месяца и Google, и Anthropic объявили о скором запуске собственных генеративных ИИ-чатботов, и складывалось впечатление, что отрасль вступает именно в ту самую гонку вооружений в сфере ИИ, которой так страшился устав OpenAI. Но на вопрос о конкуренции Альтман лишь бросил: «Ну, они поторопились с пресс-релизами» — и добавил: «Очевидно же, что сейчас они отстают».

Несмотря на этот краткий приступ самодовольства, офисы OpenAI, по которым он нас провел, излучали какую-то сектантскую добродетельность — как элитная частная школа со своим кодексом чести. Здесь была столовая, гостиная в духе 1980-х, уставленная настольными играми, и точная копия университетской библиотеки — вплоть до библиотечной лестницы и светящихся настольных ламп, воссоздававших интерьер знаменитой Bender Room из Библиотеки Грина в Стэнфорде. На одной из полок лежала стопка виниловых пластинок, венчал которую саундтрек к «Бегущему по лезвию».

Но истинной гордостью Альтмана была самая замысловатая архитектурная деталь: витая центральная лестница, которую он спроектировал так, чтобы все 400 сотрудников неизбежно пересекались друг с другом каждый день. Стоя на вершине лестницы и с жаром рассказывая об инженерных решениях, которых она потребовала, Альтман действительно напоминал жреца в собственном храме. Мне вспомнились его слова из нашего разговора.

«Совсем недавно в AGI почти никто не верил, — говорил он. — И сейчас, возможно, многие все еще не верят. Но думаю, теперь больше людей готовы допустить его возможность. Мне кажется, значительная часть мира проходит через тот же процесс, через который за предыдущие годы прошли почти все наши сотрудники, — процесс осознания всего этого. И это непросто. Это захватывает дух. Это пугает. Это просто колоссально. Поэтому я предполагаю, что в ближайшие годы этот процесс охватит весь мир, а мы постараемся стать голосом, который поможет в нем сориентироваться».

Альтман не продавал технологию, — он продавал веру. И уже преуспел в этом сверх всяких ожиданий. Готовя статью для *The Wall Street Journal* — ради которой и приехала в Сан-Франциско, я поинтересовалась у Тиля его мнением об альтмановском идеализме. Тиль ответил: «К нему стоит относиться скорее как к фигуре мессии».

Чуть больше года спустя, в апреле 2024-го, я зашла в лобби отеля Baccarat в центре Манхэттена, и Альтман был уже там, сидел в одном из роскошных кожаных кресел. Я приехала заранее, но он умудрился меня опередить, отправил свою охрану в сторонку и — как выяснилось потом — договорился оплатить счет. Увидев, что я вошла, Альтман вскочил, широко раскинул руки и поприветствовал меня теплым объятием. В этом весь он: сердечный, обаятельный, внимательный, добрый.

Одет Альтман был в свою неизменную униформу: васильковый лонгслив, темные хипстерские джинсы цвета индиго и безукоризненные серые кроссовки New Balance. Через несколько дней ему исполнялось 39, и в каштановых волосах уже кое-где мелькала седина. Он заказал эспрессо — одну из двух чашек своей дневной нормы — и выглядел искренне довольным, воодушевленным продолжительной поездкой в Нью-Йорк.

Меня это удивило, поскольку за месяцы переговоров Альтман недвусмысленно дал понять: книгу писать не стоит. Статья в *The Wall Street Journal*, которую я и моя коллега Бербер Джин подготовили годом ранее после того интервью в офисе OpenAI, привела меня к решению написать книгу. Узнав об этом, Альтман заявил, что еще слишком рано и книга будет чересчур сосредоточена на его персоне. Несколько месяцев спустя он сообщил мне окончательное решение: участвовать в этом проекте не будет. Поскольку герой моей предыдущей книги все время работы над его биографией находился в состоянии близком к вегетативному, меня это особо не смутило, и я продолжала названивать. А потом, за несколько месяцев до нашей встречи в Нью-Йорке, Альтман передумал. Он согласился помочь — немного, но попросил четко обозначить, насколько отвратительным считает весь замысел.

«Я категорически против искажения истории, когда одного человека выставляют олицетворением компании, движения или технологической революции, потому что мир устроен не так и это несправедливо по отношению к выдающейся работе других, — сказал он с нотками гнева в голосе. — Думаю, не следует поощрять подобный взгляд на вещи».

При всем благородстве этой позиции отстаивать ее стало куда труднее после того, как Альтман вернулся в кресло генерального директора всего через пять дней после увольнения. Почти