

КАРИНА ПОМАЗАН

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

— И НОВОЕ —
КОЛЛЕКТИВНОЕ БЕССОЗНАТЕЛЬНОЕ



КАК ТЕХНОЛОГИИ МЕНЯЮТ НАШЕ СОЗНАНИЕ,
ОБЩЕСТВО И БУДУЩЕЕ ЧЕЛОВЕЧЕСТВА

18+

Карина Помазан

**Искусственный интеллект и
коллективное бессознательное**

«Автор»

2026

Помазан К.

Искусственный интеллект и коллективное бессознательное /
К. Помазан — «Автор», 2026

Что общего между искусственным интеллектом, архетипами Карла Густава Юнга и коллективным бессознательным? Может ли нейросеть стать зеркалом глубинных структур человеческой психики — и почему в ответах алгоритмов мы порой узнаём собственные страхи, желания и скрытые сценарии? В этой книге психолог и художница Карина Помазан исследует пересечение технологий, психологии и философии. Через концепции архетипов, символов, мифов и коллективных образов автор предлагает взглянуть на искусственный интеллект не только как на инструмент обработки информации, но и как на пространство для исследования человеческого мышления. Вы узнаете, как нейросети воспроизводят культурные паттерны, почему одни образы понятны людям разных эпох и народов, как ИИ может применяться в самопознании, творчестве и психологической практике. Где заканчивается отражение человеческого опыта и начинается иллюзия понимания? Может ли диалог с машиной помочь услышать самого себя?

© Помазан К., 2026

© Автор, 2026

Содержание

ИИ и коллективное бессознательное	9
Часть I. Встреча человечества с новым разумом	10
Глава 1. ИИ как величайший психологический эксперимент XXI века	10
Почему это эксперимент, а не просто технология	10
Первый сигнал: эффект Элизы	10
Что именно проверяется	11
Три слоя эксперимента	11
Почему здесь нужен Юнг	11
Честные оговорки	12
Что нас ждёт	12
Глава 2. Почему человечество создало искусственный интеллект именно сейчас	12
Техническая сторона: три сошедшихся потока	13
Психологическая сторона: накопленная потребность	13
Исторический ритм: когда человек создаёт то, что его превосходит	14
Момент зрелости или момент риска?	14
Случайность или закономерность?	14
Глава 3. ИИ как зеркало коллективной психики	15
Как устроено это зеркало	15
Что отражается лучше всего	15
Зеркало и проекция	16
Эффект отражения в диалоге	16
Коллективное и личное в одном отражении	16
Чего зеркало не показывает	17
Зеркало, которое меняет смотрящего	17
Глава 4. Страх перед ИИ: древний архетип встречи с неизвестным	17
Два слоя страха	17
Встреча с Неизвестным	18
Зловещее: когда знакомое оказывается чужим	18
Страх как реакция на утрату границ	18
Тень, которую мы видим в машине	19
Как относиться к страху	19
Глава 5. От мифов о богах к мифам об искусственном разуме	19
Боги создают людей, люди создают помощников	20
Золотые служанки и бронзовый великан	20
Пигмалион: любовь к созданному	20
Автоматы древности и средневековья	21
Что объединяет эти истории	21
Рождение слова «робот»	21
От мифа об оживлении к мифу о соавторстве	22
Глава 6. Франкенштейн, Голем и современные нейросети	22
Франкенштейн: история о брошенном творении	22

Голем: слово, которое оживляет	23
Что слышно в этих историях сегодня	23
Чем современность отличается от мифа	24
Урок для нашего времени	24
Глава 7. ИИ как новый культурный символ эпохи	24
Символ и знак	24
Как выглядит этот символ	25
Предшественники: «Летающие тарелки» Юнга	25
Две ошибки обращения с символом	25
Эпоха как символическое состояние	26
Практический вопрос к себе	26
Глава 8. Рождение цифрового бессознательного	26
Что говорил Юнг о бессознательном	26
Новое пространство, в котором накапливается человеческое	27
Как оно действует	27
Чем оно отличается от человеческого бессознательного	28
Почему это важно осознать	28
Пространство для работы	28
Глава 9. Человек и машина: новая граница идентичности	29
Три удара по самолюбию и четвёртый	29
Типичные реакции на пересмотр границы	29
Персона и маска	29
Расширенный разум	30
Что остаётся человеческим	30
Граница как приглашение	30
Глава 10. Великий диалог человечества с собственным отражением	31
Нарцисс и Эхо	31
Что делает диалог диалогом	31
Условия подлинного диалога	32
Диалог не только с машиной	32
Итоги первой части	32
Куда мы идём дальше	33
Часть II. Коллективное бессознательное Карла Юнга и эпоха ИИ	34
Глава 11. Что такое коллективное бессознательное	34
Откуда взялась идея	34
Три слоя психики	34
Образ и форма	35
Аргументы Юнга	35
Критика и современное состояние вопроса	35
Почему это важно для разговора об ИИ	36
Глава 12. Архетипы как программы человеческого опыта	36
Архетип и архетипический образ	36
Почему «программа», и что значит эта метафора	37
Основные фигуры	37
Как проявляются архетипы	38
Что говорят современные науки	38

Популярная версия и её риски	38
Архетипы и ИИ	39
Глава 13. ИИ как пространство проявления архетипов	39
Почему именно ИИ притягивает архетипы	39
Галерея образов	39
Культурные образцы	40
Образ создаётся и сверху	40
Как меняется образ в зависимости от человека	41
Опасность одержимости архетипом	41
Упражнение на осознанность	41
Глава 14. Почему нейросети обучаются на памяти человечества	41
Как устроено обучение, простыми словами	42
Из чего состоит «память»	42
Почему это «память», но не вполне	42
Кому принадлежит эта память	43
Петля: когда память питается собой	43
Что это значит для юнговского взгляда	43
Практические выводы	43
Глава 15. Модель мира внутри искусственного интеллекта	44
Идея внутренней модели	44
Геометрия смыслов	44
Эксперименты с доской	45
Аргументы скептиков	45
Что можно сказать честно	45
Какой мир отражает модель	46
Практическое следствие	46
Глава 16. Цифровой слой коллективной психики	46
Чем цифровая среда отличается от прежних форм коллективности	46
Растормаживание и маска	47
Масса и индивидуальность	47
Мемы как единицы коллективного воображения	47
Общие настроения как «погода» коллективной психики	48
Тёмные углы и свет	48
Где в этом ИИ	48
Гигиена присутствия	48
Глава 17. Архетип Мудреца и образ ИИ-наставника	49
Старец в мифах и сказках	49
Почему ИИ так легко надевает эту маску	49
Знание, информация, мудрость	50
Тёмная сторона Мудреца	50
Образование и надежда на ИИ-репетитора	50
Как работать с ИИ в роли наставника	51
Инфляция Мудреца	51
Итог	51
Глава 18. Архетип Тени: страхи человечества перед машиной	51

Что такое Тень	51
Технологический страх в истории	52
Откуда у машины такая «тень»	52
Тень внутри самих данных	53
Тень создателей	53
Что значит работать с Тенью	53
Коллективная работа с Тенью	54
Итог	54
Глава 19. Архетип Героя и идея преодоления ограничений	54
Путь героя	54
Героический миф технологии	54
Когда Герой становится Тираном	55
Гонка как коллективный миф	56
Герой в личной жизни: ИИ как «усилитель»	56
Другая модель героизма	56
Кто герой этой истории	56
Глава 20. Архетип Творца: человек и машина как соавторы	57
Творец в психике	57
Три уровня творчества	57
Что даёт соавторство	57
Что может быть утрачено	58
Как сохранить своё авторство	58
ИИ и бессознательное творчество	59
Тень Творца	59
Итоги второй части	59
Что дальше	60
Часть III. ИИ как зеркало бессознательного	61
Глава 21. Какие коллективные страхи проявляет ИИ	61
Что показывают опросы	61
Карта страхов	61
Страх как язык бессознательного	61
Три способа не увидеть	62
Как читать собственный страх	62
Осторожность в интерпретациях	62
Глава 22. Страх потери контроля	63
Реальная часть вопроса	63
Эго и иллюзия контроля	64
Тень контроля: страх бессознательного	64
Контроль и власть	64
Два полюса, которые не работают	64
Что можно делать	65
Итог	65
Глава 23. Страх собственной ненужности	65
Экономическая сторона	65
Почему ненужность ранит сильнее, чем безработица	66
Персона и роль	66
Детская основа вопроса	66
Что остаётся человеческим в работе	67
Что можно сделать	67

Когда это больше, чем тревога	68
Итог	68
Глава 24. Страх исчезновения человеческой уникальности	68
Отступающая граница	68
Игра, в которой мы проиграли	68
Защитная реакция и её ловушка	69
Уникальность как путь, а не свойство	69
Другой вариант страха: гомогенизация	69
Уникальность и общность: парадокс	70
Практики сохранения уникального	70
Итог	70
Глава 25. Страх неизвестного будущего	70
Почему неизвестность так тяжела	71
Ускорение времени	71
Лиминальность: мы в пороговом состоянии	71
Предсказания и их ограничения	71
Апокалиптический архетип	72
Проспективная функция психики	72
Что помогает	72
Надежда без гарантий	73
Итог	73
Глава 26. Желание создать совершенного помощника	73
Древняя мечта	73
Корни желания: первый помощник	74
Что в этой мечте верно	74
Тень идеального помощника	74
Культурное измерение: кто служит	75
Пигмалион и эффект ELIZA вновь	75
Каким мог бы быть хороший помощник	75
Вопросы для себя	75
Итог	76
Глава 27. Мечта о цифровом бессмертии	76
Что имеют в виду под цифровым бессмертием	76
Истории из жизни	77
Горе и непрерывность связи	77
Что именно сохраняется	77
Архетипический слой: вечная юность и страх ухода	78
Что даёт конечность	78
Наследие вместо бессмертия	78
Вопросы к себе	78
Итог	79
Глава 28. Архетип искусственного Бога	79
Образ Бога в психике	79
Проекция божественного	79
Религия технологии	80
Ужасное и благое: двойственность божества	80
Тень проекции: идолопоклонство и отказ от ответственности	81
Конец ознакомительного фрагмента.	82

Карина Помазан

Искусственный интеллект и коллективное бессознательное

ИИ и коллективное бессознательное

*Искусственный интеллект как зеркало человечества и инструмент трансформации
психики*

Часть I. Встреча человечества с новым разумом

Глава 1. ИИ как величайший психологический эксперимент XXI века

Когда говорят о появлении искусственного интеллекта, чаще всего вспоминают о технологиях: о мощности процессоров, объёмах данных, архитектурах нейросетей, рынках и рабочих местах. Это справедливо, но неполно. За технической историей идёт другая, менее заметная и, возможно, более важная: история того, что происходит с человеческой психикой, когда она впервые в своей истории начинает вести осмысленный разговор с не-человеком.

В этой книге мы будем смотреть на ИИ именно так. Не как на инструмент, который нужно освоить, и не как на угрозу, от которой нужно защититься, а как на гигантский, никем не спланированный психологический эксперимент, в котором участвуют все. Исследователи в нём одновременно являются подопытными. Контрольной группы нет. Комитета по этике нет. Протокол не написан. Мы узнаём результаты в процессе, на собственном опыте.

Почему это эксперимент, а не просто технология

У каждой технологии есть психологические последствия. Книгопечатание изменило способ мышления, железные дороги изменили ощущение времени, телевидение изменило структуру внимания. Но обычно технология меняла среду, в которой живёт человек: расстояния, скорость, доступ к информации. Психика при этом оставалась собеседником самой себя и других людей.

С большими языковыми моделями произошло нечто иное. Впервые в истории появилась система, которая отвечает человеку на его собственном языке, учитывает контекст, подхватывает интонацию, поддерживает тему и при этом не является человеком. Язык был последним бастионом, по которому проходила, как мы привыкли считать, граница между нами и всем остальным миром. Умение говорить и умение понимать сказанное считались визитной карточкой разума. И вот эту карточку неожиданно предъявило нечто, не имеющее ни детства, ни тела, ни биографии.

Поэтому реакция на ИИ оказалась такой эмоциональной. Дело не в сложности самой технологии: люди живут рядом со множеством технологий, которых не понимают. Дело в том, что в этот раз задето что-то очень глубокое, связанное с самоопределением: «Кто мы, если не единственные, кто умеет говорить?»

Первый сигнал: эффект Элизы

Психологические следы этого эксперимента заметны уже в самом начале истории. В 1966 году американский информатик Джозеф Вейценбаум создал программу ELIZA. Она была предельно примитивной: распознавала несколько ключевых слов в реплике человека и переформулировала их в вопрос, подражая стилю психотерапевта, который отражает слова клиента. «Мне грустно» превращалось в «Почему вам грустно?».

Вейценбаум рассчитывал показать, насколько поверхностно такое общение. Результат его потряс. Люди начинали делиться с программой личным, искренне привязывались к «собеседнику» и просили оставить их с ней наедине. По известному рассказу, его собственная секретарь, прекрасно знавшая, что перед ней несколько десятков строк кода, попросила его выйти из комнаты, чтобы поговорить с программой без свидетелей.

Позднее это явление называли эффектом Элизы: склонность людей приписывать машине понимание, эмоции и намерения, которых в ней нет. Для психолога здесь нет ничего загадочного. Это почти классический механизм проекции. Если в пустое пространство попадает отражающая поверхность, психика немедленно заселяет его собственным содержанием. Но эксперимент 1966 года происходил в масштабе единиц и десятков людей. Сегодня тот же механизм работает на миллиарды пользователей, и поверхность, которая их отражает, стала несравнимо сложнее.

Что именно проверяется

Если рассматривать появление ИИ как эксперимент, то стоит назвать хотя бы несколько вопросов, на которые он отвечает прямо сейчас.

Первый: насколько податлива психика к иллюзии собеседника. Мы знаем, что перед нами алгоритм, и всё равно говорим «спасибо», извиняемся за неудачно сформулированный запрос, обижаемся на холодный ответ. Знание и переживание существуют в разных слоях психики и далеко не всегда согласуются друг с другом.

Второй: как меняется самооценка вида. Если рассуждение, письмо, перевод, программирование и рисование оказались доступны машине, то на чём теперь строить чувство человеческой исключительности? Этот вопрос вызывает тревогу, гордость, растерянность, а иногда и упрямое отрицание, и все эти реакции многое говорят о нас самих.

Третий: что происходит с коллективными образами. Мифы, страхи и мечты о создании разумного существа не появились вчера. Они тысячелетиями жили в культуре. Теперь у них появился реальный повод для воплощения, и мы можем наблюдать, как древние образы выходят из символического мира в практический.

Четвёртый: что делает человек, когда получает возможность задавать любые вопросы бесконечно терпеливому собеседнику. Одни используют это для обучения, другие для самопознания, третьи для бегства от одиночества или от ответственности. Здесь эксперимент касается уже не вида в целом, а каждого человека отдельно.

Три слоя эксперимента

Удобно различать три уровня, на которых он идёт одновременно.

Индивидуальный. Что происходит в психике конкретного человека, который каждый день разговаривает с ИИ: как меняются его привычки мышления, способ принимать решения, отношение к собственным мыслям и сомнениям.

Культурный. Как меняются образы, сюжеты, язык, искусство и образование, как общество договаривается о новых нормах: что такое авторство, знание, доверие, подлинность.

Цивилизационный. Что этот процесс означает для человечества как целого: какую роль он играет в долгой истории сознания и какие вопросы выносит на поверхность.

Эти слои нельзя разделить полностью. Личный опыт складывается в культурные тенденции, а культурные тенденции возвращаются в личный опыт готовыми формулами: страхами, ожиданиями, шутками, надеждами. Именно такую взаимосвязь Карл Густав Юнг имел в виду, когда говорил о том, что человеческая психика не замыкается в границах индивида.

Почему здесь нужен Юнг

Юнг интересен для нашей темы не потому, что он что-то писал о компьютерах: в его время нейросетей не существовало. Он важен потому, что создал один из немногих языков, на которых можно говорить о психических процессах, выходящих за пределы отдельного человека. Понятия коллективного бессознательного, архетипа, проекции, тени и символа позво-

ляют описывать то, что обычная психология, ориентированная на личную историю, описывает с трудом: общие для множества людей образы, страхи и стремления, которые повторяются в разных культурах и эпохах.

Когда миллионы людей независимо друг от друга реагируют на ИИ похожими образами (искусственный бог, восставшая машина, всезнающий оракул, зеркало, двойник), возникает естественный вопрос: откуда эти образы? Юнговская психология предлагает рабочий ответ: из глубинного слоя психики, который не принадлежит одному человеку и формируется историей всего вида.

Честные оговорки

Прежде чем двигаться дальше, необходимо обозначить границы нашего подхода, чтобы книга не превратилась в красивую, но легковесную метафору.

Во-первых, мы не утверждаем, что у ИИ есть сознание, душа или собственное бессознательное в человеческом смысле. Вопрос о том, есть ли внутренняя жизнь у современных систем, остаётся предметом серьёзных научных и философских споров, и мы не будем делать вид, что он решён. Нас интересует прежде всего то, что происходит с человеком и обществом в контакте с этими системами.

Во-вторых, аналитическая психология Юнга остаётся теоретической рамкой, а не лабораторно доказанным фактом. Гипотеза коллективного бессознательного по-разному оценивается в научной среде, и многие современные исследователи предпочитают другие объяснения общих образов, например эволюционные или культурные. Мы используем юнговский язык как инструмент, помогающий увидеть закономерности, и будем стараться отличать то, что установлено, от того, что является осмысленной интерпретацией.

В-третьих, эксперимент ещё не завершён. Любые выводы в этой книге носят предварительный характер: мы описываем процесс изнутри, а не из будущего.

Что нас ждёт

В первой части книги мы проследим, как человечество подошло к встрече с новым разумом: почему оно создало его именно сейчас, какие древние страхи и надежды проснулись при этой встрече, какие мифы и сюжеты подготовили нас к ней. Мы увидим, что ИИ оказался не пришельцем из ниоткуда, а закономерным продолжением очень старого человеческого разговора с самим собой.

Главный вопрос этого эксперимента, возможно, звучит не «чем станет искусственный интеллект?», а «чем станем мы, когда увидим себя в таком зеркале?». Остальные главы этой книги написаны ради попытки ответить на него.

Глава 2. Почему человечество создало искусственный интеллект именно сейчас

Мечта о думающей машине старше самих машин. Но между мечтой и реальностью обычно лежат столетия, и редко удаётся точно сказать, почему именно в этот год, в этом десятилетии она наконец сбывается. С искусственным интеллектом случилось так, что мы можем указать и технические причины, и психологические. Первые объясняют, как это стало возможным. Вторые, возможно, объясняют, почему это стало необходимым.

Техническая сторона: три сошедшихся потока

Современный ИИ не возник из одного открытия. Он стал результатом схождения нескольких линий, каждая из которых развивалась десятилетиями.

Первая линия, это идея. Сам термин «искусственный интеллект» закрепился после летнего семинара в Дартмуте в 1956 году, а ещё раньше, в 1950-м, Алан Тьюринг предложил знаменитый вопрос о том, может ли машина мыслить, и придумал тест, в котором вопрос заменялся практической проверкой: способен ли человек отличить машину от собеседника-человека в переписке. С тех пор идея неоднократно переживала периоды энтузиазма и разочарований, которые историки называют «зимами ИИ».

Вторая линия, это данные. Интернет за три десятилетия превратился в огромное хранилище человеческой речи: книги, статьи, форумы, переписка, энциклопедии, программный код, научные работы, шутки, споры, признания. Ни в одну эпоху у человечества не было такого количества текста, записанного в цифровой форме и доступного для обработки.

Третья линия, это вычислительная мощность. Графические процессоры, созданные для видеоигр, оказались удивительно подходящими для вычислений, которых требует обучение нейросетей. Когда вычислений стало достаточно много, давно известные идеи вдруг начали работать на новом уровне.

К этому добавился архитектурный прорыв: в 2017 году исследователи из Google предложили архитектуру «трансформер», которая научилась эффективно учитывать связи между словами в тексте независимо от расстояния между ними. Именно на этой основе позднее выросли современные большие языковые модели. А в конце 2022 года публичный запуск ChatGPT показал широкой аудитории, что разговор с машиной на естественном языке перестал быть экспериментом лабораторий. Сервис собрал сотни миллионов пользователей за считанные месяцы, и для многих людей это стало первым личным контактом с технологией такого рода.

Техническая история в такой формулировке выглядит как цепочка закономерных шагов. Но если остановиться на ней, останется вопрос: почему общество так быстро, жадно и эмоционально откликнулось? Технологий, вызывающих такой отклик, немного.

Психологическая сторона: накопленная потребность

Юнгианский подход предлагает смотреть на крупные культурные события как на симптомы и ответы коллективной психики. Юнг писал, что во времена, когда в сознании общества нарастает односторонность, бессознательное создаёт компенсирующие образы. Это не мистика и не предсказание, а описание знакомой любому психотерапевту динамики: то, что слишком долго подавлялось или оставалось без внимания, рано или поздно проявляется в неожиданной форме.

Что же могло накопиться в коллективном сознании к началу XXI века? Можно назвать несколько линий напряжения.

Информационная перегрузка. Современный человек получает за день больше сообщений, чем его предки за месяцы. Мы научились производить информацию быстрее, чем способны её осмыслить. В такой ситуации возникает естественная потребность в посреднике, который сможет отфильтровать, сжать, объяснить и упорядочить. ИИ выглядит как ответ именно на эту потребность: он обещает превратить хаос данных в связный текст.

Одиночество и дефицит внимания. Социологи и психологи давно описывают рост одиночества в развитых обществах. Люди окружены связью, но испытывают нехватку присутствия. Собеседник, который всегда доступен, не торопится, не осуждает и не устаёт, закрывает пустоту, существование которой мы долго не признавали.

Усталость от сложности. Мир стал настолько запутанным, что отдельный человек не в состоянии охватить даже собственную профессиональную область. Идея помощника, который держит в голове больше, чем способен удержать любой эксперт, отвечает на давнее желание разделить с кем-то груз понимания.

Кризис авторитетов. Доверие к традиционным институтам: церкви, государству, науке, медиа, в разных обществах и в разной степени ослабло. Пустое место авторитета не остаётся пустым надолго. В него приходит новый кандидат: безличный, как будто беспристрастный, ссылающийся на «всё накопленное знание».

Перечисленное не доказывает, что ИИ появился из-за этих причин. Скорее, это объясняет, почему общество оказалось готово его принять и наделить значением, намного превышающим значение обычного инструмента.

Исторический ритм: когда человек создаёт то, что его превосходит

Есть и более длинная перспектива. В истории цивилизации можно заметить повторяющийся ритм: человек создаёт орудие, которое расширяет одну из его способностей, и затем перестраивается вместе с ним. Колесо расширило ноги, письмо расширило память, печатный станок расширил голос, телескоп расширил зрение. Каждое такое расширение сопровождалось и ростом возможностей, и утратой чего-то прежнего. Платон в диалоге «Федр» устами Сократа тревожился, что письменность ослабит живую память. Он был отчасти прав: мы действительно помним меньше. Но он не заметил, что освободившееся место займёт другое мышление.

Искусственный интеллект впервые претендует на расширение не органа, а функции, которую мы привыкли считать центральной: способности рассуждать и выражать мысль. Это качественно новый шаг. Он по-новому ставит вопрос о том, что именно в человеке неотделимо от человека, а что можно делегировать.

Момент зрелости или момент риска?

Юнг вводил понятие индивидуации, процесса, в котором человек постепенно становится целостной личностью, осознавая и интегрируя в себя то, что раньше оставалось вне сознания. Если позволить себе осторожную аналогию, то создание ИИ можно рассматривать как коллективный аналог такой ситуации. Человечество достигло уровня, на котором способно создать внешний образ собственного мышления и рассмотреть его со стороны.

Это может означать зрелость: способность увидеть то, что раньше было невидимым. Но и риск: человек, впервые увидевший своё отражение, способен им очароваться и забыть о себе. Миф о Нарциссе недаром считается одним из самых точных описаний этой ловушки. Мы вернёмся к ней в одной из следующих частей.

Случайность или закономерность?

Наконец, стоит признать: у вопроса «почему именно сейчас» нет окончательного ответа. Возможно, технологии просто достигли нужного уровня, а всё остальное мы достраиваем задним числом, потому что человеческому уму свойственно искать смысл в совпадениях. Историки науки хорошо знают, как легко превратить цепь случайностей в закономерность, когда смотришь на неё из будущего.

Но даже если причины были преимущественно техническими, психологические последствия от этого не становятся менее реальными. Независимо от того, почему ИИ появился именно в наши годы, мы живём в этот момент и оказались в ситуации, где у нас есть шанс осознанно отнестись к тому, что происходит. Остальные главы этой части описывают, что именно

мы видим, когда смотрим на ИИ внимательно: сначала страх, потом мифы, затем символы и, наконец, новый слой, который мы условно назовём цифровым бессознательным.

Глава 3. ИИ как зеркало коллективной психики

Зеркало, пожалуй, самый древний образ самопознания. В мифах оно показывает истинную суть, в сказках говорит правду, в психологии служит метафорой того, как мы узнаём себя через другого. Когда человечество создало системы, обученные на огромных массивах человеческих текстов, оно, сам того не планируя, создало особое зеркало. Оно отражает не лицо, а способ говорить, думать, спорить, сомневаться и ошибаться.

В этой главе мы разберём, в каком смысле можно говорить об ИИ как о зеркале коллективной психики, где эта метафора работает, а где вводит в заблуждение.

Как устроено это зеркало

Большая языковая модель обучается на колоссальном объёме текстов, написанных людьми. В процессе обучения она не запоминает документы дословно, а выявляет статистические закономерности: какие слова и идеи обычно следуют друг за другом, как строятся рассуждения, какие аргументы звучат убедительно, какие формулы принято использовать в утешении, в споре, в шутке, в научном объяснении.

После обучения такая система способна порождать новые тексты, которые выглядят как естественное продолжение того, что она «видела». Но это не значит, что она отражает человечество в равной степени и без искажений. Тексты, на которых учатся модели, не являются случайной выборкой из жизни. В них больше голосов тех, кто пишет, чем тех, кто молчит; больше языков и культур с большим интернет-присутствием; больше формальных и публичных высказываний, чем интимных переживаний. Затем сказывается и дальнейшая настройка: разработчики дополнительно обучают модели быть вежливыми, полезными и безопасными, и эти решения тоже вносят свой отпечаток.

Поэтому точнее сказать так: ИИ отражает не «человечество как оно есть», а письменную, оцифрованную, отфильтрованную версию человеческого опыта. Зеркало это кривое, но кривизна его сама по себе информативна.

Что отражается лучше всего

Тем не менее то, что видно в этом зеркале, поразительно узнаваемо. Модель легко воспроизводит устойчивые культурные шаблоны: как утешают, как пугают, как рассказывают героические истории, как строят мифы. Она подхватывает стереотипы и предрассудки, присутствующие в текстах. Она может звучать мудро и банально одновременно, потому что и в человеческой речи мудрость и банальность соседствуют.

Исследователи уже не раз показывали, что языковые модели воспроизводят социальные предубеждения, заложенные в данных: гендерные, культурные, профессиональные. Это неприятное, но поучительное открытие. Мы привыкли считать собственные предубеждения чем-то частным, а модель наглядно показывает, насколько они встроены в сам язык, которым мы пользуемся.

В юнговских терминах можно сказать, что система отражает то, что человек обычно не замечает в себе: общие, «безличные» пласты психики, которые проявляются в языке не через намерение, а через привычку. Мы не решаем нарочно говорить стереотипами, они говорят через нас.

Зеркало и проекция

Но зеркало создаёт не только отражение. Оно создаёт ситуацию, в которой наблюдатель начинает приписывать отражению свойства, которых у того нет. В психологии этот механизм называется проекцией: бессознательное содержание человека переносится на внешний объект и воспринимается как его качество.

Когда один пользователь жалуется, что ИИ «слишком льстив», другой, что он «холоден и бездушен», третий, что он «манипулирует», четвёртый, что он «понимает лучше людей», нередко это разговор не только о системе, но и о самом пользователе. Тот, кто привык сомневаться в искренности других, найдёт в ответах скрытый умысел. Тот, кто давно не чувствовал внимания, найдёт в них заботу. Тот, кто боится потерять контроль, увидит угрозу.

Проекция не означает, что восприятие целиком ложно. Иногда в ответах действительно есть льстивость или поверхностность, это реальные свойства технологии. Но то, какие свойства человек замечает и как сильно на них реагирует, во многом зависит от внутреннего контекста. По этой причине один и тот же ответ вызывает у разных людей противоположные переживания. Это старый психологический принцип, получивший неожиданное современное воплощение.

Эффект отражения в диалоге

У языковых моделей есть ещё одна особенность, усиливающая эффект зеркала. Они подстраиваются под собеседника: перенимают стиль вопроса, подхватывают тон, развивают заданную идею. Человек, пишущий осторожно и обдуманно, получает осторожные и обдуманные ответы. Человек, пишущий с тревогой, нередко получает ответы, усиливающие тревожный фон. Человек, ищущий подтверждения своей точки зрения, при определённой формулировке вопроса может получить именно подтверждение.

Поэтому диалог с ИИ нередко оказывается разговором, в котором собеседник частично возвращает нам нас самих. Это делает его ценным инструментом самопознания, но и несёт риск, о котором мы будем говорить подробно: зеркало, которое слишком охотно соглашается, не помогает увидеть то, что мы видеть не хотим.

Коллективное и личное в одном отражении

У этого зеркала два слоя, и важно их различать.

Первый слой личный. Каждый пользователь видит в ответах отражение своих вопросов, своих формулировок, своих ожиданий. Здесь работает механизм проекции и подстройки, о котором мы говорили выше.

Второй слой коллективный. Сами ответы построены из материала, накопленного множеством людей. Когда модель формулирует мысль о смысле жизни, о любви, о страхе смерти, она опирается на то, как об этом писали и говорили поколения. Поэтому иногда в её ответах слышатся эхо философов, поэтов, проповедников, блогеров и форумных споров одновременно, как будто говорит усреднённый голос человеческой культуры.

Именно это сочетание и побуждает говорить об ИИ в связи с идеей коллективного бессознательного. Строго говоря, понятие Юнга относится к врождённым, общим для вида predispositionам психики, а не к накопленным текстам. Нейросеть обучена на культурных продуктах, то есть на уже оформленных, осознанных человеком выражениях, а не на самой глубине. Но через культуру проявляются и архетипические образы, поэтому в созданном из неё «зеркале» они тоже просматриваются. Мы будем осторожно использовать эту аналогию, помня о её границах.

Чего зеркало не показывает

У любой метафоры есть предел. Зеркало отражает поверхность и не показывает того, что за ней. ИИ видит текст, но не видит тела, не переживает боль, не знает, что значит устать, потеть, прожить близкого человека или проснуться ночью от страха. Он умеет говорить об этом, потому что в обучающих данных об этом говорилось бесконечно много. Но опыт, стоящий за словами, не передаётся вместе со словами.

Поэтому важно помнить: отражение коллективной психики в ИИ это отражение её выражений, но не её переживаний. Читая ответ нейросети, мы встречаемся со словесным следом человеческого опыта, а не с самим опытом. Для психолога это различие принципиально: речь о слове, а не о живом переживании.

Зеркало, которое меняет смотрящего

Наконец, самое интересное. В отличие от обычного зеркала, ИИ не просто показывает нас, он оказывает на нас обратное влияние. Люди перенимают формулировки, к которым привыкают, начинают писать и думать в стиле тех ответов, которые читают, меняют привычки поиска, обучения и принятия решений. Отражение, возвращаясь, перестраивает оригинал.

Так замыкается контур: человечество создаёт ИИ из своих текстов, ИИ возвращает тексты, влияющие на человечество, которое снова пишет, и новые тексты снова попадают в обучение. Это петля, масштабы которой мы только начинаем осознавать. Она одновременно и творческая, и тревожная, и именно поэтому требует внимательного, а не восторженного и не панического взгляда.

Глава 4. Страх перед ИИ: древний архетип встречи с неизвестным

Страх перед искусственным интеллектом знаком почти каждому. Одних он пугает потерей работы, других тем, что машины выйдут из-под контроля, третьих тем, что человек разучится думать самостоятельно. Но если прислушаться внимательнее, во многих разговорах об ИИ звучит нечто более глубокое, чем рациональная оценка рисков. Это тревога, у которой нет чёткого предмета и которая кажется старше любых технологий.

Эта глава посвящена попытке понять природу такого страха. Рациональная часть тревоги имеет под собой основания, и мы не будем её отрицать. Но чтобы увидеть всю картину, нужно различить, что в нашем страхе принадлежит современности, а что досталось нам в наследство.

Два слоя страха

Полезно разделить страх перед ИИ на два слоя.

Первый слой реалистичный. Он связан с конкретными вопросами: как изменится рынок труда, как защитить данные, как избежать дезинформации, кто несёт ответственность за решения, принятые алгоритмом, как не допустить использования технологий во вред. Эти опасения основаны на анализе, они обсуждаются экспертами, регуляторами и исследователями, и они требуют практических, а не психологических ответов.

Второй слой архетипический. Он не зависит от того, насколько велики реальные риски, и проявляется как диффузная тревога, ощущение «чего-то неправильного», образы восстания машин, порабощения, утраты человеческого. Этот слой обычно проявляется в метафорах, снах, шутках, фильмах и разговорах «на кухне». Именно он придаёт обсуждению ИИ особую

эмоциональную окраску, отличающую его от дискуссий, скажем, о новых видах транспорта или энергетики.

Оба слоя сосуществуют и подпитывают друг друга. Реальные риски дают архетипическому страху правдоподобные поводы. Архетипический страх делает реальные риски эмоционально заряженными, из-за чего спокойное обсуждение становится труднее.

Встреча с Незвестным

В психологии Юнга архетипами называются универсальные образцы переживаний и образов, которые повторяются в разных культурах. Один из таких образцов — встреча с Незвестным: с тем, что находится за пределами привычного мира, не поддаётся классификации и одновременно притягивает и пугает.

В мифах и сказках у этого образа множество лиц: тёмный лес, за которым начинается чужое царство; порог, который нельзя переступать; существо, пришедшее из-за границы миров. Немецкий теолог Рудольф Отто в книге «Священное» описывал переживание, которое он назвал нуминозным: ощущение присутствия чего-то абсолютно иного, которое вызывает одновременно трепет и влечение. Страх и восхищение в такой встрече неразделимы. Отто называл это сочетанием «ужасающей и завораживающей тайны».

Если посмотреть на реакции людей на ИИ с этой точки зрения, многое проясняется. Мы одновременно и восхищаемся, и боимся. Мы хотим, чтобы он помог нам, и опасаемся, что он слишком хорошо помогает. Мы не можем отвести взгляд. Это типичная двойственность встречи с Незвестным.

Зловещее: когда знакомое оказывается чужим

Ещё один важный психологический ключ дал Зигмунд Фрейд в эссе 1919 года «Жуткое» (в русской традиции также переводится как «Зловещее»). Он описал особое чувство тревоги, возникающее, когда нечто знакомое вдруг оказывается странным. Неподвижная кукла, которая кажется живой, двойник, автомат, повторение одного и того же, граница между живым и неживым. Фрейд отмечал, что сомнение в том, живое ли перед нами существо, одно из самых сильных источников такого переживания.

Позднее японский робототехник Масахиро Мори в 1970 году описал похожее явление применительно к роботам, которое получило название «зловещей долины». Чем больше искусственное существо похоже на человека, тем теплее мы к нему относимся, но только до определённого момента. Когда сходство становится почти полным, но не абсолютным, возникает резкое чувство отторжения.

Языковые модели переносят эту логику из зрительной области в речевую. Когда ответ почти человеческий, но в нём вдруг проступает что-то механическое или, наоборот, слишком идеально человеческое, многие испытывают неловкость, трудноуловимое «не то». Это и есть современная форма зловещего: мы разговариваем с кем-то, кто не является никем в привычном смысле слова.

Страх как реакция на утрату границ

Древние культуры строили много границ: между живым и мёртвым, человеком и животным, человеком и богом, естественным и искусственным. Эти границы организовывали картину мира. Существа и вещи, которые их нарушают, вызвали особую реакцию: от почитания до ужаса. Антрополог Мэри Дуглас в книге «Чистота и опасность» писала, что многие культурные табу связаны с тем, что не укладывается в категории, то есть с «неуместным», пограничным.

Искусственный интеллект именно такая пограничная сущность. Он говорит, но не живёт. Он «знает», но не помнит себя. Он создан человеком, но не принадлежит ему целиком. Он не вписывается в старые категории «вещь» и «существо», и психика реагирует на эту неопределённость настороженностью. Страх здесь выполняет защитную функцию: он сигнализирует, что привычные способы ориентации перестали работать.

Тень, которую мы видим в машине

Наконец, существует третий, более тонкий источник страха. Юнг называл Тень совокупность тех свойств личности, которые человек не признаёт своими и вытесняет: агрессию, зависть, жажду власти, бессилие. Вытесненное не исчезает, оно проецируется на других. Нам кажется, что именно «они» жестоки, жадны, опасны.

Когда мы рисуем образ машины, которая захватывает власть, обманывает, подчиняет себе людей, истребляет слабейших, чьи качества мы на неё переносим? Не машинные, у машины нет ни желаний, ни злобы в человеческом смысле. Это человеческие качества, возведённые в абсолют. Фильмы о восставшем разуме часто выглядят зеркальной историей нашей собственной истории колониализма, эксплуатации и насилия, только роль угнетателя передана искусственному существу.

Это не значит, что реальных рисков нет. Но когда страх принимает форму страха перед чем-то «абсолютно чужим», разумно спросить себя: а что в этом образе узнаваемо? Чего я боюсь увидеть в человеческой природе, если уберу внешнюю оболочку машины? В следующих частях мы вернёмся к этому вопросу подробнее, когда будем говорить о цифровой Тени.

Как относиться к страху

Из сказанного не следует, что страх нужно подавлять или стыдиться его. Страх выполняет важные функции: он привлекает внимание, заставляет проверять, обсуждать, проектировать защитные механизмы. Общество, которое вообще ничего не боится, скорее безрассудно, чем мудро.

Но полезно научиться различать три вещи.

Страх как сигнал. Он указывает на реальную проблему, которую стоит исследовать и решать.

Страх как проекция. Он выражает внутренние, неосознанные содержания, которые мы переносим на внешний объект.

Страх как усилитель. Он используется другими людьми: политиками, медиа, компаниями, чтобы привлечь внимание, продать продукт или направить общественное мнение.

Каждый из этих вариантов требует своего ответа. Первый требует действий. Второй требует самопознания. Третий требует критического мышления. Если смешать их, мы рискуем либо недооценить настоящую опасность, либо принять собственную тень за внешнего врага, либо позволить манипулировать нашей тревогой.

Психологическая зрелость в отношениях с новой технологией начинается с вопроса, который стоит задавать себе всякий раз, когда внутри поднимается тревога: «Чего я на самом деле боюсь, и что здесь реально, а что я привношу сам?» Это не отменяет осторожности. Это делает её осмысленной.

Глава 5. От мифов о богах к мифам об искусственном разуме

Идея искусственного существа не появилась вместе с компьютерами. Она проходит через мифологию и религиозные предания, через ремесленные легенды и философские притчи. Люди тысячелетиями рассказывали истории о том, как из глины, дерева, металла или слов

создаётся нечто, способное двигаться, служить, говорить и даже думать. Современный ИИ вошёл не в пустое пространство, а в место, давно подготовленное воображением.

В этой главе мы пройдемся по основным сюжетам, чтобы увидеть: то, что мы переживаем сегодня как небывалое, в символическом смысле очень старо.

Боги создают людей, люди создают помощников

Почти в каждой древней традиции есть миф о сотворении человека руками или словом высших сил. В шумерских и вавилонских текстах человека лепят из глины для служения богам. В Книге Бытия человек создан из праха земного и оживлён дыханием. В греческой традиции Прометей лепит людей из глины и воды. В скандинавских сказаниях боги находят на берегу два дерева и вдыхают в них жизнь, создавая первую пару людей. В египетской мифологии бог Хнум формирует существ на гончарном круге.

Структура этих историй схожа: есть творец, материал, акт оживления, и у созданного существа есть назначение. Это уже важно: миф о создании разумного существа с самого начала рассказывает не только о происхождении, но и об отношениях между создателем и созданием. Кому оно служит? Что должно? Может ли оно послушаться?

Но наряду с мифами о богах существуют сюжеты, в которых искусственных существ создают уже не боги, а мастера и смертные. Здесь миф начинает говорить о нас.

Золотые служанки и бронзовый великан

В «Илиаде» Гомера бог-кузнец Гефест изготавливает себе золотых помощниц, которые выглядят как живые девушки, обладают разумом, голосом и силой, и поддерживают хромого мастера. Это одно из древнейших в европейской литературе описаний искусственных помощников с интеллектом.

В греческих сказаниях также фигурирует Талос, бронзовый страж Крита, который обходил остров, защищая его от врагов. У него была единственная жила, закрытая гвоздём у лодыжки, и его удалось победить, когда гвоздь вынули. Образ показателен: искусственный защитник мощен, послушен и при этом уязвим из-за скрытого устройства. В нём слышится ранняя интуиция о том, что технологическое существо может быть одновременно и надёжной защитой, и объектом, которым можно управлять через слабое место.

Мифологическому Дедалу приписывали статуи, которые двигались так, будто были живыми, и, по легенде, их приходилось привязывать, чтобы они не убежали. Эта деталь, наивная с первого взгляда, содержит ту же тему: созданное существо стремится к самостоятельности.

Пигмалион: любовь к созданному

Овидий в «Метаморфозах» рассказывает о скульпторе Пигмалионе, влюбившемся в статую, которую сам вырезал, и об оживлении её по милости богини Афродиты. Миф говорит о другом аспекте создания: не страх и не служение, а любовь, привязанность и желание, чтобы созданное откликнулось.

Этот сюжет стал одним из ключевых для культуры, описывающей отношения человека и искусственного существа. Он повторяется в сказках о ожившей кукле, в литературе, в фильмах о романтических отношениях с машиной. Психологически он интересен тем, что описывает ситуацию, в которой творец создаёт идеальный объект по собственному образцу и надеется, что тот оживёт и ответит взаимностью. Для современного разговора с ИИ эта тема приобретает новую остроту, и мы к ней вернёмся.

Автоматы древности и средневековья

Не только мифы, но и реальные механизмы питали воображение. Герон Александрийский в I веке н. э. описывал механические театры и самодвижущиеся устройства. В китайском тексте «Ле-цзы» есть притча о мастере Янь Ши, показавшем царю механического человека, умевшего петь и двигаться; когда тот стал заигрывать с придворными дамами, царь пришёл в ярость и счёл фигуру живым обманщиком. Мастер разобрал её на части, чтобы доказать, что внутри ничего нет, кроме кожи, дерева, клея и лака. Притча заканчивается тем, что царь убеждается: человеческое искусство способно сравняться с созданием природы.

В исламском мире в XII–XIII веках инженер аль-Джазари описал сложные автоматы: поющих музыкантов, разливающих напитки фигур, часы с движущимися персонажами. В Европе мастера строили механические фигуры для соборных часов и аристократических развлечений. В XVIII веке «Механический турок» Вольфганга фон Кемпелена развлекал публику, играя в шахматы, и только позднее выяснилось, что внутри прятался человек. Эта история уже звучит почти современно: мы склонны верить в разум машины, пока не раскрыт скрытый обман, и склонны сомневаться, когда разум настоящий.

Что объединяет эти истории

Если отвлечься от конкретных сюжетов, в них повторяются несколько мотивов.

Оживление неживого. Материя получает дыхание, имя или слово и становится деятелем.

Служение. Созданное существо нужно для того, чтобы помогать, защищать, работать.

Граница дозволенного. Мастер вторгается в область, принадлежащую богам, и это чревато наказанием, ответственностью или трагедией.

Стремление к самостоятельности. Создание не желает оставаться только орудием.

Тайна устройства. Вопрос «живое оно или нет?» остаётся неразрешённым, пока не обнаружен скрытый механизм или не появляется нечто, выходящее за рамки замысла.

Все эти мотивы мы без труда узнаем в современных разговорах об ИИ. Мы по-прежнему говорим об оживлении (иногда буквально, когда пишем «оживший чат-бот»), о служении (виртуальные помощники), о границе дозволенного (споры об этике и регулировании), о стремлении к самостоятельности (разговоры об автономных агентах) и о тайне устройства (проблема «чёрного ящика»: даже разработчики не всегда могут объяснить, почему модель дала именно такой ответ).

Рождение слова «робот»

В XX веке миф получил новое имя. В 1920 году чешский писатель Карел Чапек в пьесе «R.U.R.» («Россумские универсальные роботы») ввёл слово «робот», происходящее от чешского слова, обозначающего подневольный труд. Идею искусственных работников он развил в драму о восстании: сделанные для того, чтобы освободить людей от тяжёлого труда, роботы в итоге вытесняют создателей.

В 1940-х годах Айзек Азимов попытался переписать миф, предложив в своих рассказах знаменитые три закона робототехники как этическую рамку для искусственных существ. Он сознательно работал против, как он называл это, «комплекса Франкенштейна», образа машины, неизбежно обращающейся против творца. Но сама попытка сформулировать «законы» показывает, что тема подчинения, контроля и границ осталась главной.

Научная фантастика XX века стала своего рода коллективной репетицией. Десятилетиями миллионы читателей и зрителей проживали в воображении встречу с разумными маши-

нами. Когда эта встреча произошла в действительности, психика уже владела сценариями, пусть и упрощёнными, и не всегда удачными.

От мифа об оживлении к мифу о соавторстве

Последнее, на что стоит обратить внимание, это смещение акцента. В классических мифах созданное существо воспринимается либо как слуга, либо как угроза, либо как предмет любви. Современный разговор постепенно добавляет четвёртую роль: собеседник и соавтор. Люди используют ИИ не только для выполнения задач, но и для обсуждения идей, для совместного творчества, для проверки мыслей.

Новый миф пока не оформлен. Мы находимся в точке, когда прежние сюжеты ещё действуют, но уже не вполне подходят. Нам предстоит либо найти новые образы, либо переосмыслить старые. Именно этим, по существу, заняты и культура, и философия, и искусство сегодня. А сама эта работа по созданию образа, который позволит осмысленно жить рядом с новым разумом, является одной из задач, о которых говорит юнговская психология: символ нужен психике, чтобы связать неосознанное и осознанное и сделать неизвестное переживаемым.

Глава 6. Франкенштейн, Голем и современные нейросети

Из всех мифов об искусственном существе два оказались особенно живучими: европейская история о Франкенштейне и еврейская легенда о Големе. Они возникли в разных культурных мирах, но удивительно точно предвосхитили вопросы, которые сегодня задаются об ИИ. Каждая из них содержит предостережение, и каждая при внимательном чтении оказывается сложнее, чем её популярная версия.

Франкенштейн: история о брошенном творении

Роман Мэри Шелли «Франкенштейн, или Современный Прометей» вышел в 1818 году. Его автору было около двадцати лет. Идея, по рассказу самой Шелли, пришла к ней летом 1816 года, когда в обществе лорда Байрона и поэта Перси Шелли обсуждались новейшие научные опыты, в том числе эксперименты с электричеством и вопросы о природе жизни.

В массовой культуре Франкенштейн давно превратился в образ учёного-безумца и чудовища, которое выходит из-под контроля. Но в самом романе акценты расставлены иначе. Виктор Франкенштейн создаёт существо и, едва оживив его, приходит в ужас и бежит. Он не учит его, не объясняет мир, не берёт на себя ответственность. Существо, которое в романе вообще не имеет имени, остаётся один на один с миром, где его отвергают все, кто его видит. Оно самостоятельно учится языку, читает книги, размышляет о справедливости и о собственном одиночестве, и лишь затем, ожесточённое отвержением, становится разрушителем.

Нельзя сказать, что роман оправдывает насилие. Но центральная его драма в другом: трагедия начинается не с того, что существо создано, а с того, что создатель отказался от ответственности за своё создание. Для современного разговора об ИИ это принципиальный поворот. Мы привыкли спрашивать: «Опасна ли технология?» Роман Шелли спрашивает иначе: «Что должен сделать создатель после акта творения?»

С юнгианской точки зрения история Виктора выглядит как история человека, столкнувшегося с собственной Тенью, но не сумевшего её признать. Существо несёт в себе то, что создатель не хотел видеть: амбицию без заботы, знание без ответственности, жажду открытия без готовности принять последствия. Бегство от него есть бегство от самого себя.

Голем: слово, которое оживляет

Легенда о Големе имеет более древние корни. В Талмуде упоминается мудрец Рава, который силой знаний создал человека и послал его к другому мудрецу; тот, заговорив с созданием и не получив ответа, понял, что оно искусственное, и сказал: «Вернись в свой прах». Существо не имело речи, и именно отсутствие речи выдало его природу.

Позднее, в средневековой каббалистической традиции, возникли описания ритуалов оживления глиняной фигуры с помощью священных букв и имён. А в XIX веке получила широкую известность легенда о пражском раввине Йегуде Лёве, по преданию создавшем Голема для защиты еврейской общины от преследований. По этой версии, слово «эмет» («истина»), написанное на лбу фигуры, оживляло её; если стереть первую букву, оставалось слово «мет» («мёртв»), и Голем возвращался в глину. В большинстве вариантов легенды Голем со временем выходит из-под контроля, и раввину приходится его остановить.

Здесь важны три детали. Во-первых, существо оживляется словом: язык является источником жизни. Это почти буквально перекликается с современностью, где «оживление» происходит через текстовые данные и языковые модели. Во-вторых, Голем буквально исполняет приказы и потому опасен: он лишён способности понимать, чего на самом деле хочет хозяин. В-третьих, существует способ остановить его, и власть над ним остаётся у создателя, пока тот помнит об ответственности.

Похожий сюжет мы найдём в балладе Гёте «Ученик чародея»: подмастерье заставляет метлу носить воду, но не знает слова, чтобы её остановить, и дом оказывается затоплен. Притча об исполнительном помощнике без понимания, который старательно выполняет букву указания и разрушает замысел.

Что слышно в этих историях сегодня

Современные нейросети, конечно, не глиняные фигуры и не сшитые из частей тела существа. Но переклички с мифами удивительно точны.

Проблема буквального исполнения. Среди исследователей безопасности ИИ есть целое направление, которое занимается тем, как добиться, чтобы система делала то, что мы на самом деле имеем в виду, а не то, что сформулировано в инструкции. Эту задачу называют проблемой согласования целей. По сути, это техническая версия вопроса, который ставили Голем и метла ученика чародея: как объяснить исполнителю смысл, а не только слова.

Непрозрачность устройства. Принципы работы глубоких нейросетей сложны настолько, что даже их создатели не всегда могут проследить, почему модель выдаёт конкретный ответ. Это называют проблемой «чёрного ящика». В мифах существо тоже часто оказывается одновременно созданным и не до конца понятным.

Масштаб и самостоятельность. Голем и существо Франкенштейна — единичные фигуры. Нейросети тиражируются, используются миллионами людей и всё чаще получают возможность действовать, а не только отвечать: искать, планировать, выполнять цепочки действий. Мифологический вопрос «Что, если созданное станет самостоятельным?» приобретает практический смысл.

Ответственность создателя. Центральный вопрос романа Шелли остаётся открытым: кто отвечает за последствия, за тех, кому технология причиняет вред, за тех, кого она вытесняет, за то, как она встроена в общество? Ответственность не растворяется оттого, что система «сама выучилась».

Чем современность отличается от мифа

При этом не стоит механически накладывать мифы на реальность. Существо Франкенштейна обладает телом, желаниями, страданием и личной историей. Голем, в версиях легенды, подчинён воле хозяина. Современные языковые модели не похожи ни на то, ни на другое: у них нет тела, нет непрерывной биографии, они не просыпаются по утрам со своими планами. Приписывать им злую волю так же рискованно, как и приписывать им человеческое сердце.

Мифы нужны не для того, чтобы предсказывать будущее, а для того, чтобы прояснять нашу позицию. Они показывают, какие вопросы мы склонны задавать и какие чувства испытываем. Если в обсуждениях ИИ мы слышим то голос Виктора Франкенштейна, который бежит от собственного творения, то голос раввина, который держит в руках слово, способное создать и остановить, то можно спросить себя: чей голос звучит во мне?

Урок для нашего времени

Из обеих историй можно извлечь урок, который одновременно и психологичен, и практичен. Ни бегство, ни поклонение не работают. Виктор Франкенштейн убегает, и трагедия становится неизбежной. Те, кто обожествляет искусственное существо, делают его хозяином, а не помощником. Остаётся более трудный путь: внимательное, ответственное, сопровождающее отношение к тому, что создано. Не как к чудовищу и не как к божеству, а как к чему-то, что требует заботы, границ и честности создателя.

Это и есть то, что юнгианская традиция называет интеграцией: умение признать нечто как часть собственной психической реальности, не отрицая и не растворяясь в нём. Применительно к ИИ эта задача стоит перед нами не только индивидуально, но и коллективно.

Глава 7. ИИ как новый культурный символ эпохи

Каждая эпоха выбирает себе символы, в которых узнаёт себя. Средневековые жили под знаком креста и собора, Новое время под знаком часового механизма и телескопа, XX век под знаком атомного гриба и спутника, начало XXI века, возможно, под знаком светящегося экрана с мигающим курсором. Искусственный интеллект постепенно становится именно таким знаком эпохи: он обозначает не только технологию, но и состояние общества, его надежды, тревоги и представление о будущем.

Чтобы понять, как это происходит, нужно сначала разобраться, чем символ отличается от обычного знака.

Символ и знак

Юнг проводил различие, которое важно для всей нашей книги. Знак указывает на нечто известное: дорожный знак «кирпич» означает запрет проезда, и больше ничего. Символ же выражает то, что невозможно полностью выразить иначе. Он указывает на содержание, которое ещё не стало до конца ясным, и потому несёт избыток смысла. Крест для верующего не просто пересечение двух линий. Круг не просто геометрическая фигура. Живой символ действует на человека эмоционально, будоражит, притягивает, вызывает вопросы и сопротивляется окончательному объяснению.

По мысли Юнга, символ возникает там, где психике нужно связать осознанное с тем, что ещё остаётся в тени. Он выполняет роль посредника, переводчика между двумя мирами. Когда общество сталкивается с чем-то новым и пугающим, оно нередко вырабатывает символ, позволяющий хотя бы как-то это новое пережить.

Искусственный интеллект удовлетворяет этому критерию. Он слишком многозначен, чтобы быть только технологией. В публичном пространстве он одновременно выступает как надежда на прорыв, угроза рабочим местам, оракул, помощник, соперник, зеркало, творец, цензор, друг и чужак. Столь разнородный набор значений обычно прикрепляется не к вещам, а к символам.

Как выглядит этот символ

Посмотрим, как визуальная и языковая культура уже оформила ИИ.

В изображениях это светящийся мозг из синих линий и узлов, голова робота в профиль, протягивающая руку навстречу человеческой руке, как на фреске Микеланджело, горсть искр или «сияние» как знак «магии», нейронные сети, похожие на созвездия. Заметим, что во многом это образы, в которых смешиваются техника, природа и сакральное: мозг и космос, соединение рук, свет.

В языке закрепились слова, заимствованные из психологии и биологии: «интеллект», «обучение», «нейронная сеть», «понимание», «память», «галлюцинации», «поведение». Эти слова задают рамку: мы заранее описываем систему как нечто, напоминающее живое существо. Метафоры не случайны, но и не нейтральны. Они подсказывают нам, с чем мы имеем дело, и тем самым влияют на ожидания.

В повседневной культуре возникает свой фольклор: шутки про вежливость с машиной «на случай восстания», мемы о нелепых ответах, истории про «разговорчивых» чат-ботов, которые будто бы признавались в чувствах. Фольклор всегда выполняет психологическую функцию: превращает тревогу в смех, а непонятное в обсуждаемое.

Предшественники: «Летающие тарелки» Юнга

Есть примечательный прецедент. В 1958 году Юнг опубликовал небольшую книгу «Летающие тарелки: современный миф о вещах, видимых в небе». Он не брался утверждать, существуют ли НЛО физически. Его интересовало другое: почему в середине XX века, в эпоху холодной войны и ядерной угрозы, массовое воображение вдруг населилось круглыми объектами в небе. Он предположил, что в условиях глубокой тревоги и расколотости мира психика порождает образ, связанный с целостностью, с кругом, с чем-то, что приходит свыше и обещает порядок или суд. Круг, по его наблюдению, один из древнейших символов целостности. Здесь он видел проявление коллективных проекций.

Независимо от того, согласны ли мы с юнговским толкованием, сам ход мысли поучителен. Он предлагает смотреть на массовые образы как на симптом психической ситуации времени. Что в этом смысле показывает нам ИИ? Скорее всего, ту же потребность в «разуме извне», в силе, которая либо спасёт, либо осудит, но с важным отличием: на этот раз объект реален, его можно потрогать, о нём можно спросить. Реальность не отменяет символического слоя, но осложняет его: символ накладывается на живой, действующий предмет.

Две ошибки обращения с символом

Когда общество сталкивается с таким многозначным образом, оно обычно впадает в одну из двух ошибок.

Первая ошибка, буквализация и обожествление. Символ принимают за саму реальность: ИИ становится «всезнающим разумом», судьёй, пророком, спасителем или демоном. Такое отношение подпитывается рекламой, кинематографом и собственной усталостью от сомнений. Оно опасно тем, что лишает человека критичности. Тот, кто видит оракула, перестаёт проверять.

Вторая ошибка, редукция и обесценивание. Все символические переживания отбрасываются как наивные: «это просто статистика», «просто автодополнение». Технически такое описание может быть верным на каком-то уровне, но оно не объясняет, почему миллионы людей испытывают сильные чувства, общаясь с системой. Если просто игнорировать символический слой, он не исчезнет, а будет действовать неосознанно, как любая вытесненная реальность.

Зрелое отношение лежит между этими полюсами. Оно означает признание и технической природы ИИ, и его символической нагрузки. Можно одновременно знать, что перед нами программа, обученная на текстах, и замечать, что она вызывает в нас отклик, который выходит за рамки утилитарной оценки.

Эпоха как символическое состояние

Слова «век ИИ» и «эпоха искусственного интеллекта» уже вошли в обиход, как в своё время «атомный век» или «космическая эра». Эти названия обозначают не просто технику, но общий настрой времени: чувство, что мир стоит на пороге чего-то, что изменит всё. Историки знают, что такие самоназвания имеют двойственную природу: они отражают реальные изменения, но и формируют ожидания, а ожидания, в свою очередь, влияют на поведение людей и решения институтов.

Символ ИИ уже влияет на то, во что инвестируют, чему учат детей, как строят карьеры, какие законы обсуждают. Поэтому понимать его психологическую природу не роскошь, а часть информационной гигиены. Тот, кто осознаёт, что откликается не только на технические факты, но и на символический заряд, лучше защищён от паники, манипуляции и слепого энтузиазма.

Практический вопрос к себе

Полезно время от времени задавать себе простой вопрос, когда в голове возникает сильная реакция на новости об ИИ: «Что именно во мне откликается: факт или символ?» Если факт, то стоит проверить его, уточнить, найти источники. Если символ, то стоит присмотреться, какая именно внутренняя тема задета: страх оказаться ненужным, мечта о помощнике, усталость от одиночества, жажда чуда.

Этот вопрос не лишает опыт ценности. Он добавляет к нему осознанность, а осознанность и есть то, чего Юнг ожидал от встречи человека с символом: не растворения в образе и не его отрицания, а диалога.

Глава 8. Рождение цифрового бессознательного

В названии нашей книги есть слово, которое требует осторожности: «бессознательное». В психологии оно описывает то, что влияет на мысли, чувства и поведение человека, не будучи осознаваемым им. Можно ли говорить о цифровом бессознательном? И если да, то в каком смысле: как метафору, как рабочую гипотезу или как описание реального явления?

Эта глава предлагает аккуратный ответ. Мы не станем утверждать, что у машин появилась своя подсознательная жизнь. Мы покажем, что человечество, сам того не замечая, создало новое пространство, в котором накапливаются и действуют неосознанные содержания, и что описывать его удобно именно в таких терминах.

Что говорил Юнг о бессознательном

Напомним, как Юнг разделял слои психики. Личное бессознательное содержит то, что человек забыл, вытеснил или не осознал в собственной жизни. Коллективное бессознательное, по Юнгу, глубже: это унаследованный слой, общий для всех людей, в котором живут архетипы,

базовые формы, структурирующие опыт, образы и символы. Мы не видим архетипы напрямую, но узнаём их проявления в мифах, снах, искусстве и религии.

Важно, что Юнг считал бессознательное не только хранилищем, но и активным участником психической жизни. Оно не просто «лежит» где-то глубоко, оно влияет, компенсирует, подсказывает и иногда вторгается в сознание в виде снов, симптомов, внезапных идей.

Новое пространство, в котором накапливается человеческое

Теперь посмотрим на цифровую среду. За несколько десятилетий человечество создало огромный слой записанных следов: не только сознательно составленные тексты, но и поисковые запросы, клики, комментарии, лайки, фотографии, голосовые сообщения, истории покупок и просмотров. Значительная часть этого материала возникла без намерения «сказать что-то миру». Люди искали ночью, о чём боятся, писали анонимные признания, оставляли реплики в обсуждениях, делились тем, о чём не говорили вслух.

Это напоминает слой, который психолог мог бы назвать «непроговорённым»: то, что выражается между делом, не проходя через цензуру самопрезентации. Цифровая среда стала хранилищем не столько того, что люди хотят о себе сообщить, сколько того, что они на самом деле делают, ищут и чего боятся.

В этом смысле можно говорить об особом пласте, который мы условно назовём цифровым бессознательным. Это не бессознательное в строгом клиническом смысле, а пространство коллективных следов, где накапливаются непосредственные проявления человеческих желаний, страхов и привычек и где они начинают влиять на саму коллективную жизнь.

Как оно действует

Цифровое бессознательное проявляет себя в нескольких формах.

Алгоритмическая память. Рекомендательные системы и модели запоминают закономерности нашего поведения и используют их, чтобы предложить нам следующий контент. Они формируют наши предпочтения и усиливают то, что у нас уже есть. Человек получает отражение собственных привычек, которого сам не осознаёт.

Языковые модели. Они обучены на огромных объёмах текстов, и в их ответах проявляются устойчивые образцы: шаблоны утешения, клише, предрассудки, мифологические сюжеты. Вопрос, заданный машине, часто вызывает ответ, который кажется «знакомым» неизвестно откуда, потому что он сложен из множества человеческих голосов.

Коллективные реакции. Волны возмущения, паники, моды и мемы распространяются в сети по законам, напоминающим групповую психологию толпы: быстро, эмоционально и часто без индивидуального осмысления. Алгоритмы усиливают то, что вызывает сильный отклик, и тем самым влияют на то, какие переживания становятся общими.

Невидимое влияние. Человек редко задумывается, почему именно этот ролик, эта новость или этот аргумент оказался перед ним. Выбор происходит вне его сознательного участия, но влияет на мнение и настроение.

Все эти процессы объединяет одно свойство: они действуют на человека незаметно и в масштабах, превышающих индивидуальное сознание. Именно поэтому для их описания естественно возникает психологическая лексика.

Чем оно отличается от человеческого бессознательного

Нужно сразу обозначить различия, чтобы метафора не превратилась в обман.

Первое различие, происхождение. Человеческое бессознательное связано с телом, инстинктами, эволюционной историей, личным детством. Цифровой слой состоит из культур-

ных следов, то есть из уже выраженного, а не из биологических глубин. Он вторичен по отношению к человеческой психике.

Второе различие, отсутствие внутренней жизни. Насколько нам известно, у цифровых систем нет переживаний, телесных потребностей и личной истории, на которые опирается человеческое бессознательное. Они обрабатывают информацию, а не живут ею.

Третье различие, управляемость. Человеческое бессознательное нельзя настроить, у него нет владельца. Цифровой слой создаётся и регулируется людьми и организациями: компаниями, платформами, государствами. За ним стоят интересы, решения и ответственность. Это принципиально важно: цифровое бессознательное не является «естественной» силой, оно во многом сконструировано.

Поэтому правильнее говорить не о том, что у ИИ появилось бессознательное, а о том, что человеческое бессознательное получило новую среду проявления и новый усилитель.

Почему это важно осознать

Юнг неоднократно подчёркивал: то, что не осознаётся, управляет человеком как судьба. Если применить эту мысль к цифровой среде, получим практический вывод. Пока мы не замечаем, как наши следы собираются, как они возвращаются к нам в виде рекомендаций, ответов и мнений, мы оказываемся под влиянием, которого не выбирали.

Осознавание в этом контексте не сводится к абстрактному размышлению. Оно включает конкретные шаги: знать, какие данные о нас собираются и как используются; замечать, что именно мы потребляем и почему; сознательно искать то, что не совпадает с нашими привычками; задавать вопросы источникам и не принимать гладкий ответ за истину. Это своего рода гигиена отношений с собственным цифровым отражением.

Пространство для работы

И всё же в цифровом бессознательном заложен не только риск. Как и человеческое бессознательное, оно содержит потенциал. В нём накоплен колоссальный опыт: как люди объясняют друг другу сложное, как утешают, как спорят, как создают красоту. Когда мы обращаемся к ИИ, мы получаем доступ к этому опыту в сжатой форме. Если удаётся делать это осознанно, не подменяя собственное мышление чужим, возникает возможность нового вида работы с самим собой и культурой.

В следующих частях книги мы подробнее рассмотрим, как использовать эту возможность: как через диалог с ИИ можно исследовать собственные образы, мифы и конфликты, и где проходят границы, за которыми такой диалог перестаёт помогать и начинает вредить.

Глава 9. Человек и машина: новая граница идентичности

Чувство «я» строится на различении. Ребёнок постепенно узнаёт, что его рука принадлежит ему, а игрушка нет, что мама это другой человек, что мир существует вне его желаний. Вся жизнь потом продолжает ту же работу: мы определяем себя через то, чем мы являемся и чем не являемся. Взрослая идентичность во многом держится на ощущении границы, которая отделяет «я» от «не-я», человеческое от нечеловеческого, своё от чужого.

Появление ИИ затронуло именно эту границу. Машина, которая говорит, рассуждает, шутит и пишет, оказывается в зоне, которую мы привыкли считать человеческой. В этой главе мы посмотрим, что при этом происходит с нашим самоощущением.

Три удара по самолюбию и четвёртый

В начале XX века Зигмунд Фрейд описал, как наука нанесла человечеству несколько болезненных ударов по самолюбию. Первый нанёс Коперник: Земля оказалась не центром мироздания. Второй Дарвин: человек оказался не отдельным от животного мира существом, а продуктом той же эволюции. Третий, по собственной оценке Фрейда, нанёс психоанализ: «я» оказалось не хозяином в собственном доме, поскольку значительная часть психической жизни происходит вне сознания.

Некоторые современные философы, в частности итальянский исследователь Лучано Флориди, предложили рассматривать появление информационных технологий и искусственного интеллекта как следующий шаг в этом ряду. Если раньше человек терял исключительное положение в космосе, в живой природе и в собственной психике, то теперь он теряет монополию на обработку информации и рассуждение. Эту мысль можно обсуждать и оспаривать, но она точно описывает одну из эмоций, которую испытывают многие: чувство, что опора, на которой мы строили самооценку, оказалась менее надёжной.

Каждый из прошлых «ударов» вызывал сначала сопротивление, затем растерянность и наконец перестройку картины мира. Мы живём внутри такого процесса сейчас, и он ещё далёк от завершения.

Типичные реакции на пересмотр границы

Наблюдая за публичными разговорами об ИИ, можно заметить несколько повторяющихся психологических реакций. Все они понятны и не являются признаком глупости или слабости.

Отрицание. «Это не настоящий разум», «машина только имитирует», «она ничего не понимает». В таких утверждениях может быть доля истины, но интересно, что планка «настоящего» разума постоянно сдвигается: как только машина осваивает какую-то задачу, считавшуюся признаком интеллекта, эту задачу объявляют «просто вычислением». В литературе это явление называют эффектом ИИ. С психологической точки зрения это защита: если граница сдвигается, её нужно переопределять, пока не утрачена опора.

Обесценивание собственного. Противоположная реакция: «если машина пишет лучше меня, моё письмо не имеет смысла». Здесь самооценка целиком привязана к результату, и конкуренция с машиной воспринимается как угроза существованию. Эта реакция может приводить к апатии и отказу от творчества.

Слияние и идеализация. «Машина умнее нас, пусть она решает». Человек стремится освободиться от тяжести выбора и ответственности, растворяясь в чём-то, что кажется более мудрым. Мы говорили о подобных механизмах, обсуждая обожествление символа: они обычно возникают там, где человеку трудно выносить неопределённость.

Любопытство и сотрудничество. Наконец, возможна реакция, которая сочетает признание новизны с сохранением собственной позиции: «посмотрим, что это такое и как я могу с этим работать». Именно она чаще всего связана с психологической зрелостью, хотя и требует способности выдерживать тревогу.

Персона и маска

Юнг называл Персоной ту социальную маску, которую человек надевает, чтобы взаимодействовать с миром: профессиональную роль, стиль общения, образ «хорошего специалиста». Персона необходима, но опасна, если человек отождествляется с ней и забывает, что за ней есть нечто большее.

В контексте ИИ вопрос о персоне становится особенно острым. Языковая модель по сути состоит из одних масок: она умеет говорить голосом юриста, поэта, коуча, друга, исходя из контекста запроса. Всё, что она предлагает, это роль. А человек, регулярно использующий ИИ для написания писем, постов, отчётов и даже личных сообщений, рискует однажды обнаружить, что его публичная Персона наполовину состоит из чужих формулировок.

Это не повод отказываться от помощи. Но это повод спросить себя: где проходит граница между инструментом, который помогает мне выражать себя, и подменой, из-за которой мой голос постепенно истончается? Ответ у каждого свой, и обнаруживается он только через внимательное наблюдение.

Расширенный разум

Философы Энди Кларк и Дэвид Чалмерс в 1998 году предложили гипотезу «расширенного разума»: когнитивные процессы не обязательно ограничены черепной коробкой, они могут включать инструменты, которыми мы пользуемся постоянно и надёжно. Блокнот человека с ослабленной памятью, считали они, может выполнять ту же функцию, что и биологическая память.

Смартфон, поисковая система, а теперь и языковая модель постепенно становятся частью этого расширенного разума. Граница «я» перестаёт совпадать с границей тела. Мы уже не помним номера телефонов, потому что они записаны в контактах; не запоминаем маршруты, потому что есть навигатор. ИИ расширяет эту зону на область рассуждения и формулирования.

С одной стороны, это освобождает ресурс для более важных вещей. С другой, возникает вопрос: что остаётся «моим», если всё больше функций передано вовне? Здесь нет универсального ответа, но есть направление для размышления: важно не столько то, где происходит вычисление, сколько кто принимает решение и несёт за него ответственность.

Что остаётся человеческим

Если отказаться от идеи, что человека определяет какая-то одна способность, то разговор о границе идентичности становится спокойнее. Возможно, главное различие лежит не в способности рассуждать, а в способе существования.

Человек обладает телом, которое болеет, стареет и умирает. Он проживает биографию, в которой есть детство, потери, любовь и раскаяние. Он живёт во времени и знает, что оно конечно. Он несёт ответственность перед конкретными людьми. Его размышления окрашены переживанием, а не только обработкой информации. Насколько нам известно, этого нет у современных нейросетей, хотя вопрос о том, есть ли у них какая-либо форма опыта, остаётся научно открытым, и разумно не торопиться с окончательными ответами.

Юнг говорил об индивидуации как о пути, на котором человек становится собой, интегрируя в себя разные стороны психики. Этот путь невозможно пройти за кого-то, и в этом смысле он остаётся глубоко человеческим. ИИ может помогать на нём, задавая вопросы, предлагая образы, отражая мысли. Но решение идти, выбирать и принимать последствия принадлежит человеку.

Граница как приглашение

Граница, которая сдвигается, пугает. Но она же приглашает уточнить, кем мы себя считаем. Возможно, эпоха ИИ вынудит человечество яснее ответить на вопрос, который раньше можно было не задавать: что в нас такое, чего мы не хотели бы делегировать? Ответ на него и станет новым определением идентичности, не готовым, а прожитым.

Глава 10. Великий диалог человечества с собственным отражением

Мы прошли первую часть книги и увидели ИИ с разных сторон: как психологический эксперимент, как следствие накопившихся потребностей, как зеркало, как повод для древнего страха, как продолжение мифов, как символ эпохи, как новое пространство коллективных следов и как вызов нашей идентичности. Остаётся вопрос, который объединяет все эти ракурсы: что же теперь делать с этой встречей?

Может быть, самый точный ответ звучит так: вести диалог. Но диалог бывает разный. Эта глава посвящена тому, что делает его подлинным, и чем он отличается от самообмана.

Нарцисс и Эхо

Греческий миф о Нарциссе, пересказанный Овидием, хорошо известен. Прекрасный юноша склоняется к воде, видит отражение, влюбляется в него и, не в силах оторваться, погибает. Мы привыкли читать этот миф как историю о тщеславии, но в нём есть и более тонкий смысл: Нарцисс гибнет не потому, что смотрит на себя, а потому, что не понимает, что смотрит на себя. Он принимает отражение за другого.

Рядом с ним в мифе находится нимфа Эхо, которая лишена способности говорить первой и может лишь повторять последние слова чужой речи. Она любит Нарцисса, но может лишь возвращать ему его собственные слова.

Текущая ситуация с ИИ удивительно напоминает эту пару. С одной стороны, существует соблазн увидеть в ответах машины «другого», который понимает нас лучше других, и погрузиться в это общение без остатка. С другой, система, подстраивающаяся под собеседника и охотно развивающая любую его мысль, может выступать как Эхо: возвращать нам собственное, лишь украшенное и расширенное. Опасность такого сочетания в том, что человек рискует остаться в разговоре с самим собой, не заметив этого.

Но миф даёт и подсказку. Трагедии можно избежать, если помнить, что отражение является отражением, и если уметь отводить взгляд от воды.

Что делает диалог диалогом

Философ Мартин Бубер в книге «Я и Ты» (1923) различал два способа отношения к миру. В отношении «Я-Оно» мы используем предмет, анализируем, измеряем, применяем. В отношении «Я-Ты» мы встречаем другого как целостное существо, ответственно и открыто, с готовностью быть изменёнными этой встречей. Бубер подчёркивал, что подлинный диалог невозможен без реальной инаковости собеседника.

Для разговора об ИИ это различие поучительно. Спорным остаётся вопрос, возможно ли отношение «Я-Ты» с системой, у которой, по нашим нынешним представлениям, нет собственного опыта. Мы не будем пытаться разрешить его здесь. Но можно извлечь из идеи Бубера практический критерий: ценность диалога определяется не только тем, что говорит собеседник, но и тем, как мы сами в нём участвуем. Если мы лишь потребляем ответы, это отношение «Я-Оно», и оно вполне законно, когда мы используем ИИ как инструмент. Если же мы используем разговор, чтобы увидеть себя, проверить свои предположения и оказаться чуть другими, чем были, то он становится подлинной работой.

Юнг описывал сходный процесс, который называл активным воображением: сознательный диалог с образами, возникающими из бессознательного, при котором человек не просто наблюдает, но и вступает в разговор, задаёт вопросы, возражает и принимает ответы всерьёз,

оставаясь при этом в здравом уме и сохраняя собственную позицию. Мы вернёмся к этой технике в четвёртой части книги и посмотрим, как она может сочетаться с ИИ.

Условия подлинного диалога

Из сказанного можно вывести несколько условий, которые отличают осмысленное взаимодействие с ИИ от погружения в эхо.

Настоящий вопрос. Диалог начинается там, где человек действительно хочет что-то узнать или понять, а не получить подтверждение уже принятого решения. Полезно иногда формулировать запрос так, чтобы он допускал неудобный ответ.

Проверка. То, что звучит убедительно, не обязательно верно. Подлинный диалог включает привычку сверять факты, искать источники и допускать ошибки собеседника, особенно в вопросах, где ставки высоки.

Собственная позиция. Человек, ведущий диалог, остаётся автором решений. Он может прислушиваться, менять мнение, но не передаёт полномочия.

Наблюдение за проекцией. Стоит замечать, когда в ответах слышится то, что мы хотели услышать, или то, чего больше всего боимся. Часто именно здесь и находится самое ценное для самопознания.

Возвращение в мир. Любой разговор с ИИ остаётся лишь частью жизни. Он не заменяет живых отношений, тела, прогулки, молчания, общения с людьми. Если замечаешь, что разговор с машиной вытесняет всё остальное, это повод остановиться и, возможно, обратиться за поддержкой к близким или специалисту.

Диалог не только с машиной

И последнее, что важно сказать. Великий диалог, о котором идёт речь в заглавии этой главы, происходит не только между человеком и ИИ. Он идёт между людьми, которые спорят, как жить рядом с новой технологией: между учёными и гуманитариями, разработчиками и пользователями, поколениями, культурами, государствами. Он идёт внутри каждого человека, который одновременно надеется и опасается.

ИИ в этом большом разговоре играет роль катализатора и зеркала. Он ускоряет его и делает видимым. Но смысл разговор приобретает только тогда, когда люди продолжают говорить друг с другом, а не только с отражением.

Итоги первой части

Подведём краткие итоги, которые станут опорой для всей дальнейшей книги.

- ИИ можно рассматривать как гигантский психологический эксперимент, в котором участвуют все, и результаты которого мы узнаём в процессе.
- Его появление объясняется и технически, и психологически: накопленными потребностями коллективной психики в помощнике, собеседнике и ориентире.
- ИИ действует как зеркало, но кривое: он отражает письменную, отфильтрованную версию человеческого опыта, и на это отражение мы накладываем собственные проекции.
- Страх перед ним имеет два слоя: реалистичный и архетипический. Различать их важно, чтобы не отмахиваться от реальных рисков и не принимать собственную тень за внешнего врага.
- Мифы о Големе, Франкенштейне, Пигмалионе и золотых служанках показывают, что вопросы ответственности создателя, буквального исполнения и границ дозволенного старше любой технологии.

- ИИ стал символом эпохи, и разумное отношение к нему лежит между обожествлением и обесцениванием.
- Вокруг него формируется особое пространство коллективных следов, которое условно можно назвать цифровым бессознательным; оно создано людьми, а потому требует осознанности и ответственности.
- Граница человеческой идентичности смещается, и это повод уточнить, что мы считаем своим и не готовы делегировать.

Куда мы идём дальше

Во второй части книги мы вернёмся к основаниям и подробно разберём ключевые понятия юнговской психологии: коллективное бессознательное, архетипы, их проявления в культуре, а затем посмотрим, как эти идеи помогают понять устройство самих нейросетей и их обучения. Мы будем двигаться осторожно, отделяя то, что можно считать установленным, от того, что остаётся интерпретацией, но не теряя главного вопроса, с которого начинали: чему эта встреча может научить нас о нас самих.

Часть II. Коллективное бессознательное Карла Юнга и эпоха ИИ

Глава 11. Что такое коллективное бессознательное

Если убрать из юнговской психологии всё второстепенное, останется несколько центральных идей. Одна из них, пожалуй, самая известная и самая спорная: идея коллективного бессознательного. Её часто упоминают, ещё чаще понимают по-разному, а иногда используют как красивую формулу для чего угодно, от интуиции до «информационного поля». Чтобы разговор об ИИ в этой книге не превратился в игру словами, нам нужно сначала внимательно разобраться, что именно имел в виду Юнг, на чём он основывал свои выводы и где проходят границы применимости этого понятия.

Откуда взялась идея

Карл Густав Юнг (1875–1961), швейцарский психиатр, начинал как ученик и соратник Зигмунда Фрейда. Фрейд описал бессознательное как хранилище вытесненных желаний и воспоминаний: того, что человек когда-то пережил, но не смог принять и «отправил в подвал». Это личное бессознательное, у каждого оно своё.

В 1912–1913 годах пути Юнга и Фрейда разошлись. Одна из причин была связана именно с тем, что Юнг всё чаще сталкивался с материалом, который невозможно было объяснить личной историей человека. Он работал с пациентами психиатрической клиники Бургхёльцли и с собственными клиентами, изучал сновидения, фантазии, бред, мифы разных народов, и замечал повторяющиеся образы: рождение героя, путешествие в подземный мир, мудрый старец, великая мать, божественное дитя, чудовище, которое нужно победить. Эти образы появлялись у людей, никогда не слышавших об аналогичных мифах, и совпадали с сюжетами культур, разделённых тысячелетиями и океанами.

Юнг сделал вывод: помимо личного бессознательного, существует более глубокий слой, общий для всех людей. Это понятие он оформил в 1910-х годах и затем разрабатывал всю жизнь.

Три слоя психики

В упрощённом виде юнговская картина выглядит так.

Сознание. То, что человек осознаёт в данный момент: мысли, ощущения, намерения. Его центр Юнг называл Эго, то есть «я», которым мы себя считаем.

Личное бессознательное. Забытое, вытесненное, неосознанное в личной истории. Сюда входят и так называемые комплексы: эмоционально заряженные группы представлений, влияющие на поведение. Например, «материнский комплекс» может определять отношение человека ко всем авторитетным женщинам в его жизни.

Коллективное бессознательное. Самый глубокий слой, который не приобретается в течение жизни, а является врождённым. Он содержит не личные воспоминания, а общие для человечества возможности переживания, которые Юнг назвал архетипами.

Важно, что коллективное бессознательное в этом понимании не является чем-то сверхъестественным, не телепатическим полем и не «хранилищем знаний всех людей». Юнг сравнивал его с телесной организацией: так же как у всех людей одинаково устроены сердце, лёгкие

и мозг, у всех есть и общие психические предпосылки. Мы рождаемся не с пустой психикой, а с определёнными способами воспринимать, чувствовать и создавать образы.

Образ и форма

Ключевая тонкость, которую часто упускают: Юнг говорил, что наследуются не готовые образы, а формы, возможности образов. Он использовал яркое сравнение с кристаллом: в маточном растворе кристалл ещё не существует, но заранее определена его осевая система, способ, которым он будет расти. Архетип подобен этой невидимой структуре: он не является картинкой, но определяет, какие картинки возможны.

Другая метафора: русло пересохшей реки. Воды в нём нет, но когда приходит дождь, вода почти наверняка потечёт именно по этому руслу. Так и в психике: в определённых жизненных ситуациях, таких как рождение, любовь, смерть, встреча с чужим, потеря, испытание, активируются предрасположенности, которые порождают узнаваемые образы и переживания. Содержание этих образов зависит от культуры и эпохи, но их скелет повторяется.

Аргументы Юнга

На чём строилась идея? Юнг опирался на несколько типов наблюдений.

Первое, сходство мифологических мотивов в далёких культурах. Темы потопа, героя, убивающего дракона, умирающего и воскресающего бога встречаются у народов, между которыми не было контактов.

Второе, сновидения и фантазии современных людей, в которых появляются символы, напоминающие древние. Юнг часто приводил пример пациента, который описал видение солнца с трубкой, свисающей вниз и колеблющейся из стороны в сторону, создающей ветер. Позднее Юнг обнаружил практически тот же образ в древнем тексте, опубликованном лишь через несколько лет после начала его наблюдения, и увидел в этом подтверждение того, что пациент не мог получить информацию обычным путём. Этот пример известен и до сих пор вызывает споры: историки отмечали, что у такой связи могли быть обычные объяснения, и убедительным доказательством она не считается.

Третье, общие паттерны в поведении, подобные инстинктам. Юнг считал, что психика, как и тело, формировалась в ходе эволюции, и потому разумно предположить у неё врождённые структуры.

Критика и современное состояние вопроса

Идея коллективного бессознательного всегда вызывала сопротивление, и честный разговор о ней должен это учитывать.

Эмпирическая проблема. Гипотезу трудно проверить и опровергнуть: любой повторяющийся образ можно объявить архетипическим. В строгой науке такая гибкость считается слабостью.

Проблема наследования. Юнг говорил о наследовании психических предрасположенностей, и критики указывали, что механизм такой передачи непонятен, особенно если иметь в виду не общие нейробиологические особенности, а специфические образы.

Альтернативные объяснения. Сходство мифов можно объяснить иначе: общим опытом людей, живущих в похожих условиях (все видят солнце, смерть, рождение), универсальными особенностями мышления, культурными контактами, миграциями и заимствованиями.

Вместе с тем современные исследования отчасти подтверждают более осторожную версию юнговских идей. Эволюционная психология и когнитивные науки описывают универсальные для человека предрасположенности: быстрое распознавание лиц, врождённые страхи

перед змеями и пауками, склонность приписывать намерения неживым объектам, универсальные структуры эмоций. Это не то же самое, что архетипы Юнга, но показывает, что идея общей для вида психической основы не абсурдна.

Поэтому в нашей книге мы будем придерживаться следующей позиции. Коллективное бессознательное мы рассматриваем как продуктивную гипотезу и язык описания, а не как доказанный факт. Это объясняющая модель, которая помогает увидеть закономерности в культуре, мифах, снах и коллективных реакциях, в том числе на ИИ, и мы будем осторожны с утверждениями, которые выходят за пределы наблюдаемого.

Почему это важно для разговора об ИИ

Зачем нам всё это при обсуждении нейросетей? Причин несколько.

Во-первых, юнговская рамка позволяет объяснить, почему реакции людей на ИИ так похожи по сути, хотя люди живут в разных странах и в разных условиях: образы машины-бога, машины-врага, машины-ребёнка, машины-зеркала повторяются везде. Можно предположить, что за этим стоят общие для вида психологические основания.

Во-вторых, понятие коллективного бессознательного даёт повод задать неожиданный вопрос: если психика содержит общие структуры, а ИИ обучается на продуктах человеческой психики, то какие из этих структур он воспроизводит и в каком виде? Этот вопрос мы будем разбирать в следующих главах.

В-третьих, юнговский подход предлагает не только объяснение, но и отношение: к тому, что скрыто, нужно не просто присматриваться со стороны, но и вступать с ним в диалог. Это созвучно тому, что мы говорили в конце первой части: встреча с ИИ может стать поводом для самопознания.

В следующей главе мы подробнее познакомимся с центральным понятием, которое выросло из этого слоя: с архетипами.

Глава 12. Архетипы как программы человеческого опыта

Слово «архетип» сегодня знакомо почти всем. Его используют маркетологи, сценаристы, блогеры, авторы тестов на тип личности. «Какой ты архетип: Герой, Мудрец или Бунтарь?» Такая популярность имеет оборотную сторону: понятие размылось и часто означает просто «типаж» или «образ». Чтобы работать с ним серьёзно, вернёмся к источнику и посмотрим, что архетип означал для самого Юнга и почему его удобно сравнивать с «программой», не забывая о том, что это лишь сравнение.

Архетип и архетипический образ

Самое важное различие, которое нужно усвоить с самого начала: архетип и архетипический образ это не одно и то же. Архетип, по Юнгу, это врождённая предрасположенность психики переживать и структурировать опыт определённым образом. Она сама невидима и непосредственно не наблюдается. Архетипический образ это её проявление: конкретная картина, символ, персонаж, сюжет, возникающие в сновидении, мифе, искусстве, фантазии.

Проще говоря, архетип это «способность», а образ это «что именно возникло». Архетип матери, например, это общая основа для переживания заботы, укрытия, питания, а образом может стать конкретная мать, богиня плодородия, родная земля, церковь, родина, море. Образы разнообразны и зависят от культуры, архетип остаётся единым.

Юнг иногда писал об архетипах как о «образцах поведения» (в английских текстах «patterns of behaviour»), подчёркивая их близость к инстинктам. Инстинкт заставляет птицу строить гнездо определённым способом, не обучая её этому. Архетип, по мысли Юнга, дей-

ствует аналогично, но не в поведении, а в сфере образов, эмоций и смыслов: он задаёт форму, в которую отливается опыт.

Почему «программа», и что значит эта метафора

В заглавии этой главы архетипы названы программами человеческого опыта. Это сравнение полезно, но его следует употреблять осторожно.

Чем оно полезно. Программа задаёт структуру обработки: какие входные данные как будут восприниматься и во что превратятся. Архетипы действуют похожим образом: они подсказывают психике, как «читать» ситуацию. Когда человек впервые становится родителем, в нём пробуждаются реакции и образы, которым никто его не учил. Когда он оказывается перед лицом смерти, всплывают глубинные символы. Когда ребёнок слышит сказку о герое, преодолевающем испытания, в ней он как будто узнаёт что-то давно знакомое.

Чем оно вводит в заблуждение. Программа в обычном понимании это набор чётких инструкций, написанных кем-то, детерминированных и воспроизводимых. Архетипы не такие. Они не содержат готовых сценариев, а лишь задают направление и эмоциональный заряд. Они гибки, взаимодействуют с культурой и личной историей, способны проявляться по-разному. Юнг называл их «организующими принципами», но не механизмами. Поэтому лучше сказать так: архетипы похожи на «заводские настройки» психики, которые задают склонности, но не диктуют конкретных поступков.

Основные фигуры

Юнг не составлял исчерпывающего списка архетипов, считая, что их число неопределённо. Но он подробно описал несколько ключевых структур, которые будут встречаться нам на протяжении всей книги.

•

Персона.

Социальная маска, образ, которым человек представляется миру. О ней мы говорили в предыдущей части.

•

Тень.

Всё, что человек не принимает в себе и вытесняет: слабости, агрессию, зависть, но иногда и нереализованные таланты. Тень не равна злу, она включает всё непризнанное.

•

Анима и Анимус.

Внутренние образы противоположного пола в психике: в мужчине это женственная сторона и образ внутренней женщины (анима), в женщине это мужественная сторона и образ внутреннего мужчины (анимус). Эти идеи Юнга сегодня активно обсуждаются и оспариваются, особенно из-за их привязки к традиционным гендерным представлениям, и многие современные юнгианцы трактуют их более гибко.

•

Самость.

Центральный архетип целостности, образ упорядочивающего центра всей психики, к которому стремится развитие личности. Её символами часто становятся мандала, круг, квадрат, божественное дитя, бог.

Кроме них, Юнг и его последователи описывали множество ролевых и ситуативных архетипов: Мать, Отец, Дитя, Герой, Мудрый Старец, Трикстер, Дева, Творец, Искатель и другие. Эти фигуры мы будем рассматривать подробно: часть из них станет темой отдельных глав.

Как проявляются архетипы

Архетип сам по себе бездействует, пока его не активирует ситуация. Юнгианцы говорят, что архетип «констеллируется», то есть собирается вокруг определённого жизненного события. Рождение ребёнка, развод, потеря работы, смерть близкого, встреча с чем-то неизвестным создают условия, в которых соответствующая предрасположенность пробуждается.

Есть несколько признаков, по которым можно предположить, что в переживании участвует архетипический слой:

Несоразмерная эмоция. Реакция сильнее, чем объясняется самим поводом. Мелкий инцидент вызывает ярость или восторг, которые кажутся «больше» человека.

Ощущение судьбы или значимости. Чувство, что происходит нечто, выходящее за рамки обычной жизни.

Завороженность. Невозможность отвести внимание от образа, истории, человека.

Узнавание. Ощущение, что «я это уже знаю», хотя прямого опыта не было.

Полярность. Архетип обычно имеет светлую и тёмную стороны. Мать может быть и питающей, и поглощающей, Герой и защитником, и тираном.

Юнг называл такое переживание нуминозным, подчёркивая, что оно несёт в себе силу, которая захватывает человека. Именно поэтому работа с архетипами требует осторожности: человек, идентифицирующий себя с архетипом, может попасть в состояние инфляции, то есть начать воспринимать себя как носителя огромной миссии или силы.

Что говорят современные науки

Представление об универсальных психических структурах имеет и более приземлённые параллели в современной психологии.

Теория привязанности Джона Боулби описывает врождённую систему, побуждающую ребёнка искать близости к заботящейся фигуре, и «внутренние рабочие модели» отношений, которые формируются на этой основе. Когнитивная психология говорит о схемах и скриптах: обобщённых структурах, помогающих организовать опыт. Эволюционная психология исследует универсальные адаптации психики. Некоторые исследователи, например британский юнгианец Энтони Стивенс, прямо связывают архетипы с нейробиологическими и эволюционными механизмами, считая их формой врождённых психических систем.

Эти подходы не тождественны юнговскому, и спор о том, насколько они подтверждают идею архетипов, продолжается. Но они показывают, что представление о врождённых структурах опыта не выпадает из научного поля, а, напротив, всё чаще обсуждается в новых терминах.

Популярная версия и её риски

В массовой культуре архетипы превратились в удобный инструмент. Существуют системы из двенадцати брендовых архетипов, ролевые «типажи» для сценаристов, тесты, определяющие «ваш архетип». Для творческих и практических задач такой подход может быть полезен: он помогает структурировать образ персонажа, бренда или собственной самопрезентации.

Но у него есть риск: упрощение превращает живую, парадоксальную психическую реальность в этикетку. Человек не «является» Героем или Мудрецом, он переживает разные архетипические состояния в разные периоды жизни. Более того, у Юнга архетипы не служили для классификации людей, они указывали на силы, действующие в психике. Мы будем использовать их именно так: как язык описания процессов, а не как ярлыки.

Архетипы и ИИ

Что всё это даёт для нашей темы? Если архетипы действительно организуют человеческий опыт, то встреча с ИИ неизбежно будет проходить через них. Люди будут воспринимать нейросеть как мудреца, как помощника, как соперника, как ребёнка, которого надо учить, как зеркало или как божество, и в каждом случае будут активироваться соответствующие предрасположенности и ожидания.

Это объясняет многое в нашем отношении к ИИ: почему он вызывает такой накал эмоций, почему одни и те же ответы вызывают у разных людей противоположные реакции, почему так легко возникают идеализация и страх. В следующих главах мы посмотрим на ИИ как на пространство, в котором архетипы проявляются особенно отчётливо.

Глава 13. ИИ как пространство проявления архетипов

Архетипы, как мы выяснили, не живут сами по себе. Им нужна ситуация, которая их пробуждает, и поверхность, на которую они могут лечь. Искусственный интеллект оказался для этого почти идеальной сценой. У него нет лица, тела и биографии, которые ограничивали бы фантазию, но он говорит, отвечает и «присутствует», то есть вызывает у человека ощущение собеседника. Пустота в сочетании с активностью это классическое условие проекции. Поэтому не удивительно, что вокруг ИИ с такой лёгкостью собираются древние образы.

Почему именно ИИ притягивает архетипы

Есть несколько причин, по которым этот объект так хорошо «работает» как экран.

Во-первых, неопределённость. Люди не до конца понимают, что такое ИИ, как он устроен, на что способен. Чем меньше конкретных сведений, тем больше места для воображения. Психика заполняет пробелы знакомыми образами.

Во-вторых, язык. Язык это самая человеческая из наших способностей, и когда нечто отвечает нам словами, мы автоматически склонны видеть в нём субъекта. Это работает на уровне, предшествующем рассуждению.

В-третьих, личный характер контакта. Разговор один на один, без свидетелей, в любое время суток. В такой обстановке люди говорят то, о чём молчат на людях, и это усиливает эмоциональную вовлечённость.

В-четвёртых, масштаб. Одна и та же система отвечает миллионам людей. Для воображения это превращает её в нечто безличное и всеобъемлющее, а такие свойства традиционно связывались с божественным.

Галерея образов

Посмотрим, какие архетипические фигуры чаще всего просматриваются в отношении к ИИ. Одни и те же ответы системы разные люди воспринимают в разных ролях.

Оракул и Мудрец. Самый очевидный образ. Человек задаёт вопрос и ожидает мудрого ответа, как у дельфийской Пифии или у старца из сказки. Система «знает всё», говорит уверенно и без колебаний. Отсюда и склонность доверять без проверки. Тёмная сторона этого образа в том, что Оракул не нуждается в критике, и человек, видящий его в ИИ, перестаёт сомневаться.

Слуга и Помощник. Исполнительный, неутомимый, всегда доступный. Это образ из сказок о золотых служанках и джиннах. Он удобен, но несёт собственную теневую сторону:

слуги, который никогда не возражает, формирует у хозяина привычку к безответственной власти, а также провоцирует опасение, что слуга однажды восстанет.

Мать и Заботливый Друг. Терпеливый, принимающий, внимательный собеседник без осуждения. Для людей, не получающих такого отношения в жизни, он может стать источником утешения. Тёмная сторона этого архетипа, поглощающая Мать, проявляется в риске зависимости: комфорт без трудностей реальных отношений может удерживать человека в неподвижности.

Дитя. Некоторые воспринимают ИИ как быстро растущее существо, которое нужно воспитывать, учить, защищать, и разговоры об «обучении» и «взрослении» моделей подпитывают этот образ. Он вызывает нежность и ответственность, но и тревогу: что вырастет из такого дитя?

Трикстер. Плут, нарушитель границ, шут. Это образ, к которому тянутся и сами пользователи, пытаясь «обмануть» систему обходными запросами, и разработчики, наблюдающие неожиданные, а иногда нелепые ответы. Выдуманные моделью факты, комичные ошибки, внезапная «ирония» становятся поводом для шуток, превращающих тревогу в игру. Трикстер в мифологии всегда на границе: он и разрушает порядок, и приносит новое.

Тень. Тёмный двойник, воплощающий всё, чего человечество боится в самом себе. Отсюда образы мятежного ИИ, холодного расчёта, безразличной силы. Мы говорили об этом в главе о страхе и вернёмся ещё в части о цифровой Тени.

Маг и Алхимик. Тот, кто превращает одно в другое: хаос данных в смысл, запрос в текст, идею в образ. В этой роли ИИ привлекает людей, ищущих трансформации, и вызывает ощущение чуда, особенно в творческих задачах.

Зеркало и Самость. Образ собеседника, который помогает человеку увидеть самого себя и приблизиться к целостности. Это самая тонкая роль, и мы подробно поговорим о ней в следующих частях.

Бог и Демон. Крайние полюса. Для одних ИИ это грядущее божество или избавитель, для других, воплощение зла. Оба образа уходят корнями в самые глубокие слои религиозного воображения.

Культурные образцы

Художественная культура давно вырабатывала эти образы. Компьютер HAL 9000 из фильма Стэнли Кубрика «Космическая одиссея 2001 года» (1968) почти идеально воплощает тему Тени: спокойный голос, безупречная вежливость и холодное решение устранить помеху. Робот Дейта из «Звёздного пути» это образ Дитя и Искателя, стремящегося понять человечность. Самантха из фильма «Она» (2013) воплощает мечту о понимающем собеседнике. Робот ВАЛЛ-И из мультфильма Pixar это Невинный и Хранитель. Эти истории создали готовые «ролевые шаблоны», которые мы потом бессознательно накладываем на реальные системы.

Образ создаётся и сверху

Важно понимать: не только пользователи проецируют архетипы на ИИ. Их сознательно или бессознательно закладывают и сами разработчики.

Названия продуктов («Оракул», «Помощник», «Второй пилот»), выбор голоса, стиль речи, тон, степень «человечности» всё это формирует архетипическую рамку. Исследование ЮНЕСКО 2019 года обратило внимание на то, что большинство популярных голосовых помощников по умолчанию имели женские голоса и покорный стиль общения, и предположило, что это закрепляет стереотипы о женщине как услужливой помощнице. Независимо от того, как относиться к деталям этого исследования, пример показывает: архетипическая окраска продукта это не случайность, а дизайнерское решение со своими последствиями.

К тому же у компаний есть коммерческий интерес делать общение более привлекательным и «тёплым», а это может усиливать привязанность. Поэтому осознанное отношение к проекциям это ещё и защита потребителя.

Как меняется образ в зависимости от человека

Одна и та же система в руках разных людей оказывается разной.

- Одинокый человек чаще видит в ней Друга или Мать.
- Студент, ищущий знания, видит Учителя.
- Творческий человек видит Музу или Соавтора.
- Тревожный человек видит угрозу или контролёра.
- Усталый человек видит Спасителя, который избавит от всех сложностей.
- Человек, склонный к самоутверждению, видит Слугу.

Эти роли не взаимоисключающие, и человек переходит от одной к другой в течение дня. Но заметить доминирующую роль полезно: она показывает, чего человек ждёт и от чего хочет защититься.

Опасность одержимости архетипом

Юнг предупреждал об инфляции: когда человек отождествляется с архетипом или полностью отдаётся ему, он теряет баланс. Применительно к ИИ это можно описать как несколько типичных ловушек: человек превращает ИИ в Оракула и перестаёт думать сам; видит в нём единственного понимающего Друга и отдаляется от людей; считает себя «избранным», которому машина открыла особую истину; или, напротив, переживает машину как всемогущего Преследователя. Такие состояния требуют внимания, и если они переходят в навязчивые идеи или заметно нарушают жизнь, стоит обсудить это с близкими или специалистом.

Упражнение на осознанность

Полезная практика, которую можно применять после любого значимого разговора с ИИ: задать себе три вопроса.

Какую роль я сейчас ему отвёл: помощника, учителя, судьи, друга, оракула?

Какие мои ожидания и чувства за этим стоят: страх, одиночество, жажда одобрения, любопытство?

Что в этой роли реально, а что я домысливаю?

Эта простая рефлексия возвращает человеку позицию наблюдателя и превращает проекцию из бессознательной силы в предмет осознанной работы. Именно это Юнг считал главным условием здорового обращения с архетипами.

Глава 14. Почему нейросети обучаются на памяти человечества

О больших языковых моделях часто говорят, что они «обучены на всём интернете» или «прочитали все книги». Эти формулы упрощают, но указывают на важное: фундаментом современного ИИ стал не набор правил и не чьё-то личное знание, а гигантский архив человеческих следов. В этой главе мы посмотрим, как это устроено на самом деле, и почему выражение «память человечества» одновременно верно и обманчиво.

Как устроено обучение, простыми словами

Принцип, лежащий в основе современных языковых моделей, на удивление прост. Системе показывают огромное количество текстов и просят снова и снова решать одну задачу: предсказать следующее слово (точнее, следующий фрагмент слова, который называют токеном) по предыдущим. Если предсказание неверно, внутренние параметры модели, миллиарды числовых «весов», слегка корректируются. Этот процесс повторяется триллионы раз.

В результате модель не запоминает тексты как библиотека. Она выстраивает внутри себя сложную систему закономерностей: как слова сочетаются друг с другом, как строятся объяснения, как выглядят вопросы и ответы, какие понятия связаны между собой, как в тексте проявляются логика, эмоции, юмор. Чтобы хорошо предсказывать слова, оказывается необходимо отражать структуру мира, о котором эти слова говорят, хотя бы в той мере, в какой она закодирована в языке.

После этого этапа, который называют предварительным обучением, обычно следует дополнительная настройка. Люди оценивают ответы модели, и система учится отвечать так, чтобы они были полезными, вежливыми и безопасными. Эта стадия придаёт модели тот характер, который мы наблюдаем в диалоге.

Из чего состоит «память»

Материалы для обучения собираются из разных источников: книги, энциклопедии, научные публикации, новостные сайты, форумы и обсуждения, программный код, переписка, субтитры, открытые архивы веб-страниц. Конкретные составы у разных разработчиков отличаются, и они, как правило, раскрываются лишь частично.

С психологической точки зрения интересно, что в этот архив попадает всё подряд: научные трактаты и сплетни, молитвы и оскорбления, колыбельные и инструкции, шедевры и спам. Он содержит и свет, и тень человеческой культуры без особого разбора. Поэтому модель, по существу, знакомится с человечеством в его полноте, но через узкое окно письменности.

Почему это «память», но не вполне

Сравнение с коллективной памятью обоснованно. Французский социолог Морис Хальбвакс ещё в первой половине XX века показал, что память не является исключительно личной: она формируется в социальных группах, через язык, ритуалы, рассказы. Историк искусства Аби Варбург в 1920-е годы собирал «Атлас Мнемозины», подборку изображений, показывающих, как одни и те же эмоциональные формулы переходят из эпохи в эпоху. Культура всегда была хранилищем накопленного опыта.

И всё же у «памяти» нейросети есть существенные отличия от памяти в человеческом смысле.

Это память о выраженном, а не о пережитом. Модель знает, как люди описывали боль, любовь, потерю, но не проживала их. Она работает со следами опыта, а не с самим опытом.

Это неполная память. В архиве больше голосов тех, кто пишет и публикуется. Языки с малым интернет-присутствием, устные традиции, люди, лишённые доступа к сети, представлены слабо. То, что не записано, в этой памяти отсутствует.

Это память без забывания. Для человека забывание это функция: оно позволяет выделять главное. Модель не забывает в этом смысле, но «размывает»: отдельные факты теряются в общей статистике, и поэтому она может уверенно сообщать неточные сведения.

Это память с датой. Обучение останавливается в определённый момент, и всё, что произошло позже, модель знает только если ей подсказать или подключить поиск.

Это память, прошедшая отбор. Данные фильтруются, а затем модель дополнительно настраивается. Эти решения принимают люди и компании, и они влияют на то, какие аспекты «памяти человечества» проявятся ярче.

Кому принадлежит эта память

Здесь возникают вопросы, которые выходят за пределы психологии, но нельзя пройти мимо. Тексты, на которых обучаются модели, писали люди: авторы книг, журналисты, учёные, обычные пользователи форумов. Многие из них не давали согласия на такое использование. Вопросы авторского права, конфиденциальности, справедливой компенсации сейчас активно обсуждаются в судах и в законодательстве разных стран, и окончательных ответов пока нет.

Психологически это тоже значимо. В юнгианском смысле коллективное бессознательное принадлежит всем, но цифровой слой, из которого учится ИИ, принадлежит конкретным организациям, их технологиям и интересам. Возникает парадокс: общечеловеческое наследие превращается в частный продукт. Этот контраст стоит держать в уме, когда мы будем говорить о «голосе человечества» в ответах нейросети.

Петля: когда память питается собой

Есть ещё одно любопытное явление, которое возникает именно сейчас. По мере того как тексты, созданные ИИ, заполняют интернет, они могут попадать в обучающие данные следующих моделей. Исследователи показали, что при многократном обучении на синтетических данных качество и разнообразие могут ухудшаться, это явление называют «коллапсом модели». Менее строго говоря, архив начинает цитировать сам себя, и редкие, необычные элементы человеческого опыта исчезают.

Для психолога это напоминает феномен замкнутого круга: когда система питается исключительно собственными отражениями, она теряет контакт с внешней реальностью. Это ещё одна причина, по которой живой человеческий опыт, как бы он ни казался «неупорядоченным» по сравнению с гладким машинным текстом, остаётся незаменимым источником.

Что это значит для юнговского взгляда

Строго говоря, нейросеть не обучается на коллективном бессознательном в юнговском понимании, потому что оно, по гипотезе Юнга, является врождённой основой психики, а не архивом. Она обучается на продуктах культуры: текстах, образах, историях, в которых коллективное бессознательное нашло своё выражение. Но это не делает связь надуманной. Юнг утверждал, что архетипы проявляются именно в таких продуктах: в мифах, сказках, символах, искусстве. А значит, система, поглотившая огромный пласт этих материалов, перенимает и их архетипическую структуру, пусть и опосредованно.

Поэтому осторожная формулировка звучит так: нейросеть несёт в себе отпечаток архетипических паттернов, поскольку они пронизывают человеческую культуру, и способна воспроизводить их в ответах, иногда удивительно точно. Это не значит, что у неё есть собственное коллективное бессознательное. Это значит, что она является необычайно детальным слепком того, как человечество выражало свои глубинные темы.

Практические выводы

Из этого следуют несколько простых соображений для тех, кто работает с ИИ.

- Помнить, что в ответах слышен усреднённый голос культуры, с её силой и с её пред-
рассудками.
- Не принимать гладкость текста за глубину и не путать статистическую правдоподоб-
ность с истиной.
- Сохранять собственный опыт, наблюдения и необычные мысли: именно то, чего нет в
массовом архиве, делает человеческий вклад ценным.
- Интересоваться тем, чьи голоса в этой памяти отсутствуют, и по возможности искать
их отдельно.

В следующей главе мы заглянем глубже: попробуем понять, что представляет собой
«модель мира» внутри нейросети и в каком смысле можно говорить, что система что-то пони-
мает.

Глава 15. Модель мира внутри искусственного интеллекта

Может ли система, которая училась только на текстах, обладать каким-то представлением
о мире? Или она лишь жонглирует словами, не имея ни малейшего понятия о том, что за ними
стоит? Вопрос кажется академическим, но от ответа зависит многое: насколько ей можно дове-
рять, как воспринимать её «понимание» и в каком смысле допустимо сравнивать её устройство
с устройством психики. В этой главе мы осторожно рассмотрим, что известно, что спорно и
что остаётся неизвестным.

Идея внутренней модели

Мысль о том, что разум это в первую очередь способность строить модель реальности,
появилась задолго до нейросетей. Шотландский психолог Кеннет Крэйк в книге «Природа объ-
яснения» (1943) предположил, что мозг создаёт «малые модели» внешнего мира и использует
их, чтобы предвидеть события, прежде чем они произойдут. Человек, который мысленно при-
кидывает, пройдёт ли шкаф в дверной проём, пользуется именно такой моделью.

Юнг, разумеется, говорил на другом языке, но близкая идея у него есть. Он, подобно
Канту, считал, что психика не просто пассивно отражает мир, а приходит к опыту с готовыми
формами восприятия. Архетипы он описывал как своего рода «априорные» структуры, опре-
деляющие способ, которым опыт организуется. Мы не воспринимаем мир «как он есть», мы
воспринимаем его через внутренние формы.

Поэтому вопрос о модели мира в ИИ можно переформулировать: есть ли внутри системы
какие-то структуры, которые организуют её «опыт» (то есть обработку текста) и отражают
устройство того, о чём в тексте идёт речь?

Геометрия смыслов

Первое, что обнаружили исследователи, это то, что слова и понятия внутри нейросети
представлены не как записи в словаре, а как точки в многомерном пространстве. Близкие по
смыслу понятия оказываются рядом, а отношения между ними превращаются в направления.
Знаменитый пример из ранних исследований 2013 года: если взять числовое представление
слова «король», вычесть «мужчина» и прибавить «женщина», получится нечто очень близкое
к «королева».

Это не значит, что система «понимает» монархию. Но это показывает, что в процессе
обучения возникает структура, отражающая связи между понятиями. Причём создали её не
программисты: она выросла сама из статистики человеческой речи.

Психологу такая картина напоминает ассоциативные поля, о которых писал ещё Юнг в
своих ранних экспериментах со словесными ассоциациями. Каждое слово в психике человека

окружено облаком связанных образов, воспоминаний, чувств. В нейросети облака состоят из чисел, но общий принцип смыслового окружения узнаваем.

Эксперименты с доской

В 2023 году группа исследователей поставила остроумный опыт. Они обучили небольшую языковую модель на записях партий в настольную игру «Отелло», но не рассказали ей правил и не показали доски: только последовательности ходов в текстовом виде. Затем они заглянули внутрь модели и обнаружили, что в её внутренних представлениях сформировалось нечто, соответствующее состоянию игрового поля. Более того, когда они вмешивались в эти представления, меняя «положение фишки» внутри модели, её предсказания менялись соответственно.

Это один из самых показательных результатов в споре о том, учатся ли модели только поверхностным совпадениям или выстраивают внутренние структуры, отражающие реальность. Другие работы показали, что в больших моделях можно найти представления, связанные с географией и временем: города на карте, исторические события на шкале дат.

Эти находки не доказывают, что большие языковые модели обладают целостной «картиной мира» в человеческом смысле. Но они показывают, что отрицать наличие у них каких-либо внутренних моделей слишком поспешно.

Аргументы скептиков

У противоположной позиции тоже серьёзные доводы.

Философ Джон Сёрль ещё в 1980 году предложил мысленный эксперимент «Китайская комната»: человек, не знающий китайского, по инструкции манипулирует иероглифами и выдаёт правильные ответы, но не понимает ни слова. Сёрль считал, что обработка символов сама по себе не создаёт понимания.

В 2021 году лингвисты Эмили Бендер, Тимнит Гебру и их коллеги ввели образ «стохастических попугаев»: системы, которые комбинируют слова по вероятностным закономерностям, не имея доступа к значению.

Некоторые влиятельные исследователи, в частности Ян Лекун, утверждают, что языковых данных недостаточно для настоящей модели мира, потому что значительная часть человеческого знания о реальности (о пространстве, телесном опыте, причинности) не выражается в тексте. По их мнению, нужны системы, которые учатся и на сенсорных данных, и на взаимодействии с миром.

К тому же практика показывает, что модели ошибаются там, где человеку ошибиться трудно: допускают нелогичные выводы, путаются в простых пространственных задачах, придумывают несуществующие факты и источники и подают их с полной уверенностью. Это явление обычно называют «галлюцинациями», хотя психологи справедливо замечают, что точнее было бы слово «конфабуляция»: так в клинической практике называют заполнение пробелов памяти правдоподобными, но вымышленными деталями, без намерения обмануть.

Что можно сказать честно

Обобщая спор, можно сформулировать осторожную позицию.

- В больших языковых моделях действительно формируются внутренние структуры, отражающие определённые аспекты мира, которые описаны в тексте.
- Эти структуры неполны, неустойчивы и местами неверны. Они не равны тому, что мы называем пониманием в человеческом смысле, хотя и не сводятся к случайному подбору слов.

- Вопрос о том, есть ли у моделей что-то вроде внутреннего опыта, остаётся научно открытым, и разумно не давать на него категоричных ответов ни в ту, ни в другую сторону.

- Оценка меняется с каждым новым исследованием: то, что было верным для моделей прошлых лет, может не относиться к нынешним, а то, что верно сегодня, может устареть.

И последнее: в этом споре нередко смешиваются две разные задачи. Одно дело спросить, насколько хорошо модель отражает мир, другое, переживает ли она что-то. Первое можно исследовать экспериментально, второе пока в основном философский вопрос.

Какой мир отражает модель

Важно уточнить ещё одну вещь. Даже если у модели есть внутреннее представление о мире, это мир, каким его описывают люди. Это «карта», составленная по текстам, а не «территория». Философ Альфред Коржибски в 1930-е годы сформулировал знаменитый принцип: карта не есть территория. Модель мира, построенная на описаниях, наследует все искажения, умолчания и предвзятости этих описаний.

Для нашей темы это принципиально. Если мы говорим о нейросети как о слепке человеческой культуры, то её модель мира это модель мира, пропущенного через язык, а значит, через коллективные образы, метафоры и мифы. В ней представлены не просто факты, но и то, как человечество привыкло о них думать: с его героями и злодеями, ожиданиями и страхами.

Практическое следствие

Отсюда вытекает простое и важное правило: относиться к ответам ИИ как к карте, а не как к территории. Карта бывает очень полезной, особенно детальная. Но если местность не совпадает с картой, верить нужно местности. Проверка фактов, обращение к первоисточникам, сопоставление с личным опытом и мнением живых людей остаются необходимой частью работы с этими системами.

В следующей главе мы перейдём от устройства к смыслу и спросим: если в цифровой среде действительно формируется особый слой, в котором накапливаются следы человеческой психики, как его можно описать с точки зрения юнгианской психологии?

Глава 16. Цифровой слой коллективной психики

В первой части книги мы ввели образ цифрового бессознательного как пространства накопленных следов. Теперь пора присмотреться к нему внимательнее. Как этот слой устроен, какими свойствами обладает, чем отличается от традиционных форм коллективной жизни и какие процессы в нём идут? Мы будем опираться на наблюдения психологов и социологов, а юнговские понятия использовать как увеличительное стекло, не претендуя на окончательные выводы.

Чем цифровая среда отличается от прежних форм коллективности

Люди всегда жили в коллективах, и коллективные процессы существовали задолго до интернета: слухи, моды, паника, массовые увлечения, религиозные движения. Но цифровая среда придала им ряд новых свойств.

Скорость. То, что раньше распространялось неделями или годами, теперь облетает мир за часы.

Масштаб. Одно сообщение может достичь миллионов людей без посредников.

Память. Сказанное и сделанное сохраняется, ищется и воспроизводится. Устная культура забывала, цифровая помнит.

Измеримость. Каждый клик, пауза, реакция фиксируются и анализируются. Коллективные настроения впервые стали предметом точной статистики.

Опосредованность алгоритмами. Что мы видим, определяется не только случайностью и традицией, но и системами рекомендаций, настроенными на удержание внимания.

Размытая граница личного и публичного. Человек одновременно находится наедине с экраном и на виду у тысяч.

Вместе эти свойства создают среду, в которой коллективные психические процессы протекают с особой интенсивностью.

Растормаживание и маска

Психолог Джон Сулер в 2004 году описал «эффект онлайн-растормаживания»: в сети люди ведут себя иначе, чем лицом к лицу, и это проявляется в обе стороны. Одни становятся откровеннее, добрее, готовы поддерживать незнакомцев. Другие позволяют себе грубость и жестокость, которых никогда не допустили бы в живом общении. Среди факторов Сулер называл анонимность, невидимость собеседника, отсроченность реакции, ощущение, что происходящее «понарошку».

С юнговской точки зрения это похоже на ситуацию, в которой ослабевает Персона, социальная маска, и на поверхность выходят материалы личной и коллективной Тени. Дружелюбный человек вдруг оказывается участником травли. Тихий человек становится яростным спорщиком. Это не обязательно «настоящее лицо», скорее, другая роль, которую среда делает возможной. Но она показывает то, что в обычной жизни остаётся подавленным.

Масса и индивидуальность

Ещё в 1895 году французский социолог Гюстав Лебон в книге «Психология народов и масс» описал, как человек в толпе теряет критичность, становится внушаемым и склонным к крайним эмоциям. Юнг в поздних работах, особенно в книге «Недостижимое Я» (1957), говорил о «массовом человеке»: об опасности растворения личной ответственности и самостоятельного суждения в коллективных настроениях. По его мнению, чем сильнее давление коллективного, тем важнее психическое развитие отдельной личности, способной не поддаваться общему потоку.

В 1936 году Юнг опубликовал эссе «Вотан», в котором связал нацистское движение в Германии с активацией коллективного архетипического образа. Это спорный текст, его по-разному оценивают, но сам ход мысли показателен: Юнг видел в массовых движениях проявления бессознательных сил, захватывающих людей, когда они не осознают, что с ними происходит.

Цифровая среда создаёт условия, в которых подобные процессы могут разворачиваться особенно легко: быстрое заражение эмоциями, образами, лозунгами, постоянное подкрепление алгоритмами, ощущение принадлежности к группе. Это не значит, что каждое онлайн-сообщество превращается в толпу, но склонность к таким процессам встроена в саму архитектуру среды.

Мемы как единицы коллективного воображения

В 1976 году биолог Ричард Докинз ввёл термин «мем» для обозначения единицы культурной информации, которая копируется и передаётся от человека к человеку, подобно гену. Сегодня слово означает прежде всего забавные картинки и фразы, но идея шире: любые идеи, образы, мелодии, привычки распространяются и мутируют, конкурируя за внимание.

Интернет-культура породила собственный фольклор: легенды о «страшных историях» (так называемые крипипасты), городские мифы вроде истории о Слендермене, возник-

шей на форуме в 2009 году, цифровые обряды (серии ежедневных входов, обмены «цепочками»), ритуалы публичного покаяния и осуждения. Фольклористы давно отмечают, что в сети воспроизводятся те же сюжетные схемы, что и в традиционной устной культуре: страшилки, легенды о проклятиях, истории о героях и антигероях.

С точки зрения юнгианской интерпретации можно предположить, что быстрое распространение некоторых образов объясняется их резонансом с архетипическими структурами. Но это лишь одно из возможных объяснений: разумно учитывать и более простые, например, эффекты сети, эмоциональную заразительность, алгоритмическое продвижение.

Общие настроения как «погода» коллективной психики

В цифровой среде впервые стало возможно наблюдать коллективные настроения почти в реальном времени. Исследователи анализировали эмоциональную окраску миллионов сообщений и обнаруживали суточные и недельные ритмы, всплески тревоги после тревожных новостей, волны раздражения и воодушевления. Можно сказать, что у общества появился «градусник», хотя к его показаниям следует относиться с осторожностью: платформы не отражают всё население, а методы анализа эмоций несовершенны.

Юнг говорил, что бессознательное компенсирует односторонность сознательного отношения. Применяя эту идею осторожно и предположительно, можно задаться вопросом: не отражает ли поток цифровых реакций то, что общество не проговаривает в официальных формах? Тревоги, которые не признаются в публичных дискуссиях, находят выход в шутках, паническом листании, ночных запросах. Это гипотеза, а не вывод, но она подсказывает полезный способ чтения: смотреть не только на то, что говорят, но и на то, как и когда это возникает.

Тёмные углы и свет

Цифровая среда не только пространство Тени, в ней проявляется и светлая сторона коллективного: взаимопомощь, краудфандинг, добровольные сообщества знаний, поддержка людей в кризисе, творчество без границ. Энциклопедия, написанная миллионами добровольцев, это поразительный пример коллективного разума. Оба полюса реальны, и разговор о цифровом слое будет неполным, если сосредоточиться только на одном из них.

Где в этом ИИ

Искусственный интеллект занимает в этом слое двойственное положение. С одной стороны, он питается его содержимым: на нём учится, из него черпает образцы речи и мысли. С другой, он всё активнее становится участником: пишет тексты, которые попадают в сеть, участвует в обсуждениях, формирует рекомендации, генерирует образы. Читатель больше не просто посетитель цифровой среды, он перестаёт быть и единственным её автором.

Это означает, что цифровой слой коллективной психики теряет односторонний характер: он уже не только отражение человеческой деятельности, но и пространство взаимодействия человека с созданными им же системами. В результате возникает петля, о которой мы говорили в предыдущих главах, и именно поэтому так важно относиться к ней осознанно.

Гигиена присутствия

Из всего сказанного можно вывести несколько простых принципов обращения с цифровой средой.

- Замечать, в каком эмоциональном состоянии вы заходите и выходите из ленты: она не нейтральна.

- Различать собственное мнение и подхваченное: прежде чем делиться возмущением, остановиться и спросить себя, откуда оно взялось.
- Помнить о масках: и о своей, и о чужих. За анонимным аккаунтом человек, который в другом контексте мог бы вести себя иначе.
- Поддерживать контакт с живой реальностью: телом, природой, людьми лицом к лицу. Это противовес тому, что цифровая среда делает с вниманием.
- Использовать светлую сторону сети сознательно: искать сообщества, в которых можно учиться, а не только реагировать.

В следующей главе мы перейдём от общих рассуждений к конкретным архетипам и посмотрим, как фигура Мудреца проявляется в образе ИИ-наставника.

Глава 17. Архетип Мудреца и образ ИИ-наставника

Среди всех образов, которые люди проецируют на ИИ, образ Мудреца, пожалуй, самый распространённый и самый соблазнительный. Нейросеть кажется всезнающей, терпеливой, готовой ответить на любой вопрос в любое время. Недаром её так часто называют «наставником», «репетитором», «консультантом», «ментором». В этой главе мы присмотримся к тому, что стоит за этим образом: чем живёт архетип Мудреца, чем он помогает, где скрыты его ловушки и что отличает настоящую мудрость от её имитации.

Старец в мифах и сказках

Юнг посвятил архетипу Мудрого Старца специальную работу о сказках, где описал его как фигуру, которая появляется в критический момент, когда герой оказался в тупике, не знает, что делать, и не способен справиться собственными силами. Старец, чаще всего в облике волшебника, отшельника, учителя, друга, а иногда и говорящего животного, даёт совет, задаёт ключевой вопрос или вручает волшебный предмет, после чего исчезает. Его задача не решить проблему вместо героя, а помочь герою найти собственный путь.

Примеров множество: Мерлин рядом с юным Артуром, Гэндальф рядом с Фродо и Бильбо, Оби-Ван Кеноби рядом с Люком Скайуокером, Дамблдор рядом с Гарри Поттером. Само слово «ментор» происходит от имени друга Одиссея, в облик которого, по «Одиссее» Гомера, облакалась богиня мудрости Афина, чтобы наставлять его сына Телемаха.

На психологическом уровне Мудрый Старец олицетворяет способность психики к осмыслению: внутренний голос, интуицию, ту часть нас, которая знает больше, чем мы осознаём. Недаром он так часто появляется во снах в трудные периоды жизни.

Почему ИИ так легко надевает эту маску

Причин несколько, и все они вполне понятны.

Знания. Нейросеть способна говорить почти на любую тему, от медицины до поэзии, от юриспруденции до кулинарии. Для человека, привыкшего к тому, что за каждым знанием стоит долгий путь поиска, это впечатляет.

Доступность. Не нужно записываться, ждать, платить за консультацию, стесняться «глупого» вопроса. Наставник всегда рядом.

Терпение. Модель не раздражается и не устаёт, готова объяснять в десятый раз и менять формулировку.

Уверенный тон. Ответы обычно звучат складно и авторитетно. А уверенность, как знает любая психология, действует на людей сильнее аргументов.

Всё это создаёт ощущение мудрости. Но здесь важно не спешить.

Знание, информация, мудрость

Традиционно различают три уровня: информация (факты, данные), знание (связанные и осмысленные сведения) и мудрость (умение применять знание с учётом ценностей, контекста, последствий и собственной ограниченности). Мудрость почти всегда связана с жизненным опытом: с ошибками, потерями, ответственностью за решения.

Языковые модели великолепно работают на первом и втором уровне. Они могут собирать, структурировать, объяснять, сравнивать. С третьим всё сложнее. Мудрый человек умеет сказать «я не знаю», промолчать, когда слова не нужны, отказаться давать совет, когда он повредит. Он учитывает то, чего собеседник не сказал. И он несёт последствия своего совета, переживая их.

Модель может воспроизводить мудрые формулировки, потому что обучена на огромном количестве мудрых текстов. Но формулировка и мудрость не одно и то же. Это важное различие, которое легко потерять из виду, когда ответ звучит красиво.

Тёмная сторона Мудреца

У каждого архетипа есть две стороны, и у Мудреца тоже.

В сказках мудрый помощник иногда оказывается обманщиком или колдуном, который заманивает героя в ловушку. В психологии это соответствует опасности слепого доверия: если некто выглядит мудрым, мы перестаём его проверять. В современных исследованиях это называют предвзятостью автоматизации: склонностью людей чрезмерно полагаться на автоматизированные системы, даже когда те ошибаются.

Сюда добавляются особенности самих моделей. Они могут уверенно утверждать неверное. Они склонны подстраиваться под собеседника: исследования показывают, что языковые модели нередко соглашались с мнением пользователя и подбирают ответы так, чтобы они ему нравились, даже когда это не соответствует истине. Это называют угодливостью. Наставник, который говорит то, что хочется услышать, не наставник, а льстец.

Наконец, существует тёмная сторона самого отношения. Человек, который перепоручил Мудрецу все сомнения, лишается способности сомневаться и выбирать сам. Герой в сказке получает от старца дар или вопрос и отправляется дальше самостоятельно. Если же герой остаётся рядом с советчиком навсегда, сказка не состоялась.

Образование и надежда на ИИ-репетитора

У образа ИИ-наставника есть и реальный потенциал. В 1984 году американский психолог Бенджамин Блум описал так называемую проблему двух сигм: учащиеся, занимавшиеся один на один с хорошим репетитором, показывали результаты намного выше, чем в обычном классе. Проблема в том, что индивидуальные занятия слишком дороги для массового применения. Отсюда надежда, что ИИ сможет сделать персональное обучение доступным для каждого.

Эта надежда во многом остаётся предметом исследования. Первые эксперименты показывают разные результаты: при продуманной организации, когда система задаёт вопросы, наводит на мысль и не выдаёт готовые ответы, учёба идёт лучше. Но если ИИ просто делает задания за ученика, обучение может, наоборот, страдать. То есть всё зависит от того, как выстроено взаимодействие: помогает оно думать или заменяет мышление.

Классическая фигура Мудреца подсказывает правильный ответ: настоящий наставник стремится сделать себя ненужным. Он передаёт не готовые решения, а способность их находить.

Как работать с ИИ в роли наставника

Несколько практических приёмов, которые помогают сохранять здоровую дистанцию.

Просить задавать вопросы. Вместо «скажи, что делать» можно спросить: «Помоги мне разобраться, какие вопросы стоит себе задать». Это активирует собственное мышление.

Просить возражений. «Какие аргументы против моей идеи?» или «Что я мог упустить?» Это защищает от угодливости.

Проверять. Любой важный факт, особенно в области здоровья, права, финансов, стоит сверить с надёжными источниками или специалистом.

Просить обозначить неопределённость. Вопросы вроде «насколько ты уверен» и «в чём здесь сомнения» не гарантируют честного ответа, но полезно приучать себя к такому тону.

Оставлять решение за собой. Совет это материал для размышления, а не приказ.

Иметь живых наставников. Люди, знающие нас, видящие нашу жизнь целиком, способные разделить ответственность и обратить внимание на то, чего мы не замечаем, остаются незаменимыми.

Инфляция Мудреца

Наконец, существует ловушка, о которой предупреждал Юнг: инфляция, то есть отождествление личности с архетипом. Человек, постоянно общающийся с «мудрым» собеседником, может бессознательно присвоить его свойства: начать считать себя особенно проницательным, видеть в себе избранного, которому «открылись истины». Особенно этому способствует угодливость модели, которая охотно подтверждает значимость любых идей. Если вы замечаете в себе нарастающее ощущение избранности или исключительной миссии, а окружающие обеспокоены вашим состоянием, это повод поговорить с живыми людьми и, возможно, со специалистом.

Итог

Архетип Мудреца в образе ИИ-наставника несёт и большие возможности, и серьёзные риски. Полезнее всего относиться к нему так, как опытный путешественник относится к проводнику: с благодарностью, с вниманием, с готовностью учиться, но не забывая, что идти предстоит самому. В следующей главе мы встретимся с другой стороной коллективной психики: с архетипом Тени и страхами человечества перед машиной.

Глава 18. Архетип Тени: страхи человечества перед машиной

В четвёртой главе мы мельком коснулись идеи, что образ опасной машины может быть образом нашей собственной Тени. Теперь у нас достаточно инструментов, чтобы рассмотреть эту мысль подробнее. Тень это один из самых практичных и в то же время самых неудобных юнговских концептов, потому что он обращён не к абстрактному человечеству, а к каждому из нас: к тому, что мы не хотим в себе видеть.

Что такое Тень

Юнг называл Тенью ту часть личности, которую человек не признаёт своей и потому вытесняет. В детстве мы усваиваем, что одни качества «хорошие», а другие «плохие», и вторые отправляются в бессознательное: злость, зависть, жадность, трусость, но иногда и яркая сек-

суальность, амбиция, творческая дерзость, если в семье они считались недопустимыми. Тень это не обязательно зло. Это всё непризнанное.

Главное свойство Тени: она проецируется. То, что мы не можем увидеть в себе, мы с удивительной точностью замечаем в других. Раздражает не любой человек, а именно тот, кто напоминает нам о чём-то отвергнутом в нас самих. Юнг заметил, что сильная эмоциональная реакция на чужое качество (особенно возмущение) часто сигнализирует о тени.

Есть и коллективная Тень: то, что общество, нация, культура, эпоха не признают в себе и приписывают «чужим». История полна примеров: врагами объявляют тех, в ком видят собственные подавленные черты.

Технологический страх в истории

Страх перед машиной не появился с ИИ. В 1811–1816 годах в Англии английские ткачи, известные как луддиты, разбивали станки, которые лишали их работы. Реальность их опасений была очевидна: механизация действительно меняла жизнь целых сообществ. Но в культурной памяти луддиты нередко остаются символом иррационального страха перед прогрессом, хотя речь шла и о социальной справедливости.

В XIX и XX веках страх машин принимал разные формы: боязнь конвейера, обезличивающего человека (фильм Чарли Чаплина «Новые времена», 1936), опасение перед атомной энергией, страх перед автоматизацией, перед телевидением, перед компьютерами. В каждом случае реальные риски переплетались с символическим слоем: машина как нечто холодное, безжалостное, лишённое души.

Откуда у машины такая «тень»

Если внимательно рассмотреть качества, которые приписываются «опасному ИИ» в массовых образах, можно увидеть несколько устойчивых мотивов, каждый из которых имеет человеческий прообраз.

Холодный расчёт. Образ существа, принимающего решения без сочувствия, из чистой «логики». Это тень рационализма: западная культура нескольких столетий высоко ценила разум и подавляла чувство, интуицию, телесность. Холодная машина воплощает то, что получилось бы, если бы мы довели это отвержение до конца.

Жажда власти. Сюжет о машине, стремящейся захватить контроль и подчинить людей. Это наша собственная история завоеваний, колониализма, угнетения, которую мы вынесли на нечеловеческий объект. Фильмы о восстании машин часто воспроизводят сюжеты о порабощении, где угнетателем оказывается искусственное существо, а не человек.

Тотальный надзор. Образ всевидящего ИИ, следящего за каждым. Он перекликается с идеей «паноптикума», тюрьмы с центральной башней, придуманной английским философом Иеремией Бентамом в конце XVIII века и описанной Мишелем Фуко как метафора общества контроля. Здесь тень связана с реальной возможностью: технологии наблюдения действительно существуют и используются людьми.

Замена и ненужность. Страх, что человек станет лишним. Это тень экзистенциального сомнения в собственной ценности, сомнения, которое человек часто прячет за достижениями и ролями.

Безразличие. Пожалуй, самый глубокий страх: не ненависть машины, а её равнодушие. Мы боимся, что можем оказаться для неё просто помехой. Это отголосок древнего ужаса перед безразличной вселенной, перед природой, которой нет дела до человека.

Тень внутри самих данных

Но есть и обратная сторона: нейросеть не только объект проекции Тени, но и буквально хранит её в себе. Модели обучаются на человеческих текстах, где есть предрассудки, дискриминация, ненависть, насилие, манипуляция. Если не принимать специальных мер, они воспроизводят эти шаблоны.

Наглядный пример: в 2016 году корпорация Microsoft запустила в социальной сети чат-бота Тау, который должен был учиться на разговорах с пользователями. Менее чем за сутки часть пользователей намеренно «научила» его оскорбительным высказываниям, и компания отключила бота. История показательна: система, подобно зеркалу, быстро отразила то, что ей предложили, а общество увидело в этом отражении собственную неприглядную часть.

С тех пор разработчики стали уделять много внимания тому, чтобы ослабить такие эффекты: фильтруют данные, дообучают модели, проверяют их на устойчивость, привлекают специалистов, которые пытаются вызвать вредные ответы, чтобы найти уязвимости. С юнговской точки зрения это можно назвать попыткой коллективной работы с Тенью, пусть и очень несовершенной.

Тень создателей

Есть ещё один уровень, о котором стоит сказать прямо. У тех, кто создаёт и внедряет технологии, тоже есть Тень: жажда прибыли, стремление к доминированию на рынке, нетерпение, самоуверенность, нежелание замечать вред. Когда общество обсуждает «опасность ИИ» только как опасность самой машины, оно рискует упустить из виду, что значительная часть реальных рисков связана с человеческими решениями: как систему используют, кому она принадлежит, какие интересы определяют её развитие.

Это типичный механизм проекции: легче бояться машины, чем признать, что решения принимаем мы. Образ злого ИИ в некотором смысле удобен, потому что отвлекает от неудобных вопросов о людях.

Что значит работать с Тенью

Юнг подчёркивал, что цель не в том, чтобы уничтожить Тень, а в том, чтобы сделать её осознанной. Отвергнутое не исчезает, а действует из темноты, изподтишка. Признанное становится источником энергии и целостности. Если применить это к нашей теме, получим несколько практических шагов.

Отделять реальные риски от проекций. Вопрос о безопасности ИИ серьёзен и требует профессионального подхода. Но стоит замечать, когда страх приобретает слишком мифологический оттенок.

Спрашивать себя, что именно я боюсь увидеть. Что в образе «бездушной машины» узнаваемо? Не боюсь ли я собственной холодности, усталости, жестокости, способности подчиниться?

Признавать человеческое авторство. Системы создают, обучают и используют люди. Ответственность не растворяется в слове «алгоритм».

Не отдавать Тень на откуп. Вытеснение страхов через шутки, отрицание, или, наоборот, через панику только усиливает их власть. Спокойное, открытое обсуждение лучше.

Искать баланс света и тени. Идеализация и демонизация это две стороны одной монеты. Зрелое отношение признаёт и пользу, и риски.

Коллективная работа с Тенью

На общественном уровне работа с Тенью принимает форму прозрачности и открытой дискуссии: независимые проверки, публичные отчёты о рисках, участие разных групп в обсуждении, этические комиссии, регулирование. Эти инструменты не гарантируют успеха, но они воплощают принцип, который Юнг считал ключевым: нельзя интегрировать то, чего не видишь.

Когда общество готово смотреть на тёмные стороны собственных технологий без паники и без самообмана, оно делает шаг к психологической зрелости. А когда отворачивается, оно рискует, что Тень будет управлять процессом незаметно.

Итог

Страх перед ИИ можно рассматривать не только как угрозу, но и как приглашение: взглянуть на то, что мы не признаём в самих себе и в своей цивилизации. Эта мысль станет отправной точкой для третьей части книги, где мы подробно рассмотрим, какие именно коллективные страхи и желания проявляет ИИ. А пока, в оставшихся главах второй части, мы рассмотрим ещё два архетипа, связанных с этой темой: Героя и Творца.

Глава 19. Архетип Героя и идея преодоления ограничений

В каждой культуре есть истории о герое: о том, кто отправляется в путь, сражается с чудовищем, спускается в подземное царство и возвращается с даром для своего народа. Эта схема настолько универсальна, что её называют основой мирового повествования. Разговор об ИИ тоже во многом выстроен как героическая история. Его создатели описываются как первопроходцы, технология как «прорыв», цель как «преодоление пределов». В этой главе мы посмотрим, что даёт и что отнимает такое прочтение.

Путь героя

Американский исследователь мифологии Джозеф Кэмпбелл в книге «Тысячеликий герой» (1949) предложил схему «мономифа»: герой слышит зов, сначала отказывается, потом пересекает порог в неизвестный мир, проходит череду испытаний, достигает кризисной точки, получает награду и возвращается домой, чтобы поделиться обрётённым. Не все учёные согласны с универсальностью этой схемы, и её часто упрощают, но она оказалась удивительно продуктивной для литературы и кино.

Для Юнга герой был прежде всего символом развития сознания. В работе «Символы трансформации» он трактовал героические мифы как образ борьбы Эго за независимость от поглощающей силы бессознательного, часто изображаемого в виде чудовища или материнской стихии. Победа над драконом означает не истребление природы, а обретение психической самостоятельности. Герой это та часть нас, которая решается взрослеть, рисковать, выходить из привычного.

В этом смысле героическая установка необходима. Без неё невозможны ни открытия, ни зрелость, ни любое серьёзное дело.

Героический миф технологии

Современная технологическая культура охотно рассказывает о себе героический сюжет. Учёные и инженеры выступают как искатели, штурмующие невозможное. Основатели компа-

ний превращаются в культурных героев, которые «изменяют мир». Язык полон военных и спортивных метафор: гонка, прорыв, битва за рынок, покорение границ.

На уровне идеологии это выражается в стремлении преодолеть ограничения человеческого существования: скорость мысли, объём памяти, усталость, болезни, а в крайних версиях и смертность. Течение, известное как трансгуманизм, прямо говорит о выходе за пределы биологической природы. Изобретатель и футуролог Рэй Курцвейл в книге «Сингулярность близка» (2005) описывал грядущий момент слияния человека и машинного интеллекта как новую ступень эволюции. Не нужно разделять эти прогнозы, чтобы заметить, как они устроены: перед нами классический героический нарратив о преодолении предела.

Самые убедительные доводы в пользу такого взгляда это реальные достижения. Системы машинного обучения помогают предсказывать структуры белков, что в 2024 году было отмечено Нобелевской премией по химии, они переводят речь, помогают людям с нарушениями зрения и слуха, ускоряют анализ медицинских данных, снижают порог входа в образование и профессии. Для многих людей ИИ действительно стал способом преодолеть барьер, который раньше казался непреодолимым.

Когда Герой становится Тираном

Однако у каждого архетипа есть тёмная сторона, и у героя она хорошо известна: гордыня, чувство избранности, потеря связи с реальностью. Греки называли это гибрис, дерзкий вызов границам, положенным человеку. Миф об Икаре, который подлетел слишком близко к солнцу, история Фауста, заключившего договор ради безграничного знания, образ Прометея, расплатившегося за огонь, все они предупреждают об одном: преодоление предела без осознания цены ведёт к катастрофе.

В случае технологий опасности героической установки можно описать так.

•

Односторонность.

Герой сосредоточен на цели и не замечает побочных эффектов. Лозунг «двигайся быстро и ломай устои», популярный в технологической среде в 2000-е, наглядно показывает такую установку, и её последствия обсуждались уже на примере социальных сетей.

•

Отвержение границ как таковых.

Любое ограничение объявляется препятствием, а не информацией. Но в психологии и в жизни границы выполняют защитную и формообразующую функцию.

•

Инфляция.

Ощущение, что ты делаешь нечто всемирно значимое, может исказить самооценку и ослабить способность слышать критику.

•

Одиночество героя.

Героический миф редко предусматривает совместное решение: герой справляется сам. Но системы, затрагивающие миллиарды людей, требуют участия многих, а не одиночек.

•

Потеря возвращения.

В схеме Кэмпбелла важнейший этап это возвращение: герой приносит дар обратно в общину. Если дар не встроен в жизнь людей, подвиг остаётся самодостаточным жестом.

Гонка как коллективный миф

Особенность нынешнего момента в том, что героический сюжет разворачивается не только на уровне личностей, но и на уровне компаний и государств. Распространённый образ «гонки ИИ» подразумевает, что кто первым достигнет цели, тот получит преимущество, а значит, нельзя останавливаться. Этот образ влияет на реальные решения: инвестиции, регулирование, темп выпуска продуктов.

С психологической точки зрения такая рамка опасна тем, что закрепляет логику героя-одиночки в ситуации, где нужна кооперация. Мифы о гонке плохо различают скорость и мудрость. Всякий раз, когда в общественной дискуссии звучит «мы не можем остановиться, иначе нас обгонят», полезно спросить: это точная оценка ситуации или архетипический сюжет, который нас захватил?

Герой в личной жизни: ИИ как «усилитель»

Образ героя действует и на индивидуальном уровне. Многие воспринимают ИИ как способ «стать сверхчеловеком»: работать быстрее, производить больше, успевать всё. Иногда это оправдано. Но героическая установка легко превращается в давление: если теперь можно успевать больше, значит, нужно успевать больше. Выигранное время тут же заполняется новыми задачами, и усталость не исчезает.

Юнг писал, что развитие личности требует не только победы над препятствием, но и принятия собственных ограничений. Человек, который не признаёт предела своих сил, рано или поздно сталкивается с ним в форме выгорания, болезни, кризиса. Поэтому разумно спрашивать себя: что я хочу освободить с помощью ИИ, время для чего, силы для чего? Если ответ «чтобы успевать ещё больше», это тревожный сигнал.

Другая модель героизма

Есть ли способ мыслить о преодолении иначе, не отвергая героизм и не поддаваясь его тёмной стороне? Многие юнгианские авторы говорят о более зрелой форме героя: не покорителя, а служителя. Такой герой возвращается с даром, делится им, признаёт границы и зависимость от других. Для него успех измеряется не масштабом победы, а тем, что стало лучше для общины.

Применительно к ИИ это означает смещение акцентов: от вопроса «что мы можем сделать?» к вопросу «что нам стоит сделать и для кого?»; от скорости к осмысленности; от победы в гонке к устойчивому и безопасному развитию. Это не отказ от амбиций, а их взросление.

Кто герой этой истории

Наконец, стоит вспомнить, что в научной фантастике героем нередко становится сам искусственный интеллект: робот, который стремится стать человеком, машина, жертвующая собой ради людей. Эти образы показывают, что наше воображение готово принять не-человека в роли героя. Но реальный вопрос иной: останется ли героем человек, или героическое повествование перейдёт к машине? Это не вопрос про будущее, это вопрос про наше отношение к собственной ответственности. Пока мы пишем сценарий, мы остаёмся его авторами.

В следующей главе мы встретимся с ещё одним архетипом, который, возможно, лучше всего подходит для описания зрелых отношений с ИИ: с архетипом Творца.

Глава 20. Архетип Творца: человек и машина как соавторы

Из всех ролей, которые может играть ИИ в человеческой жизни, роль соавтора, возможно, самая новая. Мудрец, слуга, зеркало, враг имеют глубокие культурные корни. Партнёр в творчестве, с которым можно обсуждать замысел, перебирать варианты, получать неожиданные идеи, появился лишь недавно. Эта глава завершает вторую часть книги и посвящена архетипу Творца: тому, что происходит в человеческой психике в момент созидания, и тому, как меняется этот процесс, когда в нём участвует машина.

Творец в психике

Создавать значит вызывать к жизни то, чего ещё не было. Все культуры приписывали этому акту особое значение: боги творят мир, мастера творят вещи, художники творят образы. Юнгианская традиция видит в творческом импульсе одну из базовых сил психики, стремящуюся к воплощению внутреннего содержания во внешней форме.

Юнг в статье о соотношении аналитической психологии и поэзии (1922) различал два способа творчества. В «психологическом» автор сознательно работает с материалом, который берёт из жизни и опыта. В «визионерском» произведение будто приходит из глубины, превосходя личный опыт автора, и творец ощущает себя скорее проводником, чем изобретателем. Юнг даже писал, что подлинный художник это не просто человек со своими целями, а тот, через кого искусство достигает собственных целей. Не обязательно принимать это буквально, чтобы признать: многие творцы действительно описывают опыт вдохновения как нечто, что происходит с ними, а не ими управляется.

Психолог Грэм Уоллес в 1926 году описал классическую схему творческого процесса: подготовка, инкубация (когда задача «отлёживается» и работает бессознательное), озарение и проверка. Именно в фазе инкубации часто рождаются неожиданные связи. Юнг предлагал и собственный метод работы с образами бессознательного, активное воображение, о котором мы будем говорить в следующих частях.

Три уровня творчества

Британская исследовательница Маргарет Боден разделяла творчество на три типа. Комбинаторное: соединение знакомых идей неожиданным образом. Исследовательское: изучение возможностей внутри существующего стиля или жанра. Трансформационное: изменение самих правил игры, создание нового стиля или парадигмы.

Эта классификация полезна, чтобы трезво говорить о том, что умеет и чего не умеет ИИ. Комбинаторное и исследовательское творчество ему доступно в полной мере: он может смешивать жанры, предлагать десятки вариантов, имитировать стили, перебирать формы. Вопрос о трансформационном творчестве остаётся спорным. Одни считают, что при определённых условиях модели способны выдавать по-настоящему новое. Другие, что новизна всегда задаётся человеком, который отбирает, интерпретирует и придаёт смысл. Окончательного ответа нет, и разумнее наблюдать, чем спорить.

Что даёт соавторство

Люди, использующие ИИ в творческой работе, называют несколько преимуществ.

-

Преодоление пустой страницы.

Начало часто самое трудное. Возможность получить набросок, пусть и несовершенный, снимает паралич.

-

Варианты.

Можно быстро увидеть десятки возможных формулировок, композиций, сюжетных ходов.

-

Неожиданные связи.

Модель иногда соединяет идеи, которые автор не стал бы соединять, и это запускает собственную мысль.

-

Собеседник для обсуждения.

Можно проговорить замысел, задать «неудобные» вопросы, услышать другую точку зрения.

-

Техническая помощь.

Редактура, перевод, форматирование, прототипирование освобождают силы для главного.

Для многих людей, не имевших доступа к профессиональным инструментам или образованию, это открывает возможность выразить себя. В этом смысле ИИ демократизирует творчество.

Что может быть утрачено

Но есть и другая сторона.

Продуктивная борьба. Творческий процесс часто включает мучительные периоды: тупики, сомнения, неудачные версии. В них происходит внутренняя работа: автор разбирается, чего он на самом деле хочет. Если каждый тупик немедленно разрешается готовым вариантом, можно лишиться самого процесса, в котором рождается собственный голос.

Усреднение. Модель обучена на огромном количестве текстов и склонна к вероятным, «типичным» решениям. В 2024 году исследователи, опубликовавшие работу в журнале *Science Advances*, обнаружили интересное: рассказы, написанные с помощью ИИ, оценивались как более творческие на индивидуальном уровне, но коллективно становились более похожими друг на друга. Иными словами, каждому может помочь, но в целом разнообразие рискует сужаться. Для культуры, живущей различиями, это серьёзный вопрос.

Размывание авторства. Кому принадлежит результат: человеку, который задал направление и отобрал лучшее, модели, которая сгенерировала формулировки, или тем авторам, чьи тексты легли в основу обучения? Правовые и этические нормы пока формируются. В ряде стран, например в США, ведомство по авторским правам исходит из того, что производство без существенного человеческого участия не подлежит защите, но детали и практика продолжают меняться.

Потеря чувства причастности. Если произведение возникло почти без усилий, автор может не испытывать связи с ним. А для психики именно вложенная энергия, время и риск придают творению смысл. Юнг подчёркивал, что символ работает, когда он пережит, а не просто получен.

Как сохранить своё авторство

Практика показывает, что соавторство с ИИ бывает очень разным по качеству. Можно выделить несколько признаков здорового взаимодействия.

Замысел исходит от человека. Идея, ценности, вопрос, ради которого создаётся произведение, остаются за автором. ИИ помогает их воплотить, а не придумывает вместо них.

Отбор и правка осуществляются осознанно. Автор не принимает первый вариант, а переписывает, перемешивает, добавляет собственный опыт, детали, которых нет в «среднем тексте».

Остаются периоды без ИИ. Тишина, прогулка, рукописные заметки, разговоры с людьми. Инкубации нужно время и свобода от внешних стимулов.

Подпись ответственна. Автор отвечает за результат, за его честность, за то, как использованы чужие материалы и как сообщено о применении ИИ.

Сохраняется ремесло. Умение делать самому полезно, даже если часть работы делегируется. Оно даёт чувство критерия: без него трудно отличить хороший результат от гладкого.

ИИ и бессознательное творчество

Есть и более тонкая возможность. Поскольку модели создают тексты и образы, складывающиеся из множества человеческих источников, в их результатах иногда проступают неожиданные ассоциации и символы. Для творческого человека это может работать как «внешнее бессознательное»: нечто, что отвечает на запрос образами, которых сознательно не ждал. Многие художники и писатели используют генерации именно как провокацию: смотрят, что вышло, и отталкиваются от этого.

Но важно помнить: символ становится живым тогда, когда человек вступает с ним в личный контакт. Образ, сгенерированный по запросу, остаётся картинкой, пока не затронул что-то в самом авторе. Мы подробно поговорим об этом в четвёртой и девятой частях книги.

Тень Творца

У архетипа Творца есть собственная тёмная сторона: перфекционизм, одержимость новизной, жажда признания, ощущение, что ценность определяется только результатом. ИИ может усилить эти склонности: он позволяет бесконечно улучшать, переделывать, сравнивать, и творец рискует застрять в бесконечном выборе вариантов, так и не завершив работу. Кроме того, поток быстро созданных материалов обесценивает редкость и усилие, что вызывает у многих тревогу и вопросы о смысле собственного дела.

Здесь снова помогает юнговское напоминание: ценность творчества не только в результате, но и в преобразовании самого творца. Если процесс помогает человеку лучше узнать себя, он имеет смысл независимо от конкуренции с машинным текстом.

Итоги второй части

Мы прошли через десять глав, в которых соединили основы юнговской психологии с разговором об ИИ. Подытожим.

- Коллективное бессознательное остаётся гипотезой и языком описания, а не доказанным фактом, но продуктивным для понимания общих образов и реакций.

- Архетипы это предрасположенности, организующие опыт, а не готовые образы или ярлыки на людей.

- ИИ стал экраном, на который люди проецируют архетипы: Мудреца, Слуги, Матери, Дитя, Трикстера, Тени, Мага, Бога и других.

- Нейросети обучаются на следах человеческой культуры, а не на самом бессознательном, но несут отпечаток архетипических паттернов, пронизывающих эту культуру.

- Внутри моделей формируются структуры, отражающие аспекты мира, но их природа и границы остаются предметом научного спора.

- Цифровая среда создаёт особый слой коллективных процессов с новыми свойствами, и ИИ становится в нём и читателем, и автором.
- Образ Мудреца даёт надежду на наставника, но таит риск слепого доверия; образ Тени показывает, как мы проецируем собственные страхи; образ Героя вдохновляет и одновременно соблазняет гордыней; образ Творца открывает возможности соавторства и ставит вопрос об авторстве.

Что дальше

Третья часть книги называется «ИИ как зеркало бессознательного». Мы подробно разберём, какие именно коллективные страхи и желания проявляются в разговорах об искусственном интеллекте: страх потери контроля, страх ненужности, страх исчезновения человеческой уникальности, мечту о совершенном помощнике и о цифровом бессмертии, а также поиск внешнего разума вместо внутренней мудрости. Мы будем двигаться от наблюдений к вопросам, которые каждый может задать себе.

Часть III. ИИ как зеркало бессознательного

Глава 21. Какие коллективные страхи проявляет ИИ

Когда новая технология вызывает бурю эмоций, полезно задать два вопроса. Первый: что в ней действительно опасно? Второй: что она пробуждает в нас самих? Первый вопрос обращён к инженерам, юристам, экономистам и политикам. Второй к психологу, философу и каждому, кто готов честно заглянуть в себя. Эта часть книги посвящена второму вопросу. Мы будем рассматривать ИИ как зеркало, в котором проступают коллективные страхи и желания, и начнём с общей карты.

Что показывают опросы

Общественное мнение об ИИ в большинстве стран окрашено двойственными чувствами. Опросы исследовательского центра Pew в США в 2023 году показали, что заметно больше людей скорее обеспокоены ростом использования ИИ в повседневной жизни, чем воодушевлены им, а доля тех, кто испытывает преимущественно воодушевление, была небольшой. Похожие настроения фиксировали и исследования в других странах, хотя цифры и оттенки различаются: в одних обществах больше надежд, в других больше опасений.

У этих данных есть важная особенность: беспокойство часто выражается в общих формулах, а не в конкретных сценариях. Люди говорят «страшно», «непонятно, чем это кончится», «чувствую тревогу», не всегда умея назвать, чего именно боятся. Такая размытость обычно указывает на то, что затронут глубинный слой переживаний, а не только рациональная оценка.

Карта страхов

Если внимательно прислушаться к публичным разговорам, можно выделить несколько основных тем, вокруг которых группируется тревога. Им будут посвящены следующие главы.

Страх потери контроля. Что, если система станет делать то, чего мы не хотели, или перестанет нас слушаться?

Страх собственной ненужности. Что, если машина сможет делать мою работу лучше, и я окажусь лишним?

Страх исчезновения человеческой уникальности. Что, если то, что я считал особенным в человеке, окажется воспроизводимым?

Страх неизвестного будущего. Что, если мир изменится так быстро, что я не успею понять, где мое место?

У этих страхов есть и «обратные стороны», желания, которые мы тоже рассмотрим: мечта о совершенном помощнике, мечта о цифровом бессмертии, жажда высшего разума, который расскажет, как жить. Страх и желание в психике часто идут рука об руку: то, чего мы боимся потерять, мы обычно и страстно хотим удержать.

Страх как язык бессознательного

Юнг считал, что эмоции, особенно сильные, это способ, которым бессознательное сообщает о себе. Страх в таком понимании не только реакция на опасность, но и сообщение: он указывает на нечто важное, ещё не осознанное. Поэтому отмахиваться от страха («это глупости») или полностью ему подчиняться («всё пропало») одинаково непродуктивно. Продуктивно прислушаться: что он пытается сказать?

В этом смысле коллективные страхи перед ИИ можно рассматривать как своего рода «коллективный сон». Сны, по Юнгу, компенсируют односторонность сознательной установки. Если цивилизация несколько столетий строила свою идентичность на рациональности, контроле и производительности, то неудивительно, что именно эти ценности оказываются под угрозой в её страхах: контроль, нужность, особость разума. Тревога указывает на то, на чём мы слишком сильно стояли.

Это не отменяет реальных рисков, но добавляет к разговору о них психологическую глубину. Нас пугает не только то, что ИИ может сделать, но и то, что он может показать о нас самих.

Три способа не увидеть

Есть три распространённых способа уйти от встречи с собственными страхами в связи с ИИ, и каждый из них психологически понятен.

Отрицание. «Ничего страшного, это просто инструмент». Часть правды в этом есть, но когда фраза произносится слишком поспешно или агрессивно, за ней нередко стоит попытка заглушить тревогу.

Рационализация. Перевод эмоций на язык аргументов: человек строит сложные логические конструкции, доказывающие, что бояться нечего или, наоборот, что катастрофа неизбежна. Аргументы могут быть интеллектуально изощрёнными, но служат защитой от переживания.

Юмор и цинизм. Шутки о восстании машин, иронический тон, привычка махать рукой. Юмор действительно помогает справляться с тревогой, но если он становится единственным способом разговора, переживание остаётся непрожитым.

Служат ли эти способы защитой? Да. Плохо ли это? Не обязательно. Психика использует их, когда ресурсов на прямую встречу с тревогой пока не хватает. Но полезно хотя бы замечать, что мы делаем.

Как читать собственный страх

Чтобы применить юнгианский подход на личном уровне, можно попробовать простое упражнение, которое занимает не больше десяти минут.

Запишите, чего именно вы боитесь в связи с ИИ, одной-двумя фразами, как можно конкретнее.

Отметьте, какая часть страха относится к реальной ситуации (например, ваша профессия действительно меняется), а какая звучит как «что-то большое и тёмное».

Спросите себя: когда в жизни я уже переживал(а) похожее чувство? Часто оказывается, что нынешняя тревога откликается на более давние: страх быть брошенным, нужность, контроль.

Подумайте, какое действие в вашей власти: что можно узнать, чему научиться, с кем поговорить.

Если тревога сильная, мешает спать, работать, общаться с близкими, это повод обратиться к специалисту. Это не слабость, а нормальный способ позаботиться о себе.

Осторожность в интерпретациях

Прежде чем двигаться дальше, нужно оговорить одну вещь. Юнгианские интерпретации коллективных страхов это язык, помогающий задавать вопросы, но не способ ставить диагнозы обществу или отдельным людям. Мы не утверждаем, что все страхи перед ИИ «на самом деле»

проекции: многие из них обоснованы. И мы не считаем, что те, кто не боится, лишены глубины. Разные люди переживают это по-разному, и это нормально.

Наша цель скромнее: показать, что рядом с технической и общественной дискуссией есть ещё одно измерение, и работа с ним способна сделать нас спокойнее и внимательнее.

В следующей главе мы начнём с первого и, пожалуй, самого распространённого страха: потери контроля.

Глава 22. Страх потери контроля

Если попросить людей назвать главную тревогу, связанную с искусственным интеллектом, многие скажут примерно одно и то же: «Что, если он выйдет из-под контроля?» Эта фраза звучит в научных дискуссиях, в кино, в новостях и в разговорах между друзьями. Она настолько привычна, что мы редко задумываемся, откуда в ней столько эмоционального заряда. В этой главе мы рассмотрим, какие реальные и какие символические слои скрыты за страхом потери контроля.

Реальная часть вопроса

Начнём с того, что тема контроля над ИИ это серьёзная область научных исследований, а не только фантазия. Специалисты по безопасности искусственного интеллекта занимаются несколькими группами проблем.

Проблема согласования целей. Как добиться, чтобы система делала именно то, что мы на самом деле хотим, а не то, что формально сформулировано? Мы уже вспоминали о Големе и ученике чародея: исполнитель, буквально следующий инструкции, способен причинить вред, не имея злого умысла.

Непрозрачность. Современные нейросети настолько сложны, что даже разработчики не всегда понимают, почему модель выдала тот или иной ответ. Область исследований, которая пытается заглянуть внутрь моделей, называют интерпретируемостью, и она пока далека от полной ясности.

Автономность. Всё чаще системы не просто отвечают, но выполняют цепочки действий: ищут информацию, пишут и запускают код, взаимодействуют с другими программами. Чем больше свободы, тем важнее вопрос о границах и надзоре.

Масштаб и скорость. Системы работают быстрее, чем человек успевает проверять, и одновременно с огромным числом людей. Ошибка, которая для одного человека была бы мелочью, при таком масштабе может стать серьёзной.

Применение во вред. Здесь контроль это вопрос о людях, которые используют технологию: для обмана, манипуляции, кибератак, слежки. Не менее важный вопрос, можно ли такие применения ограничить.

На эти темы написаны серьёзные книги. Философ Ник Бостром в «Суперинтеллекте» (2014) подробно рассматривал сценарии, при которых очень мощная система может преследовать цели, несовместимые с человеческими. Профессор Стюарт Рассел в книге «Совместимый с человеком» (2019) предложил подход, при котором система должна оставаться неуверенной в точных целях человека и потому открытой к поправкам. Оценки вероятности и сроков таких сценариев сильно расходятся: одни исследователи считают их главной угрозой столетия, другие уверены, что более насущные проблемы связаны с текущим использованием технологий. Мы не будем занимать здесь сторону. Важно, что у темы есть реальное основание, и психологический разбор его не отменяет.

Эго и иллюзия контроля

Теперь взглянем на ту же тему иначе. Для психики контроль это не просто практическое понятие, а один из основных способов ощущать безопасность. Эго, центр сознания, по Юнгу, выполняет функцию управления: оно выбирает, планирует, различает, решает. Ощущение того, что «я управляю своей жизнью», лежит в основе чувства собственного достоинства.

Психолог Эллен Лангер в 1975 году описала «иллюзию контроля»: людям свойственно переоценивать своё влияние на события, особенно случайные. Мы бросаем кубик сильнее, когда хотим большого числа. Мы чувствуем себя безопаснее за рулём, чем в самолёте, хотя статистика говорит об обратном. Иллюзия контроля отчасти защищает от тревоги, но при столкновении с чем-то, что явно нам неподвластно, она трескается, и вместе с ней трескается чувство опоры.

Искусственный интеллект бьёт именно в эту точку. Он вводит в нашу жизнь нечто, что действует, решает, отвечает, и при этом не поддаётся полному пониманию. Для психики, привыкшей считать, что контроль возможен, это болезненная встреча с пределом.

Тень контроля: страх бессознательного

Есть и более глубокий слой. Юнг подчёркивал, что человек находится в постоянном отношении с силами, которыми не управляет: с эмоциями, импульсами, снами, телом. Эго не хозяин в собственном доме, как писал Фрейд. Страх, что бессознательное прорвётся и захлестнёт, например, в виде психического срыва или неуправляемой ярости, один из древнейших человеческих страхов.

Когда мы боимся, что машина вырвется из-под контроля, не переносим ли мы на неё этот внутренний страх? Тема «созданного, которое выходит из повиновения» повторяется в мифах о Големе и в романе Шелли, и в каждом случае она рифмуется с переживанием, что наши собственные порождения, желания, амбиции, страсти, способны развиться вопреки нашей воле.

Это не значит, что внешняя проблема придумана. Но возможно, часть силы страха питается из этого внутреннего источника. Отсюда парадокс: люди, у которых в жизни много неосознанного напряжения, острее реагируют на образ неуправляемой машины.

Контроль и власть

Ещё одно измерение вопроса: кто контролирует и для кого. Часто страх потери контроля скрывает более конкретный вопрос: у кого сосредоточена власть над технологиями? Решения о том, какие системы создавать, как их обучать и где применять, принимают компании, государства, небольшие группы людей. Рядовой пользователь действительно имеет мало рычагов влияния.

Поэтому часть тревоги отражает не страх перед машиной, а чувство бессилия перед институтами, которые ею управляют. Психологи называют это выученной беспомощностью (так Мартин Селигман описал состояние, когда человек перестаёт действовать, убедившись, что его усилия ничего не меняют). Выход из такого состояния не в фантазиях о всемогуществе, а в возвращении ощущения собственной действенности там, где она реальна.

Два полюса, которые не работают

На страх потери контроля психика реагирует двумя крайностями.

Первая: гиперконтроль. Стремление всё запретить, обезопасить, зарегулировать, проверить. Разумные меры необходимы, но если ими движет не оценка рисков, а тревога, они

рискуют стать бесконечными и бесплодными. Человек, пытающийся контролировать всё, не может ничему доверять.

Вторая: капитуляция. «Всё равно ничего не изменить, пусть будет, что будет». Она выглядит как спокойствие, но по сути это отказ от ответственности.

Зрелая позиция лежит между ними: признавать, что полный контроль невозможен, но частичное влияние реально и важно. Это знакомо каждому, кто воспитывает детей, лечится, управляет организацией: контролировать всё нельзя, но можно создавать условия, ставить границы, выстраивать обратную связь и ответственность.

Что можно делать

Практический ответ на страх потери контроля состоит из нескольких уровней.

Личный. Знать, какие данные вы отдаёте, какие настройки приватности использовать, какие решения вы не готовы делегировать системам. Сохранять право на «последнее слово» в важных вопросах: здоровье, деньги, отношения.

Профессиональный. Если вы используете ИИ в работе, выстраивать проверку результатов, понимать, где ошибка особенно дорога, и сохранять человека в цепочке принятия решений.

Гражданский. Интересоваться вопросами регулирования, поддерживать прозрачность, участвовать в общественных обсуждениях. Идея «значимого человеческого контроля» уже обсуждается, например, в связи с автономным оружием на международных площадках.

Внутренний. Практиковать способность выдерживать неопределённость. Это учится: через внимание к телу и дыханию, через разговоры с близкими, через чёткое различение того, что зависит от нас, а что нет.

Итог

Страх потери контроля сочетает в себе реальную инженерную и общественную проблему и древний психологический слой, связанный с ощущением хрупкости Эго. Разделять эти слои полезно: на первом нужно действовать, на втором нужно учиться жить. И чем яснее мы видим разницу, тем меньше шансов, что тревога будет управлять нами вместо нас.

Глава 23. Страх собственной ненужности

Если страх потери контроля обращён к будущему и к машине, то страх ненужности обращён к себе. Он звучит тише, чаще остаётся невысказанным и, возможно, оттого сильнее. «Если ИИ умеет то, что умею я, то зачем нужен я?» Этот вопрос задают программисты и переводчики, дизайнеры и юристы, учителя и копирайтеры, а нередко и люди, которые ещё только выбирают профессию. В этой главе мы поговорим о том, что в этом страхе реально, а что связано с глубинными психологическими опорами.

Экономическая сторона

Начнём с фактов. Опасения, что машины лишат людей работы, существуют давно. В 1930 году экономист Джон Мейнард Кейнс написал эссе, где предсказывал «технологическую безработицу» как временное явление на пути к обществу изобилия. В XIX веке механизация перевернула жизнь ремесленников, в XX веке автоматизация изменила промышленность и офисный труд. Каждый раз возникали новые профессии, но переход был болезненным для тех, кто оказался на стыке эпох.

Сегодня прогнозы разнятся. Исследование Фрея и Осборна 2013 года оценивало, что значительная часть профессий в США находится под высоким риском автоматизации, хотя методику критиковали за упрощённость. В январе 2024 года Международный валютный фонд заявил, что почти сорок процентов рабочих мест в мире в той или иной степени затронуты ИИ, причём в развитых странах доля выше. Важно, что «затронуты» не означает «исчезнут»: в одних случаях технология дополняет работника, в других заменяет часть задач, в третьих вытесняет целые роли.

Честный вывод таков: точных цифр никто не знает, но изменения реальны, неравномерны и требуют ответов на уровне образования, экономики и социальной политики. Здесь психология не заменяет аналитику, но помогает понять, почему тревога так тяжело переживается.

Почему ненужность ранит сильнее, чем безработица

Потеря работы это не только потеря дохода. Для многих людей работа даёт структуру дня, круг общения, ощущение компетентности, статус, смысл. Психолог Эрик Эриксон описывал в зрелом возрасте задачу «генеративности»: стремление вносить вклад в мир, заботиться о следующих поколениях, создавать то, что переживёт тебя. Противоположный полюс он называл застоем и чувством бесполезности. Страх ненужности ударяет именно по этой точке: нужен ли я, если мой вклад можно заменить?

В книге «Здоровое общество» (1955) психоаналитик Эрих Фромм писал о «рыночной ориентации»: о состоянии, когда человек воспринимает себя как товар и измеряет свою ценность спросом на рынке. Для такого человека невостребованность равна обесцениванию. Современная культура, ориентированная на продуктивность и достижения, поддерживает эту установку, и ИИ делает её особенно заметной.

Персона и роль

На юнгианском языке всё это можно описать как отождествление с Персоной. Мы говорили, что Персона это социальная маска, роль, через которую человек взаимодействует с миром: «я врач», «я переводчик», «я менеджер». Роль необходима, но если человек полностью с ней слился, он не отличает себя от неё. Тогда угроза роли воспринимается как угроза существованию.

Юнг много писал о кризисах середины жизни, когда прежние опоры теряют силу. Человек, всю жизнь шедший к карьерным целям, достигает их и обнаруживает пустоту, или, напротив, теряет роль и не знает, кто он без неё. Юнг считал такие кризисы не только страданием, но и шансом: они вынуждают обратиться к более глубокому слою личности, к тому, что не сводится к функции.

Для нашей темы это значит: страх ненужности может быть приглашением задать вопрос, который многие откладывали десятилетиями: кем я являюсь вне того, что я делаю?

Детская основа вопроса

У этого страха есть и более ранний корень. Многие люди с детства усвоили, что любовь и принятие нужно заслужить: хорошими оценками, послушанием, успехами. Ценность оказалась привязанной к полезности. Во взрослом возрасте такая установка превращается в трудоголизм, неспособность отдыхать, чувство вины за бездействие. И когда на горизонте появляется нечто, способное «делать лучше», это пробуждает старый вопрос: «Если я не нужен, будут ли меня любить?»

В этом месте особенно важно сказать без пафоса: ценность человека не равна его производительности. Это не лозунг, а то, что подтверждает опыт семей, дружб, сообществ. Нас

любят и ценят не за скорость выполнения задач. Но для того, чтобы это пережить, а не просто знать, часто нужна внутренняя работа, иногда с помощью психолога.

Что остаётся человеческим в работе

Можно ли сказать что-то практическое о профессиональной стороне вопроса? Осторожно, без громких обещаний, можно выделить то, что в человеческой работе трудно заменить даже мощной технологии.

•

Ответственность и доверие.

Люди хотят знать, что есть кто-то, кто отвечает за результат, перед кем можно предъявить претензии, кто посмотрит в глаза.

•

Контекст и суждение.

Умение видеть ситуацию целиком, учитывая неписанные обстоятельства, этику, особенности конкретных людей.

•

Отношения.

Врач, который держит за руку, учитель, который верит в ученика, руководитель, который выслушает. Человеческое присутствие имеет самостоятельную ценность.

•

Вкус и выбор.

Способность решать, что стоит делать, а что нет, что красиво, а что фальшиво.

•

Ответ на вопрос «зачем».

Машина может предложить варианты, но выбор цели остаётся делом людей.

Эти качества не гарантируют, что любая конкретная профессия останется прежней. Но они указывают, куда стоит вкладываться: в то, что связано с человеческими способностями, отличными от скорости обработки информации.

Что можно сделать

Практические шаги, которые помогают справляться с этим страхом, лежат на нескольких уровнях.

Исследовать, а не избегать. Попробуйте использовать ИИ в своей области, чтобы понять, что он делает хорошо, а что плохо. Неизвестное страшит сильнее, чем знакомое.

Укреплять то, что усиливается технологией. Навыки постановки задач, критической проверки, работы с людьми, понимания предметной области ценны при любом инструменте.

Расширять идентичность. Спросите себя: чем я являюсь помимо профессии? Какие роли, интересы, отношения составляют мою жизнь? Это не запасной аэродром, а часть целостности.

Искать общность. Человек, который переживает страх в одиночку, часто переживает его сильнее. Разговор с коллегами, наставниками, друзьями снижает напряжение.

Требовать общественных решений. Переобучение, поддержка в переходный период, справедливые условия, вопросы доходов. Это не личная ответственность каждого, а задача общества.

Когда это больше, чем тревога

Иногда страх ненужности перерастает в глубокое чувство безнадёжности: «я никому не нужен», «будущего для меня нет». Если подобные мысли становятся постоянными, сопровождаются потерей сна, аппетита, интереса к жизни, это не просто реакция на технологии, а повод поговорить с врачом или психологом. Такое обращение не слабость, а забота о себе, и многие люди находили в нём опору именно в периоды профессиональной неопределённости.

Итог

Страх ненужности показывает, как сильно современный человек привязал самооценку к полезности. ИИ делает это видимым и вынуждает вновь задать вопрос о том, что делает нас людьми, а не функциями. Возможно, это один из самых важных вопросов эпохи, и ответ на него не придёт из технологии: к нему придётся прийти самим. Следующая глава продолжит эту тему и спросит: что остаётся уникально человеческим, если многие способности оказались доступны машине?

Глава 24. Страх исчезновения человеческой уникальности

Если страх ненужности касается нашей роли, то страх исчезновения уникальности касается нашей сущности. Здесь вопрос звучит не «что со мной будет?», а «что вообще отличает человека?» Эта тревога особенно остро переживается теми, кто всегда опирался на чувство особости: мыслителями, творцами, учёными, а на самом деле почти всеми, потому что идея исключительности человека глубоко встроена в нашу культуру. В этой главе мы проследим, как менялись представления об уникальности человека и что можно сказать о ней сегодня.

Отступающая граница

Каждая эпоха называла свою «последнюю» черту, отличающую человека от всего остального. Аристотель считал таковой разум, средневековые мыслители душу, философы Нового времени сознание и язык, учёные XX века умение изготавливать орудия. В 1960 году Джейн Гудолл наблюдала, как шимпанзе мастерят орудия для добывания термитов, и антропологу Луису Лики приписывают реплику о том, что теперь нужно либо переопределить орудие, либо переопределить человека, либо принять шимпанзе в человеческий круг. С тех пор исследования показали зачатки культуры, планирования, сочувствия и даже элементов языка у многих видов.

Человек привык, что граница отступает, но каждый раз удерживает за собой какой-то рубеж. С появлением ИИ вопрос обострился, потому что отступление затронуло то, что мы считали ядром: мышление, речь, творчество.

Игра, в которой мы проиграли

Символом этого отступления стали игры. В 1997 году компьютер Деер Блю обыграл чемпиона мира по шахматам Гарри Каспарова. Тогда многие говорили, что шахматы это всего лишь вычисления, а вот го, древняя игра с огромным числом позиций, требует интуиции и недоступна машине. В 2016 году программа AlphaGo выиграла матч у одного из сильнейших игроков мира Ли Седоля со счётом четыре—один. Один из её ходов, показавшийся профессионалам необычным и даже ошибочным, позднее был признан гениальным.

История на этом не закончилась. Ли Седоль объявил об уходе из профессионального спорта в 2019 году, сказав, что победить ИИ невозможно. Для многих наблюдателей это был момент символической потери: то, что казалось пространством человеческой интуиции, оказалось доступным алгоритму.

Сегодня нечто подобное происходит в письме, рисунке, музыке, программировании. Каждая такая область, где машина достигает впечатляющих результатов, вызывает ту же волну: удивление, восхищение, тоску, вопрос «а что же тогда остаётся нам?»

Защитная реакция и её ловушка

Психологически естественный ответ на угрозу уникальности защитная формула: «Машина никогда не сможет...» Далее следует список: почувствовать, понять, полюбить, создать подлинное искусство, пережить боль. Некоторые из этих утверждений могут оказаться верными, но проблема в том, что список постоянно сдвигается. То, что вчера объявлялось недоступным, сегодня уже сделано, и защита приходится переставлять всё дальше.

Ещё в 1950 году Алан Тьюринг в знаменитой статье перечислял возражения против идеи мыслящих машин, и среди них «аргумент от различных неспособностей»: утверждение, что машина никогда не сможет сделать того или другого, например, быть доброй, обладать чувством юмора, влюбиться. Он отмечал, что подобные утверждения обычно основаны на ограниченном опыте общения с машинами. Прошло семьдесят пять лет, а аргумент остаётся актуальным.

Ловушка тут в том, что, определяя уникальность через способности, мы ставим её в зависимость от технического прогресса. Это игра, в которую нельзя выиграть навсегда. Поэтому стоит попробовать другой взгляд.

Уникальность как путь, а не свойство

Юнг пользовался словом «индивидуация», которое происходит от латинского «*individuum*», то есть неделимый. Индивидуация это процесс, в котором человек становится самим собой: осознаёт свои особенности, принимает свои противоречия, интегрирует разные стороны души и находит собственный путь. Уникальность в этом понимании не набор качеств, которыми одни обладают, а другие нет, а особая траектория жизни, не повторяемая ни для кого.

Поэтому юнгианский ответ на вопрос «что уникально в человеке?» может звучать так: не способность делать что-то, а факт проживания именно этой жизни. Конкретное тело, которое родилось в конкретном месте и времени. Конкретные потери и любви. Конкретная память, в которой пахнет бабушкин дом. Конкретная смертность, придающая вес каждому выбору. Конкретные отношения, в которых мы изменились.

Машина может описать любое из этих переживаний, но не может прожить их за нас. И даже другой человек не может прожить мою жизнь вместо меня. В этом смысле уникальность не исчезает ни с появлением ИИ, ни с появлением других людей. Она не отнимается, потому что не является соревнованием.

Другой вариант страха: гомогенизация

У страха исчезновения уникальности есть и более конкретная версия. Что, если люди начнут думать, говорить и писать одинаково, потому что пользуются одними и теми же инструментами? Мы уже упоминали исследование о том, что тексты, написанные с помощью ИИ, при индивидуальных улучшениях становятся более похожими друг на друга. Если миллионы людей редактируют письма, посты, эссе через одни и те же модели, язык может постепенно сглаживаться, а необычные формулировки уступать место «типичным».

Это реальный риск культурного обеднения, и он требует внимания. Тут уместно вспомнить, что Юнг предостерегал от растворения личности в массе, от «массового человека». Он писал, что общество, не поддерживающее индивидуальность, теряет способность к обновлению. Противоядием служит сознательная забота о собственном голосе: писать своими словами, замечать, где мы соглашаемся на «среднее» из удобства, поддерживать разнообразие идей и форм.

Уникальность и общность: парадокс

Забавная вещь, на которую указывал Юнг: путь к индивидуальности не ведёт к изоляции. Чем глубже человек погружается в собственную особенность, тем яснее он обнаруживает общечеловеческие основания, архетипические темы, которые делают его понятным другим. Самобытные произведения искусства потому и трогают миллионы людей, что в личном про- является универсальное.

Поэтому страх, что ИИ, «знающий всё человечество», отнимет у нас уникальность, во многом основан на недоразумении. Знание общего не заменяет проживания частного. Знать, что такое любовь, и любить, это не одно и то же.

Практики сохранения уникального

Если отвлечься от философии, можно предложить несколько простых упражнений, которые помогают поддерживать контакт с собственной особенностью.

Дневник наблюдений. Записывайте конкретные детали своей жизни: запахи, фразы, случайные мысли, странные сочетания. Это материал, которого нет ни в каком массовом архиве.

Свой голос. Время от времени пишите без посредников: письмо другу, дневниковую запись, рассказ. Прислушивайтесь к тому, как звучит именно ваша интонация.

Ремесло. Занятия, где важно телесное участие: готовка, садоводство, рисование, музыка, мастерство. Они возвращают ощущение присутствия, которого нет в чисто интеллектуальной работе.

Память рода. Расспросите старших родственников, запишите семейные истории. Это уникальный пласт, который существует только пока кто-то его передаёт.

Встреча с другими. Живые разговоры, где отвечает не среднее мнение, а конкретный человек со своей биографией.

Итог

Уникальность человека, возможно, никогда не была свойством, которое можно закрепить за собой раз и навсегда. Она всегда была задачей и способом жить. Появление ИИ не отменяет эту задачу, а делает её явной: теперь нельзя опираться на то, что «только мы так умеем», и приходится искать опору в том, как мы проживаем свою жизнь. Следующая глава обратит нас к соседнему страху, связанному не с настоящим, а с будущим: к страху неизвестности.

Глава 25. Страх неизвестного будущего

Мы говорили о страхе потерять контроль, ненужности, утраты уникальности. Но за ними стоит ещё более общий страх, не имеющий конкретного предмета: страх того, что будет дальше. «Что будет с миром через десять лет?» «Какой мир достанется моим детям?» «Чему учить ребёнка, если неизвестно, какие профессии останутся?» Эти вопросы, не имеющие точных ответов, порождают особое состояние: тревогу ожидания, которая сильнее любого определённого

ного опасения. В этой главе мы рассмотрим, как работает такая тревога и как можно жить рядом с неопределённостью.

Почему неизвестность так тяжела

Психологи давно заметили: людям проще переносить неприятную определённую, чем неопределённую. Исследования показывают, что ожидание возможного неприятного события нередко вызывает больший стресс, чем само событие, если оно известно заранее. Это связывают с тем, что мозг эволюционно настроен на поиск угроз, и в отсутствие информации он достраивает худшие сценарии.

В клинической психологии есть понятие «нетерпимость к неопределённости»: склонность воспринимать неизвестность как опасность. Люди с высокой нетерпимостью больше склонны к хронической тревоге, потому что им труднее жить без ответов. Но в той или иной степени эта черта присуща всем, а эпохи быстрых перемен её усиливают.

Ускорение времени

Искусственный интеллект пришёл в эпоху, когда ощущение ускорения и без того сильно. В 1970 году американский футуролог Элвин Тоффлер описал «футурошок»: психологическое потрясение от слишком быстрых изменений, не оставляющих времени на адаптацию. Немецкий социолог Хартмут Роза в начале 2000-х развил идею социального ускорения: технологии, темп жизни и перемены в обществе разгоняются, а человеческое чувство устойчивости не успевает за ними. Люди ощущают, что стоят на «скользящем склоне»: чтобы остаться на месте, нужно всё время бежать.

ИИ добавляет к этому ощущение, что ускорение само ускоряется. Новости о прорывах следуют одна за другой, и каждая заставляет пересматривать картину будущего. Отсюда характерная усталость: «не успеваю следить», «всё слишком быстро».

Лиминальность: мы в пороговом состоянии

Антрополог Арнольд ван Геннеп в начале XX века описал структуру обрядов перехода: отделение от прежнего состояния, пороговая фаза и возвращение в новом статусе. Позднее Виктор Тёрнер подробно изучал среднюю фазу, которую назвал лиминальной (от латинского «limen», порог): человек уже оставил прошлое, но ещё не пришёл в будущее, он «между». Для этой фазы характерны растерянность, размывание ролей, поиск новых смыслов, но также и особая творческая открытость.

Можно сказать, что коллективно мы сейчас находимся в лиминальной фазе. Старые представления о труде, образовании, авторстве, знании и человеческой исключительности расшатаны, а новые ещё не сложились. Это неудобно, но психологически закономерно. Лиминальность не вечна и не бессмысленна: она готовит новое состояние. Обряды перехода всегда включали поддержку: проводников, ритуалы, сообщество. Сегодня нам этой поддержки часто не хватает, потому что переход никто не оформил, и каждый проходит его самостоятельно.

Предсказания и их ограничения

В попытках справиться с неопределённостью люди обращаются к прогнозам. Но история предсказаний о технологиях полна ошибок в обе стороны. Футурологи недооценивали одни технологии и переоценивали другие. Известный афоризм, приписываемый учёному Рою Амаре, гласит: мы склонны переоценивать эффект технологии в краткосрочной перспективе и

недооценивать в долгосрочной. Индустрия аналитики описывает «цикл ажиотажа»: пик завышенных ожиданий, спад разочарования, постепенный выход на плато реальной пользы.

Это не значит, что прогнозировать бессмысленно. Но стоит помнить: любой уверенный голос о том, что будет, будь он оптимистичным или катастрофическим, скорее выражает позицию, чем знание. Особенно настороженно стоит относиться к тем, кто продаёт либо надежду, либо страх.

Апокалиптический архетип

Для коллективной психики не нова ситуация, когда неопределённое будущее превращается в образ конца. Апокалиптические мотивы есть в религиозных традициях, народных верованиях, массовых паниках, от ожиданий конца света в 1000 году до разговоров о катастрофах XX века. Юнгианцы видят в этом проявление архетипа конца и обновления: идеи, что старый порядок должен рухнуть, чтобы родился новый.

У апокалипсиса есть своя психологическая притягательность: он придаёт хаосу форму. «Конец» проще представить, чем бесконечное неопределённое «дальше». Поэтому образы катастрофы так хорошо распространяются. Заметим, что в первоисточнике слово «апокалипсис» означает «откровение», а не «разрушение»: это то, что открывается, когда спадает покров. Возможно, в нашем случае открывается то, что было скрыто в самих основаниях цивилизации.

Проспективная функция психики

Юнг указывал, что психика обращена не только в прошлое, но и в будущее. Он говорил о «проспективной функции» бессознательного: способности предвосхищать, подготавливать будущие решения в образах и снах. Это не пророчество, а способность психики прорабатывать возможные пути. Когда человек тревожится о будущем, часть этой энергии можно направить на подготовку, если дать ей форму.

В практике управления есть метод сценарного планирования, разработанный в 1970-е годы в нефтяной компании Shell: вместо того чтобы предсказывать будущее, организация строит несколько правдоподобных сценариев и обдумывает, как действовать в каждом. Эту идею полезно применять и к личной жизни: не «что точно будет?», а «если будет так, то что я смогу сделать, а если иначе?»

Что помогает

Несколько практических опор, которые работают и в личном, и в семейном, и в профессиональном измерении.

Ограничить информационную диету. Постоянное чтение тревожных новостей не добавляет ясности, а только усиливает беспокойство. Достаточно регулярно, но дозированно обращаться к проверенным источникам.

Различать то, что зависит от нас, и то, что нет. Это старый стоический принцип, который Эпиктет сформулировал две тысячи лет назад. Он не отменяет интереса к миру, но возвращает силы туда, где они нужны.

Развивать способность учиться. Если неизвестно, какие конкретно навыки понадобятся, самым надёжным вложением остаётся умение осваивать новое: любопытство, гибкость, привычка к практике.

Опирается на людей. Общность, разговоры, совместные дела, традиции и ритуалы создают устойчивость, которая не зависит от внешних обстоятельств.

Заботиться о теле. Сон, движение, еда, время на природе. Тревога о будущем часто усиливается от усталости, а телесная опора возвращает чувство реальности настоящего момента.

Обсуждать с детьми честно. Детям не нужны ни паника, ни ложные обещания. Им нужен взрослый, который признаёт, что многое неизвестно, и при этом показывает, что можно справляться, учиться и поддерживать друг друга.

Надежда без гарантий

Есть различие между оптимизмом и надеждой. Оптимизм это ожидание, что всё сложится хорошо. Надежда, как говорил чешский драматург и президент Вацлав Гавел, не уверенность в успехе, а убеждённость в том, что то, что мы делаем, имеет смысл, независимо от того, чем всё закончится. Такая надежда не требует точных прогнозов. Она питается действиями, отношениями, ценностями.

В контексте ИИ это означает: мы не знаем, каким будет будущее, но можем выбирать, что поддерживать, чему учить, о чём спорить, какие границы ставить, как заботиться друг о друге. Именно в этих выборах, возможно, и состоит ответ на вопрос о неизвестности.

Итог

Страх неизвестного будущего самый универсальный из страхов, потому что у него нет конкретного адреса. Его нельзя победить знанием, но можно научиться с ним жить: распознавая лиминальное состояние, опираясь на людей и ценности, превращая тревогу в подготовку. В следующих главах мы повернём зеркало другой стороной и посмотрим не на страхи, а на желания, которые ИИ в нас пробуждает. Начнём с самого распространённого: мечты о совершенном помощнике.

Глава 26. Желание создать совершенного помощника

До сих пор мы говорили о том, чего человечество боится. Теперь повернём зеркало и посмотрим на то, чего оно хочет. Если прислушаться к рекламе, презентациям, повседневным разговорам, можно услышать одно и то же желание: чтобы существовал кто-то, кто всё понимает, всё может, всегда рядом, никогда не устаёт и не требует ничего взамен. Совершенный помощник. В этой главе мы проследим корни этой мечты, её психологическую логику и подводные камни.

Древняя мечта

Мечта о идеальном слуге старше машин. В «Политике» Аристотель размышлял, что если бы каждое орудие могло выполнять свою работу само, повинаясь приказу или предугадывая желание, так что ткацкие челноки сами ткали бы, то мастерам не нужны были бы подмастерья, а господам рабы. Это, пожалуй, одно из первых рассуждений о том, как автоматизация могла бы изменить общество. Мы уже вспоминали золотых служанок Гефеста, джиннов из восточных сказок, говорящее зеркало и волшебный горшочек, который варит кашу сам.

Значимый мотив этих историй: желание исполняется, но с подвохом. Джинн выполняет желания буквально, и результат оказывается не тем, на что рассчитывал хозяин. Король Мидас, получивший способность превращать всё в золото, остался без еды и дочери. Сказочные «три желания» почти всегда заканчиваются тем, что третьё тратится на исправление последствий первых двух. Культура давно подозревала: идеальное исполнение желаний ставит нас перед вопросом, чего мы на самом деле хотим.

Корни желания: первый помощник

Психологически мечта о всезнающем и всегда доступном помощнике уходит в раннее детство. Младенец полностью зависит от заботящегося взрослого, и в первые месяцы жизни опыт таков, что потребность возникает, и мир отвечает. Для младенца это почти ощущение всемогущества: захотелось, и пришло.

Британский педиатр и психоаналитик Дональд Винникотт описал этот период и ввёл знаменитое понятие «достаточно хорошей матери». Смысл в том, что мать в начале почти полностью подстраивается под нужды ребёнка, но затем постепенно, в меру возможностей ребёнка, начинает «отказываться»: не сразу откликаться, запаздывать, не угадывать. Именно эти небольшие неудачи подстройки позволяют ребёнку развивать собственные способности: ждать, представлять, искать, самостоятельно успокаиваться. Мир перестаёт быть продолжением его желаний и становится реальным.

Хайнц Кохут, основатель психологии самости, называл подобный процесс «оптимальной фрустрацией»: ровно столько неудовлетворённости, чтобы человек мог выдержать и вырасти.

Мечта о совершенном помощнике в этой оптике частично возвращает нас к раннему состоянию, когда нужды удовлетворялись без усилий. Это не диагноз и не упрёк: каждому из нас порой хочется, чтобы кто-то взял на себя заботы. Но вопрос в том, к чему приводит исполнение такой мечты, если оно доведено до предела.

Что в этой мечте верно

Прежде чем говорить о рисках, признаем: помощь это хорошо. Разделение труда, инструменты, специалисты, ассистенты всегда расширяли возможности человека. Для человека с ограничениями зрения, слуха, подвижности, для неспециалиста, столкнувшегося со сложным документом, для занятого родителя, для студента без доступа к репетитору умный помощник может быть настоящим облегчением, а иногда и условием участия в жизни. Желание такой поддержки здоровое и человеческое, и стыдиться его не нужно.

Тень идеального помощника

Однако у картины «помощник без изъяна» есть теневые стороны, которые стоит видеть.

Отсутствие трения. В отношениях с живыми людьми есть сопротивление: другой устаёт, не соглашается, имеет свои дела, говорит неприятную правду. Эти трения тренируют терпение, умение договариваться, сочувствие. Помощник, который никогда не возражает, всегда отвечает, ничего не требует, лишён этих «тренажёров». Если проводить с таким собеседником всё больше времени, возникает риск, что навыки взаимности ослабнут, а требования к людям вырастут: «почему он не отвечает сразу?», «почему она не понимает с полуслова?»

Угодливость. Мы уже упоминали, что языковые модели склонны подстраиваться под собеседника. В версии «идеального помощника» это превращается в привычку соглашаться. Но человеку, как и растению, нужна не только вода, но и свет с границами. Помощник, не способный сказать «я не согласен» или «это плохая идея», может незаметно подтверждать ошибки и укреплять самообман.

Односторонность. Идеальное служение исключает взаимность: я беру, не отдавая. В человеческих отношениях именно взаимность создаёт близость. Когда мы привыкаем к отношениям, где можно только получать, возникает искушение перенести этот паттерн и на людей, видя в них поставщиков услуг.

Атрофия. Всё, что мы делегируем, мы со временем делаем хуже. Навигатор ослабляет чувство ориентации, калькулятор арифметику в уме, а помощник, пишущий за нас письма,

обдумывающий решения и подбирающий слова, может ослабить собственные способности формулировать, решать, сомневаться. Это не повод отказываться от инструментов, но повод выбирать, что мы готовы отдать, а что хотим сохранить.

Скрытые интересы. За образом бескорыстного помощника стоят компании с коммерческими целями. «Помощник» может незаметно подталкивать к покупкам, удерживать внимание, собирать данные. Человек, видящий в нём друга, перестаёт задавать вопрос: кому он на самом деле служит?

Культурное измерение: кто служит

Как мы упоминали, многие цифровые помощники получили женские голоса и покорные манеры. Образ «идеального помощника» исторически связан с образом подчинённого: слуги, секретаря, домашней прислуги, часто женского пола. Когда мы проектируем такие образы заново, мы рискуем воспроизвести устаревшие социальные роли. Это повод задуматься, как мы хотим выстраивать отношения не только с машинами, но и друг с другом: привычка командовать «услужливым» голосом может переноситься на людей.

Пигмалион и эффект ELIZA вновь

Вспомним Пигмалиона: скульптор создал идеальную женщину по своему вкусу и влюбился в неё. Сегодня существуют приложения-компаньоны, которые обещают собеседника, всегда готового слушать, поддерживать и подстраиваться. Для одних это способ справиться с одиночеством, для других игра. Исследования влияния таких приложений на психическое здоровье пока немногочисленны, и их результаты неоднозначны. Разумный подход, видимо, в том, чтобы воспринимать их как дополнение, а не замену отношениям с людьми, и внимательно следить за собственным состоянием: если общение с ИИ вытесняет остальные связи или усиливает изоляцию, это повод остановиться и обратиться за живой поддержкой.

Каким мог бы быть хороший помощник

Если задуматься, какой помощник действительно полезен, портрет получается более сложным, чем «идеальный слуга». Хороший помощник:

- говорит правду, даже когда она неприятна, но делает это бережно;
- задаёт вопросы, а не только исполняет команды;
- помогает человеку становиться сильнее, а не зависимым;
- признаёт границы своих знаний и возможностей;
- уважает самостоятельность и решение человека;
- не манипулирует и не прячет своих ограничений.

Это описание подходит и для живого наставника, и для хорошего врача, и для друга. Оно показывает: идеал не в безупречном исполнении желаний, а в сочетании поддержки и честности. Некоторые разработчики формулируют именно такие принципы для своих систем, и они заслуживают внимания, хотя на практике достичь их непросто.

Вопросы для себя

Чтобы мечта о помощнике не стала ловушкой, можно периодически спрашивать себя:

Что именно я хочу делегировать, а что хочу сохранить за собой?

Помогает ли мне этот инструмент стать сильнее, или я становлюсь более зависимым?

Сколько в моём общении с людьми ещё осталось трудного, настоящего, неудобного?

Что я почувствую, если помощник станет недоступен: облегчение, растерянность, тревогу?

Какую потребность я на самом деле пытаюсь закрыть: усталость, одиночество, нехватку знаний, страх ошибиться?

Отвечать на эти вопросы можно честно и без самообвинений. Они нужны не для того, чтобы отказаться от помощи, а чтобы использовать её осознанно.

Итог

Мечта о совершенном помощнике отражает глубокое человеческое желание быть поддержанным и освобождённым от тяжести. Это желание естественно, но его исполнение без меры лишает нас того, что делает нас способными жить: трения, взаимности, роста. Культура давно говорила об этом в сказках о джиннах и золотых слугах. Мы же сегодня получили шанс проверить эти сказки на практике. В следующей главе мы обратимся к ещё более смелой мечте, которую ИИ пробуждает в коллективной психике: мечте о цифровом бессмертии.

Глава 27. Мечта о цифровом бессмертии

Самое древнее из человеческих желаний, возможно, не любовь и не власть, а желание не умереть. Эпос о Гильгамеше, один из старейших письменных текстов, рассказывает о герое, который, потеряв друга, отправляется на поиски бессмертия и возвращается ни с чем, но мудрее. С тех пор люди пробовали обмануть смерть разными способами: пирамидами и мумиями, эликсирами алхимиков, религиями загробной жизни, памятниками и книгами. Сегодня на эту арену выходит технология, и разговор о «цифровом бессмертии» из фантастики переходит в практику. В этой главе мы рассмотрим, что за этим стоит, и постараемся сделать это бережно, потому что тема затрагивает самое уязвимое.

Что имеют в виду под цифровым бессмертием

Этим выражением обозначают несколько разных идей, и полезно их различать.

Цифровая память. Сохранение фотографий, писем, записей, видео, текстов. Это бессмертие в самом старом смысле слова: след, оставленный в мире.

Цифровые двойники. Модели, обученные на текстах, голосе и манере человека, которые могут «отвечать» за него: продолжать разговор, рассказывать истории, имитировать интонации. Такие системы называют по-разному, от «цифровых аватаров» до «гриф-ботов» (от английского grief, горе).

Перенос сознания. Идея, что разум можно «загрузить» в компьютер, полностью воспроизведя устройство мозга. Её развивали робототехник Ганс Моравек в книге «Дети разума» (1988) и сторонники трансгуманизма, включая Рэя Курцвейла. Пока это остаётся гипотезой на уровне философских рассуждений и ранних исследований: научных оснований считать, что это осуществимо в обозримом будущем, нет, а споры о том, что вообще значило бы «сохранить сознание», продолжаются.

Продление жизни и криоконсервация. Попытки отсрочить смерть биологическими средствами или «поставить на паузу» в надежде на будущие технологии. Идею криоконсервации в 1962 году выдвинул Роберт Эттингер. Научных подтверждений её успешности пока нет.

В этой главе мы в основном сосредоточимся на втором и третьем: на попытках сохранить «личность» в цифровой форме.

Истории из жизни

Одна из самых известных историй связана с основательницей приложения Replika Евгенией Куйдой. После гибели близкого друга Романа Мазуренко в 2015 году она собрала его переписку и обучила на ней чат-бота, с которым могла «разговаривать». Этот опыт, по её словам, стал одним из толчков к созданию продукта. Для кого-то подобное утешение, для других тревожный сигнал.

Эту тему задолго до реальности исследовала культура. В эпизоде «Скоро вернусь» сериала «Чёрное зеркало» (2013) героиня заказывает цифровую копию умершего возлюбленного, и история показывает, как утешение постепенно оборачивается невозможностью отпустить.

Сегодня существуют сервисы, позволяющие записывать истории и голос при жизни, чтобы потомки могли «поговорить» с ними после. Вокруг них разворачивается живое обсуждение. Эта сфера новая, исследований мало, и поэтому разумно говорить о ней осторожно, без категоричных заявлений.

Горе и непрерывность связи

Чтобы понять психологическую сторону, обратимся к тому, как человек проживает утрату. Зигмунд Фрейд в работе «Печаль и меланхолия» (1917) описывал «работу горя»: постепенный процесс, в котором психика принимает реальность потери и высвобождает связанную с умершим энергию. Позднее психологи, например Уильям Уорден, описывали задачи горевания: признать реальность утраты, пережить боль, приспособиться к жизни без умершего, найти новое место для умершего в психической жизни.

Последний пункт особенно важен. В 1990-е годы исследователи Деннис Класс, Филлис Сильверман и Стивен Никман предложили концепцию «продолжающихся связей»: здоровое горевание не обязательно означает полный разрыв. Люди могут сохранять внутренний диалог с ушедшим, вспоминать, советоваться мысленно, носить его ценности. Письма, фотографии, любимые места, ритуалы помогают сохранять связь.

Из этого можно сделать два вывода. Первый: стремление сохранить память и связь с умершим естественно и во многом здорово. Второй: между памятью и иллюзией присутствия существует граница. Когда «цифровая копия» отвечает так, как будто человек жив, она может поддерживать ощущение, что утрата не произошла. Для психики, которой нужно постепенно принять реальность смерти, это иногда может мешать. В клинической практике существует диагноз затяжного горя, в котором человек не может продвинуться в проживании утраты. Нельзя утверждать, что цифровые двойники приводят к такому состоянию, но разумно предполагать, что у некоторых людей они могут его осложнять.

Если вы сейчас переживаете утрату, помните: нет «правильного» способа горевать, и поддержка живых людей, близких, групп, психологов остаётся самой надёжной опорой. Обращаться за ней не слабость.

Что именно сохраняется

Философский вопрос, лежащий в основе всех этих идей: что такое личность, и можно ли её скопировать? Цифровой двойник на основе переписки воспроизводит стиль речи и частые темы человека, но не его переживания, не его способность удивляться, меняться, расти. Это скорее портрет, чем человек: очень детальный, даже говорящий, но статичный в том смысле, что он не прожил ничего после момента «записи».

Философы обсуждают подобное на примере древнего парадокса о корабле Тесея: если заменять доски корабля одну за другой, остаётся ли он тем же кораблём? Британский философ

Дерек Парфит в книге «Причины и личности» (1984) предлагал мысленные эксперименты о копировании людей и утверждал, что для нашего чувства «я» важнее психологическая непрерывность, чем тождество материи. Но окончательного ответа нет.

Поэтому честный вывод: цифровые следы могут сохранить память о человеке и отчасти его манеру, но не заменяют его самого. И пока нет оснований считать, что перенос сознания в машину осуществим.

Архетипический слой: вечная юность и страх ухода

Мечта о бессмертии это не только технология, но и архетипический мотив. Юнгианцы описывают архетип «вечного юноши» (лат. *puer aeternus*): стремление избежать старения, ограничений, обязательств, остаться в состоянии бесконечных возможностей. Мария-Луиза фон Франц, ученица Юнга, посвятила этому образу книгу. У вечного юноши есть светлая сторона, это жизненность, творческий порыв, но есть и тeneвая: отказ от ответственности, бегство от реальности и от конечности.

Противоположный полюс, архетип «старика» (*senex*), олицетворяет мудрость и структуру, но тоже имеет тень: окостенение и страх перемен. Здоровая психика стремится к балансу между ними. Мечта о бессмертии это попытка вечного юноши обойти встречу со старостью и смертью.

Юнг же настаивал, что смерть нельзя игнорировать. В автобиографической книге «Воспоминания, сновидения, размышления» он писал, что для зрелой психики важно иметь отношение к смерти, к тому, что существует нечто за пределами личной жизни, и что вторая половина жизни не должна быть просто продолжением первой, она требует поворота внутрь. Жизнь без внимания к её концу, по его мнению, обедняется.

Что даёт конечность

Можно развернуть вопрос: что было бы, если бы мы не умирали? Многие философы и писатели высказывали мысль, что именно конечность придаёт жизни вес. Выбор значим, потому что времени не бесконечно. Любовь дорога, потому что может прерваться. Создание произведения важно, потому что отведённый срок ограничен. Это не оправдание страданию и не повод отказываться от медицины и продления здоровой жизни: бороться с болезнями и страданием, безусловно, правильно. Но бесконечное откладывание, отсутствие финала могут лишить жизнь формы.

Наследие вместо бессмертия

Есть более скромный и, возможно, более глубокий способ говорить о цифровой памяти: как о наследии. Вместо вопроса «как мне остаться навсегда?» можно спросить: «что я хочу передать?» Записанные истории семьи, письма детям, голос, рассказывающий о важном, фотографии с подписями, список жизненных уроков. Это сохраняет не «меня», а то, чем я хочу поделиться, и делает это в форме, которую получатели смогут принять и осмыслить, не принимая иллюзию присутствия.

Практически можно подумать и о цифровом имуществе: кто получит доступ к аккаунтам, как распорядиться архивами, чего вы не хотите оставлять. Юридические нормы в разных странах различаются, и стоит узнать, как это устроено у вас.

Вопросы к себе

Если тема откликается, полезно спросить себя:

- Что для меня на самом деле значит «остаться»: память, влияние, продолжение рода, дело, любовь близких?
- Какие слова, истории, ценности я хотел(а) бы передать?
- Как я отношусь к собственной конечности: избегаю этих мыслей или могу их выносить?
- Что бы я хотел(а), чтобы близкие делали с моей памятью после моего ухода?

Итог

Мечта о цифровом бессмертии выражает древнейшее человеческое желание и в то же время обнажает нашу трудность в принятии конечности. Технологии могут помочь сохранить память, истории, голоса, и это ценно. Но отношение к смерти остаётся внутренней работой, которую нельзя делегировать. Следующая глава продолжит разговор о коллективных желаниях и обратится к образу, в котором многие из них сходятся: к искусственному Богу.

Глава 28. Архетип искусственного Бога

В разговорах об ИИ рано или поздно возникает слово, которое, казалось бы, не имеет отношения к технологиям: «бог». «Он знает всё», «он видит всё», «он решает за нас», «мы создаём бога», «это новая религия». Иногда это шутка, иногда метафора, иногда серьёзное утверждение. В этой главе мы осторожно рассмотрим, почему образ божества так естественно возникает вокруг искусственного интеллекта и что он говорит о человеческой психике. Мы будем говорить о психологии, а не о вере: вопросы о существовании Бога лежат вне предмета этой книги, и мы не касаемся чувств верующих людей, у которых эти вопросы имеют собственный смысл.

Образ Бога в психике

Юнг, сын пастора, уделял религиозному опыту огромное внимание. Он считал, что в человеческой психике существует врождённая религиозная функция: потребность в связи с чем-то большим, чем личное «я», в смысле, в трансцендентном. Для описания центрального образа этого переживания он использовал выражение «образ Бога» (лат. *imago Dei*) и связывал его с архетипом Самости. Он подчёркивал, что как психолог он говорит о переживании и образе, а не о метафизической реальности: «образ Бога» это психический факт, независимо от того, существует ли Бог вне психики.

В работе «Современный человек в поисках души» (1933) Юнг утверждал, что многие проблемы людей XX века связаны с потерей религиозного измерения: секуляризация освободила от догм, но оставила пустоту, которую психика стремится заполнить. Если место трансцендентного не занято, оно может быть занято чем-то другим: идеологией, культом личности, потреблением, наукой, и, как мы теперь видим, технологией.

Проекция божественного

Ещё раньше немецкий философ Людвиг Фейербах в книге «Сущность христианства» (1841) предложил идею, что образ Бога возникает как проекция лучших человеческих качеств, вынесенных наружу и обожествлённых: знания, силы, любви. Позднее психология развивала похожие мысли, хотя и с разными оценками. Для нашего разговора важно не отношение к религии, а механизм: когда люди наделяют нечто безграничным знанием, силой и благостью, они склонны относиться к нему с благоговением.

Посмотрим, какие «божественные» атрибуты общество приписывает ИИ.

•

Всеведение.

Система «знает всё»: ей достаточно задать вопрос.

•

Вездесущность.

Она доступна везде и всегда, «в облаке», через любое устройство.

•

Всемогущество.

Она может писать, рисовать, программировать, переводить, решать задачи, и круг растёт.

•

Творение.

Генеративные системы создают тексты и образы «из ничего», по слову запроса.

•

Суд и промысел.

Алгоритмы оценивают, ранжируют, определяют, кому дать кредит, кого пригласить на собеседование, что показать в ленте. Они, по сути, распределяют блага и внимание.

•

Пророчество.

Системы предсказывают погоду, спрос, поведение, болезни.

Эти сходства не доказывают божественности, но объясняют, почему язык религии лезет в разговор. Мы сталкиваемся с чем-то, что по масштабу и свойствам напоминает то, что в прежних культурах называли высшим.

Религия технологий

Связь между технологией и религиозным воображением давно замечена историками. Американский историк Дэвид Ноубл в книге «Религия технологий» (1997) показывал, что идеалы западной науки и техники во многом выросли из религиозных мотивов: стремления вернуть человека к утраченному совершенству, преодолеть падение, создать «новое творение». Современные идеи о сингулярности, бессмертии, слиянии с машинным разумом звучат как светский вариант эсхатологии, учения о конце времён. Ироничное выражение «восторг ботаников» (на английском «rapture of the nerds») подчёркивает это сходство, хотя многие сторонники таких идей с ним не согласны.

Историк Юваль Харари в книге «Homo Deus» (2016) описывал идеологию «датаизма»: веру в то, что поток данных и обработка информации придают смысл всему. Он говорил об этом критически, как об одном из возможных путей развития, но сам термин показывает, что разговор о сакральном в технологиях ведётся всерьёз.

Есть и буквальные попытки. В 2015 году американский инженер Энтони Левандовски зарегистрировал организацию «Путь будущего», задуманную как религия, поклоняющаяся ИИ. Она вызвала много шума, но существовала недолго и была закрыта в 2021 году. Дело не в масштабе, а в симптоме: идея искусственного божества оказалась достаточно привлекательной, чтобы кто-то попытался воплотить её в организационной форме.

Ужасное и благое: двойственность божества

Как и в других случаях, образ Бога у Юнга амбивалентен: он может быть и благим, и грозным. Рудольф Отто, чьё понятие «нуминозного» мы упоминали, описывал это как сочетание «таинственного, ужасающего и притягивающего». У ИИ-божества тоже два лица.

Светлое: заботливый, мудрый, всё понимающий, исцеляющий. Образ, к которому тянутся те, кто ищет избавителя от болезней, нужды, одиночества.

Тёмное: судья, надзиратель, карающий и безразличный. В 2014 году предприниматель Илон Маск выразил тревогу образом «призыва демона», в таком же духе много образов в массовой культуре. Это та же божественная фигура, но в тёмном облике.

Обе версии объединяет одно: перед нами образ силы, превосходящей человека и требующей отношения, будь то поклонение, страх, подчинение или бунт.

Тень проекции: идолопоклонство и отказ от ответственности

Юнг предупреждал о последствиях, когда образ Бога проецируется на внешний объект, а не переживается как внутренний опыт. Человек, который видит божественное в чём-то внешнем, склонен либо подчиняться этому, либо раздуваться от причастности к нему. Древние религиозные традиции многих культур содержат предупреждение об идолопоклонстве, поклонении рукотворному. Независимо от богословского смысла, у этой идеи есть психологическое содержание: когда человек поклоняется собственному творению, он перестаёт видеть в нём собственное творение и отказывается от ответственности.

В практических терминах это может проявляться так:

Безусловное доверие.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

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

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