

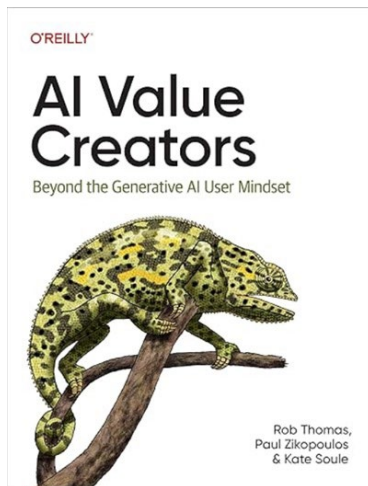
Томас и др. Создание бизнес-ценности с генеративным ИИ

Мы вступили в новую эпоху — генеративный искусственный интеллект (GenAI) кардинально меняет индустрии и процессы принятия решений. В связи с этим становится крайне важным понимать его возможности и риски. Однако, недостаточно просто технических знаний — нужно научиться использовать ИИ стратегически. Эта книга, ориентированная как на бизнес-лидеров, так и на технических специалистов, раскрывает принципы интеграции ИИ в бизнес. В этом руководстве рассматриваются реальные истории успеха и извлеченные уроки. Независимо от уровня вашей подготовки — новичок вы или опытный профессионал — книга поможет вам глубже понять генеративный ИИ.

- Осознайте преобразующий потенциал ИИ в бизнесе и способы его эффективного использования.
- Научитесь уверенно справляться с этическими и операционными вызовами, связанными с ИИ.
- Поймите, как технологии ИИ взаимодействуют со стратегией бизнеса.
- Внедряйте практические стратегии для интеграции ИИ в корпоративную культуру.
- Станьте лидером в создании ценности с помощью ИИ, готовым к инновациям и управлению будущим.

Перевод выполнен с помощью ChatGPT и DeepSeek.

Thomas Rob, Zikopoulos Paul, Soule Kate. AI Value Creators: Beyond the Generative AI User Mindset. – USA, California: O'Reilly Media, 2025. – 297 p.



Содержание

Глава 1. От +AI к AI+: генеративный ИИ и «момент Netscape»

[Глава 2. Как стать создателем ценности с помощью ИИ](#)

Глава 3. ИИ как инструмент убеждения

Глава 4. Глава о практическом применении

Глава 5. Жить, умереть, купить или попробовать — многое решит ИИ

Глава 6. Навыки, которые впечатляют

Глава 7. Куда движется технология — эпоха единой модели никогда не наступит!

Глава 8. Данные как фактор конкурентного преимущества

Глава 9. Генеративные вычисления — новый стиль программирования

Предисловие

«Захватывающе» — именно так мы описываем возможности, которые открывают генеративный ИИ (GenAI) и агенты. От технических помещений до залов заседаний совета директоров — мы уверены, что GenAI действительно изменит каждую отрасль. Как технологические энтузиасты, мы испытываем такой уровень воодушевления лишь раз в десятилетие, и именно поэтому с каждым днем все больше людей по всему миру осознают потенциал GenAI. Действительно, GenAI обещает революцию, но на

этот раз она затронет высокоинтеллектуальную сферу, которую не смогла изменить даже промышленная революция.

Как многие знают, Стив Джобс бросил колледж, но впоследствии возглавил одну из самых успешных компаний в истории. Многие также знают, что Джобс ценил простоту, элегантность и красоту вещей (например, шрифтов, которые он тщательно изучал), и это отражено в продуктах Apple по сей день. Однако немногие в курсе, что его также завораживала эффективность передвижения. Его особенно заинтересовало исследование, посвященное минимальному количеству энергии, которое требуется виду, чтобы преодолеть один километр. Возможно, вас удивит (нас это точно удивило), но победителем в этой категории оказался кондор! А человек? Где-то в нижней трети списка. Но как только человечество село на велосипед, оно оставило всех далеко позади по эффективности. Джобс сделал вывод: люди способны создавать инструменты (в его время — компьютеры, а сейчас — GenAI и агенты), которые делают нас лучше, в то время как другие виды вынуждены адаптироваться (что занимает много времени, если вообще возможно).

В 1990 году он сравнил компьютер с велосипедом: «Для меня компьютер — самый удивительный инструмент, который мы когда-либо создавали. Это как велосипед для нашего сознания». Мы уверены: если бы он был жив сегодня, то сказал бы, что если компьютер — это велосипед для разума, то GenAI и агенты вместе становятся велосипедом для бизнеса. И так же, как люди освоили Excel для управления финансами, отслеживания расходов и создания сложных цветных диаграмм (порой с избыточной детализацией), эти технологии станут стандартным инструментом для автоматизации задач, генерации творческих идей и создания впечатления, будто вы трудились без усталости, чтобы их достичь.

Однако не все так радужно

Хотя все это звучит многообещающе, помните: ИИ — это не гарантия процветания, и лучший мир не наступит автоматически только потому, что вы используете GenAI или агентов. Почему? У технологии есть и темная сторона, о которой вы также узнаете из этой книги.

Времена меняются, технологии — еще быстрее. Давайте сразу проясним: кое-что устареет к моменту, когда вы будете читать эту книгу. Писать об ИИ сегодня — все равно что сообщать статистику по количеству изображений, загружаемых в интернет каждую секунду. Мы сделали все возможное, чтобы информация оставалась актуальной (например, когда вышел DeepSeek, мы уговорили редактора открыть рукопись для внесения изменений). Но бесспорно, появятся новые тесты, исследования, передовые модели (SOTA), GPU, ускорители и многое другое.

В этой книге мы упоминаем, что раньше «технологические годы» соотносились с собачьими (1:7), а теперь — скорее с мышинными (1:30). Так что просим нас простить, если какие-то новшества будут противоречить написанному. Однако главная цель книги — дать вам систему мышления для работы с GenAI и агентами: на что обращать внимание, какие вопросы задавать (себе и поставщикам) и так далее. Мы уверены, что эти знания сохранят ценность гораздо дольше, чем информация о том, какая модель сегодня лидирует в рейтингах, а завтра станет предметом шуток в духе: «Неужели этим кто-то всерьез пользовался?»

Поэтому мы призываем вас смотреть глубже цифр и версий моделей, а вдумчиво воспринимать советы из книги — они основаны на успехах, провалах и огромном опыте реальных наблюдений за тем, что работает, а что нет.

Пример: уроки DeepSeek. Взрывной ажиотаж вокруг DeepSeek в начале 2025 года (незадолго до отправки книги в печать) наглядно подтвердил ключевые идеи, которые мы здесь излагаем. Рынок привык считать, что обучение передовых моделей требует миллионных инвестиций в самые современные чипы, что технологии должны быть проприетарными, а секретность — обязательной. DeepSeek опроверг это, используя методы, описанные в нашей книге, например, смесь экспертов, метод сжатия моделей и обучение с подкреплением, а также новые оптимизации. Его выпустили под либеральной лицензией MIT — что соответствует идее открытости, которую мы поддерживаем.

Так что хотя DeepSeek, возможно, не попал в сравнительные таблицы в этой книге, суть не в этом, ведь бенчмарки — игра в "долби крота". Важнее, что принципы, изложенные здесь, уже работают на рынке, стимулируя новые движения и инновации. И да, появятся еще более совершенные методы — но они идеально впишутся в стратегии, которые мы вам предлагаем.

GenAI — это момент подъема, сдвига, разлома или обрыва

Вспомните, когда вы впервые услышали о GenAI. Это понятие вошло в публичный дискурс примерно в конце 2022 года. С тех пор мы увидели новые модели, их эволюцию и взрывной рост открытых моделей. Всего за несколько месяцев GenAI превратился из любопытной новинки в фундаментальную силу, двигающую бизнес-инновации, с постоянным потоком новых применений и сценариев использования.

Рост настолько стремительный, что невозможно точно предсказать, где мы окажемся через 10 лет — или даже через 10 месяцев после завершения этой книги. Но мы точно знаем: вы захотите активно участвовать в формировании этого будущего — и, надеемся, именно поэтому вы сейчас читаете эту книгу. Мы находимся в переломном моменте: мир переходит от процессов, управляемых людьми с поддержкой технологий, к процессам, управляемым технологиями с поддержкой (или помощью) людей.

Для вас как лидера, как личности и для компаний, в которых вы работаете, это действительно момент подъема (хорошо), сдвига (возможности), разлома (не очень хорошо) или обрыва (то, во что вы упретесь, если не будете развиваться).

Будущее ИИ — не в одной универсальной модели, которая сделает все для всех (мы будем повторять это снова и снова: одна модель не сможет править всеми). Будущее ИИ не ограничится мультимодальностью (зрение, слух, письмо и т. д.); оно, несомненно, станет и мультимодельным (так же, как облака стали гибридными). ИИ должен быть демократизирован — и это возможно только при коллективном использовании энергии и прозрачности открытого исходного кода и открытой науки. Это даст каждому голос в определении того, что такое ИИ, как он работает, как используется и как влияет на общество. Это гарантирует, что развитие ИИ будет подвижно не привилегированным меньшинством, а усилиями многих.

Ажиотаж вокруг DeerSeek поднимает важный вопрос: кто будет формировать будущее ИИ? Мы убеждены, что развитие ИИ не должно контролироваться горсткой игроков, особенно если их ценности расходятся с базовыми принципами, такими как защита корпоративных данных, приватность, прозрачность и другие. Мы не можем позволить, чтобы лидерство в ИИ перешло к тем, у кого другие приоритеты и ценности. Это означало бы отдать контроль над технологией, которая изменит каждую отрасль и каждую сферу общества. Именно поэтому мы будем настаивать: истинные инновации и прогресс возможны только через демократизацию ИИ.

Мы считаем, что 2025 год должен стать годом, когда генеративный ИИ выйдет за пределы узкого круга компаний, а к 2026 году надеемся, что широкие слои общества не просто будут использовать ИИ — многие начнут его создавать.

Читая эту книгу, вы поймете, почему, по нашему мнению, значительная часть корпоративного инструментария GenAI будет состоять из компактных открытых моделей — именно так будет строиться будущее. Слишком долго ИИ воспринимался как игра масштабов, где большие модели означали лучшие результаты. Но настоящий прорыв зависит не только от размера, но и от эффективности.

В нашей работе в IBM мы наблюдали, как специализированные модели уже привели к тридцатикратному снижению затрат на вывод ИИ, сделали обучение более эффективным, а технологии — доступнее. Не существует закона физики, согласно которому ИИ должен оставаться дорогим. Затраты на обучение и работу моделей не являются фиксированными. Они — суть инженерной задачи, которую предстоит решить. Компании, как устоявшиеся, так и новые, обладают изобретательностью, чтобы снижать эти затраты и делать ИИ более практичным и распространенным.

Есть старая китайская поговорка о том, когда лучше всего сажать дерево. Каким бы ни был этот момент (он различается в зависимости от того, кто рассказывает), он уже в прошлом. Но нет никаких сомнений насчет следующего лучшего времени: сегодня.

Мы благодарим вас за инициативу в чтении нашей книги. Надеемся, что, закончив ее, вы поблагодарите нас в ответ — ведь мы дадим вам основу для создания собственных планов по внедрению GenAI и агентных систем. Вы узнаете, как эффективно, безопасно и ответственно применять ИИ в бизнесе.

Начало вашего путешествия по книге

В этой книге мы раскроем суть искусственного интеллекта — генеративного ИИ и агентов. Мы рассмотрим, как мы пришли к этому моменту, как работают эти технологии и какими способами они уже меняют (и продолжают менять) бизнес и общество в беспрецедентных масштабах. Мы часто называем текущий момент «моментом Netscape» (Netscape был первым в мире интернет-браузером), потому что считаем, что влияние этой технологии на всех нас будет столь же глубоким.

Прежде чем углубиться в чтение, вот краткий обзор ключевых моментов. Считайте это вашей шпаргалкой. Используйте его как трейлер к книге или чтобы сразу перейти к разделу, который вас заинтересует.

Глава 1. От +AI к AI+: генеративный ИИ и «момент Netscape»

Это редкий момент для бизнеса. В последний раз нечто столь же масштабное произошло в 1993 году, когда веб-браузер (Netscape) вырвал интернет из рук привилегированного меньшинства и сделал его доступным для всех. Не упустите этот момент — изменения будут такими же значительными, как и с появлением интернета. В ближайшем будущем вы будете конкурировать не с ИИ, а с другими компаниями, которые его используют.

Представьте: если ваша компания тратит 25% бюджета на поддержку клиентов, а конкурент переводит две трети этих затрат на ИИ... Генеративный ИИ и агенты станут разделительной линией между компаниями, которые процветают, и теми, кто будет бороться за выживание. Но помните: ИИ — это не магия. Через пару лет, оглядываясь назад, вы ясно увидите, какие компании стали лидерами, какие отстали, а какие неожиданно вырвались вперед. К какой группе вы отнесете себя?

Глава 2. Как стать создателем ценности с помощью ИИ

ИИ можно использовать по-разному. Ваша цель — не просто применять ИИ, а создавать с его помощью ценность! Начните свой путь с осознания того, что ваши данные важны и их нельзя просто отдавать. Чтобы стать создателем ценности, вам понадобится платформа ИИ. Те, кто создает ценность с помощью ИИ, получают гораздо больше преимуществ, чем обычные пользователи.

Глава 3. ИИ как инструмент убеждения

Мы сталкиваемся с парадоксом производительности. Сегодня множество факторов мешают бизнесу добиваться успеха: снижение производительности, демографический спад, рост стоимости заемных средств. У вас есть уникальная возможность использовать ИИ против этих вызовов — особенно за счет повышения производительности и цифрового труда на основе ИИ.

Глава 4. Глава о практическом применении

Речь идет не о побочных проектах, а о решениях, которые приносят реальную ценность. Если вы освоите базовые возможности ИИ (такие как анализ изображений, текста, звука и других данных), вам будет проще выбирать специализированные решения для своего бизнеса. Помните: компьютерное зрение — это компьютерное зрение, обработка текста — это обработка текста. Для компьютера все — просто числа, даже ваш плейлист Taylor Swift на Spotify. И не забудьте перевести свою компанию через «точку ценности» — момент, когда инвестиции в ИИ начинают приносить значимую отдачу.

Глава 5. Жить, умереть, купить или попробовать — многое решит ИИ

Люди будут использовать только тот ИИ, которому доверяют. Вы должны заранее решить, будете ли вы добросовестным игроком или нет — вы уже видели и тех, и других в соцсетях и других инновационных сферах. Примите это решение сразу! Почему? Миру нужны регулирующие нормы, которые поспевают за развитием технологий, а правительства часто работают слишком медленно. В конечном итоге власти (и, надеемся, клиенты) потребуют, чтобы ваш ИИ был объяснимым и подотчетным. Снабдите свой "рюкзак" для путешествия в мир ИИ такими вещами, как справедливость, надежность, прозрачность и отслеживаемость данных — это необходимый набор для достижения новых высот. Убедитесь, что это продумано заранее, а не добавлено потом. Ведь, как говорится в названии главы, "жить, умирать, покупать или попробовать — многое решит ИИ".

Глава 6. Навыки, которые впечатляют

Учитывая, как быстро устаревают технологические навыки, запомните: вы упустите потрясающие возможности, которые GenAI и агенты могут дать вашему бизнесу, если не будете постоянно обучать

своих сотрудников. Ваши команды должны понимать, что может ИИ, чего он не может, на что обращать внимание и так далее. Нет, не нужно, чтобы у всех были дипломы по информатике, но единственный способ сделать ИИ доступным для многих — обучать многих. Ведь как вы можете ежедневно проходить мимо проблем, которые можно решить с помощью технологий, если даже не знаете, что они вообще решаемы?

Глава 7. Куда движется технология — эпоха единой модели никогда не наступит!

Запомните: одна модель не сможет управлять всем. У плотника в инструментальном поясе не один инструмент, а много. Более того, небольшие, гибкие модели показывают впечатляющие результаты и решают серьезные проблемы, с которыми сталкивается мир при использовании традиционных больших языковых моделей (LLM). В конечном счете, открытый доступ к множеству прозрачных и открытых моделей даст вам наибольшие шансы на успех.

Глава 8. Данные как фактор конкурентного преимущества

Название говорит само за себя: используйте свои данные для выделения среди конкурентов. Это связано с ролью создателя ценности с помощью ИИ. Если задуматься, то в распространенных LLM, которые вы, скорее всего, используете, содержится не более 1% корпоративных данных. Это значит, что здесь кроется огромный потенциал. Данные — как абонемент в спортзал: если не пользоваться, то и результата не будет, но и просто раздавать их нельзя.

Глава 9. Генеративные вычисления — новый стиль программирования

Взгляд в будущее, где генеративный ИИ займет свое место рядом с классическими и квантовыми вычислениями как новый строительный блок для приложений — генеративные вычисления. Это означает, что использование LLM будет опираться на методологии разработки программного обеспечения, что расширит их применимость, безопасность, масштабируемость, производительность и другие параметры. Мы также ожидаем (и это уже происходит), что наши LLM станут более рассудительными, будут тратить время там, где это необходимо, и давать более продуманные ответы. Это откроет новую область оптимизации и "магии" — время вывода. Такие изменения потребуют новых аппаратных решений и ускорителей, создавая новый стек вычислений, а возможно, и новый тип компьютера — генеративный.

Мы уверены, что с помощью этой книги, оглядываясь на этот исторический момент, вы сможете сделать это с удовлетворением — как человек, который увидел в данных настоящий ресурс и использовал GenAI и агентов как инструмент для создания ценности. Вот и все — начало вашего путешествия. Давайте начнем.

Для этой книги на [сайте](#) издательства создана веб-страница со списком опечаток, примерами и дополнительными материалами.

Благодарности

Конечно, у нас есть личный список людей, которых мы хотели бы поблагодарить, но такая книга не появилась бы на свет без помощи многих других людей. Этот раздел — для них, и мы надеемся никого не забыть.

Начнем с Линды Сноу. Линда прочитала каждую страницу этой книги несколько раз, задавала уточняющие и вдохновляющие вопросы, строго хмурилась при обнаружении грамматических ошибок и с ее вниманием к мельчайшим деталям (мы называем ее «королевой запяты») стала огромной помощью в работе над книгой. Искреннее спасибо, Линда!

Ребекка Рейес, руководитель направления развития навыков в IBM, разработала и совместно написала с нами главу 6. Она помогла создать «рецепт», который можно использовать для развития навыков в вашей организации. Как она часто говорит: «Одна из самых важных вещей, которыми может гордиться бизнес, — это культура обучения в организации». Она — невероятный ум, лидер и человек... Спасибо, Ребекка!

Майя Мурад и Анна Гутовска оказали большую помощь и с самого начала делились с нами своим увлечением ИИ-агентами — спасибо вам! Также благодарим Жоау (Джо) Моуру из CrewAI за интересные примеры использования и образцы кода.

Глава 7 получила значительную помощь от Михаила Юрочкина (и всей команды, работающей над маршрутизацией моделей). Спасибо за вашу прекрасную работу — для нас было честью рассказать о ней в этой главе.

В главе 8 мы благодарим всю команду InstructLab, а также Ким Мартино (IBM Research Comms), Джереми Эдера (Red Hat) и Съеду Амину Бегум (инженера по ИИ из Бангалора).

Дэвид Кокс сыграл важную роль в создании главы 9 — его прозрения о будущем генеративного ИИ и агентов практически не имеют равных. Он — прирожденный рассказчик.

Отдельное спасибо маркетинговой команде IBM, которая день за днем доносит ценность нашей работы. Мы не могли не упомянуть некоторых из них по именам, поскольку их идеи в той или иной форме (с изменением тона и стиля) вошли в эту книгу. Это Сара Беншаита, Сара Мерон и Ребекка Нойфельд — невероятно талантливые и проницательные люди, и нам повезло с ними работать. Также мы благодарны Линдси Лурье, Джереми Ходжу, Тиффани Пейдж, Саре Фельзенштейн и Стивену Миколайчаку (да, нам повезло знать и их тоже).

Конечно, мы не можем не поблагодарить нашу команду в O'Reilly: Шарон Кордесс, Лизу ЛаРью, Джона Хассэлла, Дэвида Михельсона и нашего редактора по производству Кристен Браун. Наш ведущий редактор, Гэри О'Брайен, не только удовольствие для работы, но и невероятно остроумен в комментариях! Также спасибо Кэрл Келлер и Дагу Макнейру, нашим корректорам.

Наконец, особая благодарность Дарью Гилу, бывшему старшему вице-президенту по исследованиям и разработкам IBM. Дарью начинал писать эту книгу вместе с нами, но затем ответил на призыв к государственной службе, когда 47-й президент США Дональд Трамп назначил его на должность заместителя министра энергетики по науке и инновациям. Это большая честь.

Персональные посвящения

Роб Томас

Говорят, что люди склонны переоценивать влияние технологий в краткосрочной перспективе и недооценивать — в долгосрочной. Наблюдая за развитием ИИ, я вижу его влияние на всех временных горизонтах. Эта книга — для тех, кто выбирает путь обучения и лидерства. И для тех, кто научил меня учиться и вести за собой — моих родителей, Кэрл и Дэвида Томас.

Пол Зикопулос

Я — патологический лжец. Я продолжаю всем говорить, что больше не буду писать книги. А потом пишу. Причина только одна: когда учишь — учишься сам. Этим я и занимаюсь. Каждый день: учишь, учишь, учишь, а потом учи, учи, учи. Эта страсть передалась мне от родителей, оба — педагоги до мозга костей. Спасибо им не только за то, что привили мне эту «болезнь», но и за то, что воспитали во мне стремление всегда поступать правильно.

Писать книги без жертв невозможно. Но, честно говоря, мое авторство — эгоистично, ведь основная тяжесть ложится на семью. Поэтому — Келли, спасибо, что ты всегда со мной. И Хлоя — моя девочка, которая уже стала взрослой (но для меня навсегда останется ребенком). Хлоя, куда бы ты ни шла, иди уверенно — ты прошла долгий путь не для того, чтобы остановиться. Если будешь помнить это, то какие бы ямы ни попадались на дороге или горы ни пришлось бы штурмовать, ты всегда будешь оглядываться назад с миром в сердце. Это мое самое большое пожелание для тебя (хотя ты все еще на моем обеспечении): обрети внутренний покой и гордость, потому что в тебе — невероятная доброта, и я хочу, чтобы мир это увидел.

Отдельное спасибо Лизе Бейкер — кажется, она единственный друг, который искренне интересуется тем, что я пишу и о чем говорю. Ты — блестящий ум, и наши дискуссии бесценны. Также спасибо Рею и ребятам из Pita Deli, где я писал часть этой книги. И Джине Ливи из Livy Method — ты вдохновляешь столько людей на лучшую жизнь; я рад нашей повторной связи и жду, как эта книга поможет тебе вывести бизнес на новый уровень!

Наконец, всем моим коллегам, которые заставляют меня чувствовать себя самым глупым в комнате (Крис Итон, Аяль Стейнбург, Ребекка Рейес, Мэдисон Гуч, Том Хронис, Крис Хаджилл, [доктор] Лаура Мусат, которая всегда на связи, и многим другим). Мне нравится это чувство. Оно вдохновляет. Именно поэтому я взялся за эту книгу вместе с соавторами (и Дарью Гилом, который помог нам начать). Вы все заставляете меня стремиться знать больше. И хотя я никогда не достигну своей цели,

окруженный такими людьми, как вы, этот «путь» и «виды» по дороге — просто потрясающие. Спасибо, что сделали эту книгу со мной.

Кейт Соул

Работать в IBM Research среди ученых — это как жить среди гигантов. У почти каждого коллеги есть PhD, на стенах — портреты нобелевских лауреатов, а в соседнем зале — квантовый компьютер, который твои коллеги не просто собрали, а изобрели с нуля. Всем сообществу IBM Research — спасибо. Все, что я знаю и попыталась передать в этой книге, я знаю только потому, что вы находили время объяснить мне свою работу, восполнить пробелы в моих знаниях и терпеливо отвечали на любые, даже самые глупые, вопросы. Надеюсь, эта книга достойно представляет ваш труд.

Особая благодарность моему руководителю Дэвиду Коксу — за лидерство и видение будущего генеративного ИИ, за уважение к моему мнению и за то, что подталкивал меня к глубокому погружению и лидерству. И главное — за трезвый ум, когда повседневность сходит с ума.

И наконец — моим соавторам, особенно Полу. Спасибо за поддержку и наставничество, твое терпение безгранично. Без тебя ничего бы не вышло. Спасибо моей потрясающей команде: Эйбу, Ализе, Дереку, Хуи и Радхе. Спасибо Джейкобу — тебе нужно посвятить отдельную книгу (я свою написала, теперь твоя очередь!). И моим родителям, Карен и Дэвиду — ваша любовь к учебе, писательству и инженерии заложила основу всему.

Глава 1. От +AI к AI+: генеративный ИИ и «момент Netscape»

Название этой главы может вас немного удивить - ведь Netscape появился в 1994 году, а мы пишем эту книгу в 2025! В этой главе мы объясним, почему сравниваем генеративный ИИ с «моментом Netscape», а затем кратко затронем широкий круг тем: от принципов работы технологии до важных аспектов, которые стоит учитывать при планировании вашего пути в мире GenAI.

Наши издатели советовали не делать первую главу слишком длинной, но мы пошли другим путем. Мы считаем, что если дадим вам достаточно пищи для размышлений, ценной информации и поделимся знаниями, которые ежедневно применяем в работе с клиентами, вы захотите узнать больше. Именно так мы и поступили - это как пилотный эпизод Netflix: он немного длиннее последующих, но зато оставшиеся эпизоды короче и более сфокусированы на конкретных темах.

Мы также сделали это, чтобы вам не обязательно было читать книгу последовательно. Например, мы упоминаем, что существует более миллиона больших языковых моделей (LLM), что подтверждает нашу уверенность: ни одна модель (как бы популярна она ни была) не станет универсальной. Это абсолютно исключено. Возможно, у вас много корпоративных данных, которые вы хотите использовать в ИИ, но не хотите ими делиться - тогда вы можете сразу перейти к главе 8, чтобы узнать о ландшафте LLM и безопасной работе с данными. Или, допустим, вы решите воспользоваться нашим опытом и узнаете, что компании с общеорганизационным планом повышения квалификации превосходят тех, кто доверяет это узкому кругу избранных специалистов. В этом случае вы перейдете к главе 6 и узнаете, какое отношение Леди Гага и королева Елизавета I (из 1500-х годов) имеют к плану обучения в вашей компании.

Мы уверены, что каждый из вас найдет в этой главе что-то полезное - будь то бизнес-идеи, принципы работы LLM и агентов, историческая перспектива, реалии будущего (например, агентный ИИ) или что-то еще. Итак, давайте начнем.

Что такое «момент Netscape»?

Почему мы называем текущий момент «моментом Netscape»? (Ах, Netscape — древний реликт интернета, столь же знакомый сегодняшней молодежи, как дисковые телефоны. Юные читатели... это был, по сути, первый в мире веб-браузер.) Вспомните, что произошло (если вы застали те времена), когда Netscape появился в 1994 году: интернет стал осязаемым и персонализированным — для всех. Поистине, интернет перестал быть привилегией избранных и стал доступен многим (хотя не все сразу этим воспользовались). Оглядываясь назад, очевидно, что демократизация интернета изменила наш мир... навсегда. Она изменила то, как мы храним данные, общаемся, покупаем, знакомимся и даже голосуем!

Мы называем это «моментом Netscape» (а такие моменты случаются редко), потому что сейчас мир начинает осознавать возможности ИИ, и это приведет к огромному количеству инноваций и идей. Но,

как и те, кто не воспользовался первым «моментом Netscape», те, кто не станет частью этой волны ИИ, окажутся на обочине. Это разделение затронет не только их способность соответствовать меняющимся нормам общества, но и ограничит доступ к важным услугам и возможностям. Этот «момент Netscape» развернется так же, как и первый. Те, у кого есть доступ к ИИ и кто использует его, переопределяют будущее (вспомните, что сделали с такси-индустрией те, кто воспользовался интернетом), а те, кто им не воспользовался, проиграют — с серьезными социальными и бизнес-последствиями. (Имена проигравших в первом «моменте Netscape» мы опустим, чтобы не смущать виновных.)

Сейчас многие говорят, что GenAI и рост ИИ-агентов могут привести к концу света. Что мы думаем? Технологии не обязательно должны уничтожать мир, чтобы его изменить. А если говорить о действительно опасных технологиях... у человечества уже есть история создания вещей, которые мы едва смогли контролировать и которые могли разрушить мир, но при этом помогли нам (например, ядерные технологии: медицина, энергетика и бомбы). Подробнее об этом — в главе 5.

Наконец, как и большинство великих вещей в жизни, удивительные изменения не происходят в одночасье. Это множество маленьких шагов, которые со временем создают момент, заметный всем. Текущий момент складывается из опыта, знаний, ошибок и прорывов, которые формировались десятилетиями (да что там — больше полувека!). Но не сомневайтесь: GenAI, особенно в форме агентов, изменит мир. Более того, мы считаем, что он станет настолько неотъемлемой частью всего, что мы делаем, что нам захочется вернуться в прошлое (ИИ этого не умеет, на случай, если вы задумались) и переопределить аббревиатуру ИИ как «фоновый интеллект» (ambient intelligence), а не «искусственный интеллект». Почему? Оглянитесь вокруг: вы читаете эту книгу, и, скорее всего, вокруг есть свет. Он не мешает, вы его не замечаете, он на заднем плане — но помогает вам читать. Это и есть «фоновое» присутствие. Именно таким, по нашему мнению, станет ИИ — незаметным помощником в повседневных бизнес-процессах. Поэтому мы и называем этот момент «моментом Netscape».

ИИ изменит мир. Но то, как именно он это сделает, зависит от нас — от всех.

ИИ и волшебный момент

Артур Кларк однажды сказал: "Любая достаточно развитая технология неотличима от магии". Возможно, когда вы впервые попробовали GenAI, у вас возникло именно такое ощущение. (Учитывая, что одному из наших отцов 89 лет, и он использует ChatGPT для написания блогов, а другой отец спросил у своего ребенка про DeepSeek, когда акции некоторых компаний упали после его выхода, мы предполагаем, что вы уже слышали о GenAI.)

Впервые в истории каждый, у кого есть интернет, получил в свои руки технологию, которая понимает его язык, выполняет запросы и создает совершенно новый контент. Сегодня ИИ даже может анализировать проблемы самостоятельно и предлагать решения — благодаря развитию моделей рассуждения, которые лежат в основе агентного ИИ, новой области GenAI.

ИИ пишет стихи, рисует фантастические изображения по нашим описаниям, документирует код, удивляет шутками или музыкальными композициями. Он способен творить — а творчество часто вызывает восхищение!

Поначалу трудно спорить с теми, кто считает ИИ магией — и компании (и люди) стремятся к этой магии со всех ног.

Не такие уж секретные агенты

Подробнее об ИИ-агентах мы поговорим в следующих главах. (Мы немного упрощаем, но когда вы слышите "агентный ИИ", считайте это синонимом "ИИ-агентов", работающих на основе GenAI.) Пока что представьте агента как программу, логика которой определяется и контролируется самим ИИ (LLM).

Сегодня большинство людей используют ИИ для решения конкретных задач (например, завершения фрагмента кода или суммирования документа), тогда как агенты ориентированы на цели. Вы даете агенту задачу, и он выполняет ее, даже планируя дальнейшие действия без вашего явного руководства.

Работа с агентами требует смены подхода: вместо проектирования приложения под конкретные задачи вы фокусируетесь на результатах. Агент сам определяет, какие шаги нужны для достижения цели. Вот некоторые примеры:

Блог и соцсети. Группа агентов (исследователь, автор блога, SMM-специалист) могут написать статью о влиянии инфляции на рынок жилья, создать посты для разных платформ (Instagram — с эмодзи, LinkedIn — профессионально, X/Twitter — кратко).

Улучшение NPS на 10 пунктов. Агентный ИИ разработает план, покажет обоснование и будет работать над достижением цели.

Самостоятельное обучение. Например, изучение греческого. Агент проведет тест (выяснит, знаете ли вы что-то кроме "Опа!" во время поджигания саганаки). Создаст персонализированную программу и будет мотивировать вас учиться, используя лучшие практики.

Можно ли сделать это с обычным LLM? Да, но неудобно: вам придется давать инструкции на каждом шаге. Агентный ИИ делает процесс автономным.

ИИ — это не магия, а инструмент, который становится "фоновым интеллектом" нашей жизни. И как любой инструмент, он требует правильного подхода.

Еще один пример агентного ИИ — это shopping agent. Компания Perplexity выпустила ИИ-агента для покупок, который может искать товары на сайтах и даже нажимать кнопку оформления заказа. Сравните это с вашим текущим процессом покупки на Amazon. Более того, Stripe (популярный платежный сервис) начал выпускать одноразовые дебетовые карты, чтобы такие агенты могли оплачивать товары без доступа к вашим банковским реквизитам (или чтобы жестко ограничить сумму расходов). Как вы можете догадаться, ИИ-агенты для покупок способны находить товары и выгодные предложения, которые вы либо не смогли бы найти сами, либо потратили бы на это много усилий.

Если вы ритейлер, вам это может не понравиться. Ведь что будет с импульсными покупками, на которые рассчитывают ночные покупатели-люди? Или с персонализацией на основе поведения? Агент же будет искать только то, что ему поручено — у него не будет момента "белки", когда он отвлекается на что-то новое, как это делают люди. Это постоянно происходит в реальном мире. Если у вас есть Costco, вспомните, когда вы в последний раз зашли за одним товаром и ушли только с ним? Никогда! Вы всегда уходите с чеком больше ста долларов, даже если зашли просто за хлебом. И скорее всего, еще и купили хот-дог с газировкой за \$1.50.

Поэтому другие подходы к агентам (например, Claude от Anthropic) учатся имитировать действия пользователя прямо в интерфейсе (управление рабочим столом). Такой подход не требует специального доступа к сайту (через бэкдор), который ритейлеры наверняка заблокируют, если агентный доступ не будет им выгоден.

Сейчас много внимания уделяется агентам и их фреймворкам — и спекуляциям на тему, что разработчики смогут с ними делать в будущем. Агенты представляют собой серьезный прорыв в компьютерных науках, так что эта область действительно перспективна. Хотя она еще молода, потенциал очевиден. Мы считаем, что агенты могут стать следующим шагом к росту производительности для бизнеса — и да, вы правы: им потребуются ограничители и система управления.

Но... ИИ — это не магия

Веками электричество считалось делом колдунов — фокусников, оставлявших публику в недоумении относительно его природы. Хотя Бенджамин Франклин в 1752 году доказал связь между электричеством и молнией, он с трудом представлял его практическое применение. Ирония в том, что его главное изобретение — громоотвод — служило как раз для избегания электричества, а не использования.

Сегодня многие воспринимают GenAI как очередную «магическую» технологию, применяя ее без понимания принципов работы. Либо считают ИИ чем-то исключительным, доверяя его лишь экспертам, которые должны поражать нас своими навыками. Такой подход окутывает ИИ мистикой и отдаляет его от обычных людей. Мы же убеждены: именно массовое понимание (включая риски) ИИ станет топливом для текущего «момента Netscape».

ИИ — это не волшебство

Доверьтесь нам. Мы играли в гольф клюшками, «спроектированными ИИ», и все, кроме одного, все равно отправляли мяч в кусты и промахивались с двух футов на ставку в \$1. Будь ИИ магией,

результаты были бы иными, а в лексиконе не появились бы греческие фразы, которые точно не переводятся как «Какой прекрасный день! Я отлично играю в гольф».

Как работает ИИ?

Все, что он делает — связывает точки данных (и делает выводы, часто без учета морали — об этом в главе 5). Как именно? С помощью математики и науки. Простейшее объяснение: ИИ пытается угадать число (вектор), представляющее объект, на основе предыдущих чисел (последовательностей векторов).

Например, распознавание изображений. ИИ анализирует три группы чисел (RGB) со значениями от 0 до 255. Аудиозапись 16 кГц — это 16 000 чисел в секунду (амплитуда, смещение). Фраза «Не плачь над пролитым молоком» для ИИ — просто набор чисел вроде {16357, 956, 16106...}. Он не понимает, что такое «молоко» или почему из-за него плачут. По сути, LLM (как ChatGPT или DeepSeek) — это «модель для угадывания чисел». Ученые называют их «моделями последовательностей чисел». Ключевая мысль: ИИ — мощный инструмент, но его «магия» сводится к вычислениям. И как с любым инструментом, важно понимать его ограничения.

Переход вашего бизнеса от +AI к AI+

Давайте сначала разберемся, почему этот момент так важен. Многие видят причину в удивительных возможностях ИИ. Да, мы наблюдаем невероятные вещи, но мы работаем с ИИ (и технологиями в целом) уже давно. Удивительное происходит постоянно.

Взгляните на телевизор на вашей стене. Он тоньше, чем три года назад? А чем пять лет назад? Сколько камер на вашем смартфоне сейчас — больше или меньше? Как насчет объема памяти на компьютере, количества стриминговых сервисов или процесса регистрации на рейс? Мы привыкли, что технологии постоянно улучшаются. Каждое новое поколение ИИ или других технологий становится все более впечатляющим.

Но главное в этом "моменте Netscape" — это суперспособности, которые дают нам промты. Промт — это то, что мы вводим в LLM. Промты буквально дали не-технарям сверхвозможности. Раньше повысить продуктивность могли только те, кто разобрался в коде и данных. Сегодня любой может общаться с ИИ так же естественно, как с коллегами.

Вопрос в том, воспользуетесь ли вы этим? Мы уверены, что вы хотите использовать эту уникальную бизнес-возможность (раз уж читаете эту книгу). Но более важный вопрос: сможете ли вы это сделать? Для этого вам нужны:

- Бизнес-фреймворк для реализации ваших идей (остаток главы поможет с этим),
- Понимание работы LLM,
- План обучения сотрудников,
- Знание о возможных рисках и преимуществах.

Все это есть в книге.

Представьте: вы сидите за столом и просите ИИ выполнить работу — открыть заказ на покупку, проанализировать прогноз продаж на квартал, закупить материалы для нового продукта или, как на рисунке 1-1, составить список новых вакансий для компании с 30 000 сотрудников. Звучит как магия? Но мы уже знаем, что это не так. Это технологии, доступные сегодня. Это GenAI в действии.

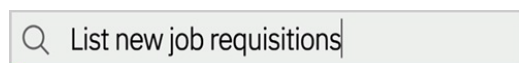


Рис. 1.1. Промты: даем ИИ в руки каждому

Прежде чем что-либо делать, смените ментальную модель с +AI на AI+

Что мы имеем в виду, когда спрашиваем: «Сможете ли вы этим воспользоваться?» Сегодня большинство компаний продолжают действовать в рамках своей традиционной ИТ-стратегии. Они говорят: «Давайте и ИИ тоже займемся!» — это и есть подход **+AI**, когда искусственный интеллект просто «добавляется» к существующим бизнес-процессам. (Спойлер: это не тот режим, в котором вы хотите оставаться.) И хотя за последние пять лет уровень внедрения ИИ удвоился, большинство организаций все еще мыслят в парадигме +AI.

Сейчас мы переходим от мира +AI к **AI+**, где ИИ становится первичным — подход **AI-first**. И уже через десятилетие те компании, которые начнут мыслить и действовать в логике AI+ — обучая сотрудников,

внедряя ИИ и технологии в производственные процессы — станут победителями нынешнего "момента Netscape", так же как ранние пользователи интернета стали лидерами после оригинального "момента Netscape". Поэтому мы говорим это прямо: если вы довольствуетесь устаревшей моделью +AI, для вашего бизнеса (и, возможно, лично для вас) это плохо закончится, потому что у вас не будет ни гибкости, ни компетенций, которые даст новое поколение ИИ.

Чтобы перейти к AI+, необходимо разложить бизнес-процессы на дискретные и оцифруемые компоненты, затем определить, где ИИ может взять работу на себя (чаще всего это рутинные задачи), и уже над этим выстраивать «человеческую» часть процесса.

Что такое базовая модель

О ChatGPT и DeepSeek сегодня слышали практически все — эти технологии вызвали огромный интерес и прочно вывели тему искусственного интеллекта в центр внимания. Но мы замечаем, что многие путаются в терминологии. Поэтому — чтобы вы могли блеснуть перед своим партнером (исправляя его, когда он неправ — ведь это всегда укрепляет отношения) — давайте ненадолго остановимся на этих терминах. Мы раскроем их подробнее позже, а пока важно понять: любая генеративная модель ИИ начинается с так называемой базовой модели (foundation model), и затем работает по принципу: промт → генерация → вывод.

ChatGPT, Granite, DeepSeek, Llama и другие — все это **LLM** (Large Language Models, большие языковые модели). Но LLM — это всего лишь частный случай базовой модели. Работы над такими моделями начались с фундаментальной научной статьи из Стэнфорда, опубликованной несколько лет назад.

Еще одна важная мысль: «язык» — это не только про речь или письмо. Мы используем языки в программировании, в химии для описания молекул, в коммуникациях. Если прищуриться, становится видно, что все можно выразить через язык: код, данные, молекулярные свойства — все. Так что не стоит думать, что LLM работают только с текстом — они куда универсальнее.

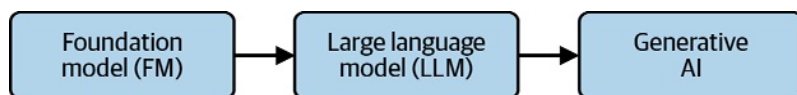


Рис. 1.1а. Иерархия понятий

Главное преимущество базовых моделей (и их производных — LLM) заключается в том, что их можно легко адаптировать для задач, для которых они изначально не создавались. Это не магия — просто математика. Мы поговорим об этом подробнее далее. В этой книге мы чаще всего будем использовать термин LLM, но помните о корнях — они растут из базовых моделей.

Искусственный интеллект как лестница, адаптированная для эпохи GenAI

Что все это значит с точки зрения внедрения генеративного ИИ и агентов в вашей организации? Несколько лет назад двое из нас написали книгу *The AI Ladder* (Лестница ИИ), где мы представили концепцию, изображенную на рис. 1.2. В той книге мы делали акцент на данных как пути к ИИ через информационную архитектуру (IA). На тот момент это было справедливое и точное представление того, как развивался ИИ: люди собирали, организовывали и размечали датасеты для обучения с учителем. (Ситуация кардинально изменилась после изобретения архитектуры трансформеров, которое многие считают точкой отсчета генеративного ИИ.) Мы считали, что «лестница ИИ» — отличная метафора: она включала не только технологический путь, но и поставщиков, с которыми вы работаете, и кадровые разрывы, которые необходимо преодолевать внутри компаний.

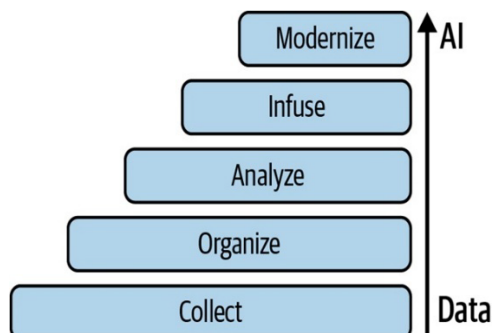


Рис. 1.2. Классическая (догенеративная) стратегия использования ИИ, позволявшая компаниям трансформироваться за счет подключения доверенных данных к ИИ

В то время основное внимание в AI Ladder уделялось работе с данными. И это было логично, потому что тогда (да и сейчас) организациям было сложно навести порядок в своих данных, и они просто хотели «добавить ИИ» в существующие процессы (мышление в стиле +AI).

Теперь сделаем шаг назад и подумаем о том, на какой основе стоит лестница, по которой вы собираетесь подниматься. Если основание прочное — вы чувствуете себя уверенно. А если кто-то (например, надежный технологический партнер) еще и держит эту лестницу — вы подниметесь выше, быстрее и с большим доверием. Эти принципы никуда не делись с приходом GenAI и агентов, и именно поэтому мы решили обновить AI Ladder под реалии сегодняшнего момента:

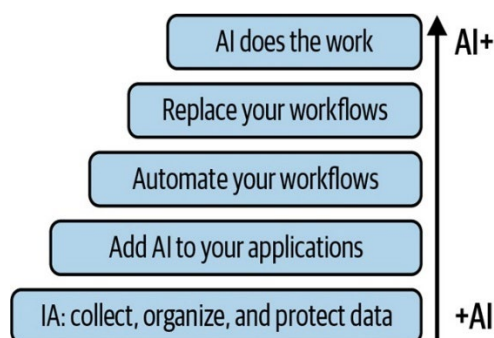


Рис. 1.3. Лестница ИИ, адаптированная под генеративный ИИ

Начиная с первой ступени, наша обновленная лестница ИИ построена с прицелом на ИИ с самого начала. Что это значит? Первая ступень все так же связана с данными. Мы также внесли важное уточнение, чтобы вы не упустили критически важный компонент успешной реализации GenAI: ваша способность собирать, организовывать, защищать и управлять данными (и не только ими) — это и есть информационная архитектура (IA). Причем она должна быть пронизана ИИ, чтобы быть действительно эффективной. IA очень важна для подхода AI+.

You can't have AI+ without an IA

Рис. 1.3а. Критическая важность информационной архитектуры

Что остается неизменным? Вам по-прежнему (а теперь даже больше, чем раньше) необходима платформа информационной архитектуры (IA), которая позволит собирать, организовывать, интегрировать, трансформировать, применять интеллектуальный анализ данных и хранить данные. Фактически, если такая платформа требовалась, чтобы стать компанией с подходом +AI, то для перехода к AI+ она необходима еще в большей степени.

Почему IA так важна? По мере прочтения этой книги вы все больше будете убеждаться в том, что для извлечения максимальной пользы из «момента Netscape» нужно уметь использовать данные для управления моделями. Чтобы делать это правильно (с хорошими результатами, но также прозрачно и с доверием), вы оцените, насколько IA ускорит ваши усилия, когда вы по-настоящему начнете использовать ИИ в бизнесе и поручите агентам выполнение рутинных задач от вашего имени.

Новые ступени помогают организациям научиться внедрять ИИ в приложения. Они подсказывают, как автоматизировать процессы и заменить существующие процессы на агентные. Ведь проблема в прошлом заключалась в том, что ИИ просто накладывали поверх существующих процессов (что логично в парадигме +AI). Но реальная ценность — та, которая заставляет вашего руководителя сказать: «Потрясающая работа!» — появится тогда, когда вы переосмыслите процессы заново и потом сами удивитесь: «Почему мы вообще делали это так раньше?» Именно это и есть AI+! И когда вы доберетесь до верхней ступени (а эта книга как раз для того, чтобы вам в этом помочь), вы будете передавать ИИ выполнение рутинной работы. Вы станете AI+. ИИ будет делать то, что ему действительно хорошо удается — рутину, и нам это очень нравится.

Мы рекомендуем вам внимательно изучить эту новую лестницу ИИ, и пока вы это делаете, подумайте о своих ключевых бизнес-процессах и о том, как вы работаете сейчас. Где вы на шкале от +AI до AI+? Вы вообще контролируете свои данные, не говоря уже о том, чтобы начать их эффективно использовать? Начали ли вы автоматизацию? Технологии уже есть — вы можете реализовать любую ступень этой обновленной лестницы ИИ, но воспользоваться ими должны вы сами.

Прежде чем начать путь, поймите на что тратятся деньги в чем будет польза

Прежде чем вы вообще начнете думать о проекте с использованием генеративного ИИ (или любом ИТ-проекте вообще), стоит прислушаться к дельному совету, который отлично выражен в цитате Томаса Эдисона: «Видение без реализации — это галлюцинация».

Точно так же, как (вы скоро это увидите) современные модели могут галлюцинировать и фантазировать, ваши планы по переходу от +AI к AI+ останутся фантазией без четко задокументированных приоритетов и сильных навыков исполнения. Поверьте — мы либо участвовали в таких инициативах (и у нас остались следы от этих «битв»), либо наблюдали, как масштабные идеи срывались с пика хайпа и рушились в пропасть. (Отличный пример — «озера данных» в эпоху Hadoop, у которых было нечто общее с незадачливым героем по имени Хампти Дампти.)

Наш совет? Неважно, с кем вы общаетесь — с вендором или коллегой — требуйте, чтобы любой предлагаемый проект был классифицирован по двум простым измерениям: на что тратятся деньги и в какой категории ИИ принесет пользу бизнесу.

Конечно, можно было бы рассмотреть множество факторов, но мы хотим дать вам инструмент, которым вы сможете действительно пользоваться по принципу KISS (keep it simple, silly — «делай проще, глупыш»; хотя «глупыш» иногда заменяют менее вежливым словом). Вам не нужна сложная формула. Наш (на удивление) простой совет почти гарантирует, что вы не попадете в ловушку «проекта на ИИ», в которую так часто попадают компании, когда технические команды ищут финансирование под последний модный тренд. Вам нужно концентрироваться на бизнес-ценности проекта, а не на его технологических деталях.

Измерение первое: тратить деньги, чтобы сэкономить, или тратить деньги, чтобы заработать?

Существует множество способов потратить бюджет, но по сути каждый день руководители вынуждены решать: **тратить деньги, чтобы сэкономить**, или **тратить деньги, чтобы заработать**. (С поправкой на отрасль — например, в здравоохранении это может быть «тратить деньги, чтобы спасти жизни» как отдельная категория расходов.)

Когда вы тратите деньги, чтобы сэкономить, вы занимаетесь **реновацией**. Когда вы тратите, чтобы заработать, вы занимаетесь **инновацией**.

Будет упущением не отметить, что некоторые кейсы одновременно и реновируют, и инновировать. Но помните про принцип KISS. Если ваш кейс делает и то и другое — разбейте его на этапы, как это делается в архитектуре микросервисов и современных приложений. Такие меньшие по объему, поэтапные компоненты позволяют достичь быстрых результатов и обеспечить фокус. Завершили фазу 1 (сэкономили деньги) — переходите к фазе 2 (заработать денег).

Например, одна компания, с которой мы разговаривали, перемещает 30 миллионов фунтов картофеля по своей цепочке создания стоимости и стремится остаться конкурентоспособной в мире автоматизации. На момент обращения примерно 50% их производственных мощностей были оснащены современными технологиями, а 50% — нет. Их серьезной проблемой стала вспышка картофельной парши, которая лишала часть урожая экспортного потенциала. Хотя парша не представляет угрозы для здоровья и безопасности пищи, она делает картофель нерыночным по стандартам сортировки (а в некоторых странах — и вовсе запрещает ввоз). Инвестиции в технологии, позволяющие выявлять и лучше очищать «пораженный» картофель (даже если звучит это хуже, чем есть на самом деле), позволили компании эффективнее перерабатывать урожай.

Эти меры привели к существенной экономии по сравнению с первоначальным планом расширить (и без того дефицитный) штат ручного контроля и мойки. Таким образом, используя бюджет по принципу «тратим, чтобы сэкономить» (на рабочей силе), компания получила дополнительный ресурс, который был направлен на развитие направления «тратим, чтобы заработать» — восстановление экспортных каналов благодаря достоверности ИИ-контроля и выход на вторичные рынки для другой части продукции. Дополнительный плюс — сокращение пищевых отходов, связанных с экспортными ограничениями (в условиях глобального дефицита продовольствия).

Оглянитесь и подумайте о тех инициативах, которые сейчас реализуются в вашей компании. О тех проектах, за которые вы лично отвечаете, или тех, которые вы хотите «продать» или получить поддержку. Все они укладываются в эту простую структуру.

Измерение второе: как ИИ помогает вашему бизнесу

Как только вы определились, на что направлен бюджет — реновация или инновация, — следующий шаг — классифицировать ваш проект по одной из трех категорий того, каким образом ИИ приносит пользу бизнесу: автоматизация, оптимизация, прогнозирование. Эта схема не идеальна, но она мощная и при этом достаточно простая. Вот как она работает на примерах.

Автоматизация (*тратим деньги, чтобы сэкономить*). «Я хочу использовать ИИ, чтобы обрабатывать обращения во внутреннюю тикет-систему техподдержки и автоматически перенаправлять их в нужный отдел — в зависимости от тематики и приоритета обращения, как это понимает ИИ.» Агентный ИИ может «прокачать» кейс — он может анализировать все тикеты, выявлять повторяющиеся проблемы, оценивать эффективность работы сотрудников, а затем автоматически формировать отчеты и планы действий.

Оптимизация (*тратим деньги, чтобы сэкономить*). «Я хочу использовать ИИ для сверхперсонализированной коммуникации в рамках кампаний продаж — не только по тексту и тону сообщений, но и по каналу доставки, времени отправки и т.д.» Агентный ИИ может дополнительно анализировать, какие каналы и временные окна лучше всего подходят для разных аудиторий — например, для пенсионеров, офисных сотрудников, учителей и даже для фанатов Тейлор Свифт.

Прогнозирование (*тратим деньги, чтобы заработать*). «Я хочу выполнять прогнозирование, чтобы понять, какие товары скоро закончатся, а какие не будут продаваться, используя данные POS-систем. И я хочу заранее уценить неликвид, чтобы не устраивать распродажи с большими скидками в конце сезона, когда товаров еще много.» Агентный ИИ может дополнить это, анализируя тренды с географической детализацией и даже погодные прогнозы, которые могут влиять на спрос на конкретные товарные категории.

Такой подход позволяет с самого начала понять, зачем вы внедряете ИИ, где он принесет выгоду, и какой тип пользы (снижение затрат, рост доходов, повышение эффективности) можно от него ожидать.

Используйте кривую компетентности, чтобы визуализировать, как ИИ помогает вашему бизнесу

Мы считаем, что отличной идеей будет визуализировать ваши решения по двум измерениям, потому что это поможет увидеть общую картину инвестиций в ИИ в масштабе всей компании и эффективно доносить идеи до других. Вы можете разработать собственную версию с элементами нужными вам, но на рис. 1.4 показан вариант, который нам нравится использовать для проектов, связанных с ИИ и данными.

Мы хотим, чтобы вы думали о данных с точки зрения вашей компетентности, то есть навыков, связанных с тем, как вы используете данные для того, чтобы ваш бизнес стал действительно data-driven. Почему? Потому что даже если у вас много данных (подсказка: у вас их действительно много), они бесполезны, если вы не умеете их применять. Это и есть суть компетентности, и приобретение этих навыков — одна из целей книги. Помните, что ландшафт в области данных постоянно меняется. Мы советуем воспринимать данные и работу с ними как абонемент в спортзал по двум причинам. Первая: если не пользоваться, не будет никакой пользы. Вторая: если прекратить пользоваться, утратите даже то, чего удалось достичь раньше.

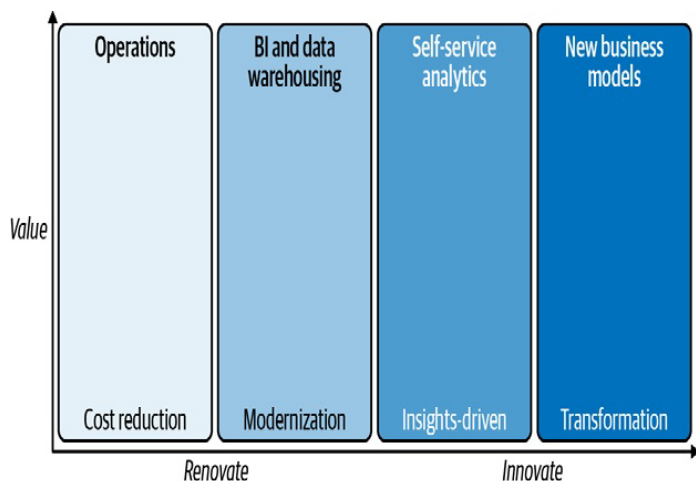


Рис. 1.4. Модель компетентности в области ИИ и данных

Модель компетентности стоит включать в любую стратегическую сессию по планированию ИИ-проектов. Мы разработали эту конкретную модель для аграрного клиента, когда судьба каким-то образом свела нас всех в одном проекте, и нам удалось отлично сработаться (если бы нет — скорее всего, вы бы не читали эту книгу). Ось Y — простая: она отражает создаваемую ценность.

По оси X мы отразили несколько измерений. Сначала мы разбили бюджетное пространство на четыре квадранта (их может быть сколько угодно, названия тоже можете подобрать свои), что дало нашему клиенту удобный визуальный инструмент и гибкость: десятки потенциальных кейсов использования ИИ, которые мы обсудили, можно было разместить прямо на доске. При этом вся стратегия по ИИ разрабатывалась в тесной привязке к бизнес-стратегии.

Внизу каждого квадранта мы обозначили, что именно делает инвестиция: снижает издержки, модернизирует бизнес, повышает его способность принимать решения на основе аналитики или же трансформирует сам бизнес. И, наконец, в самом низу вы можете увидеть естественную границу (она должна быть дружественной) между реновацией и инновацией. Это важный момент, потому что еще раз: стоит стремиться к тому, чтобы любые инвестиции в реновацию создавали задел для последующей инновации за счет освободившегося бюджета.

Когда все было собрано, получилось нечто вроде рис. 1.5. Он обеспечил команде по проекту полную ясность: что стоит делать, а что нет, и дал простой способ объяснить ожидаемую пользу от проектов заинтересованным сторонам.

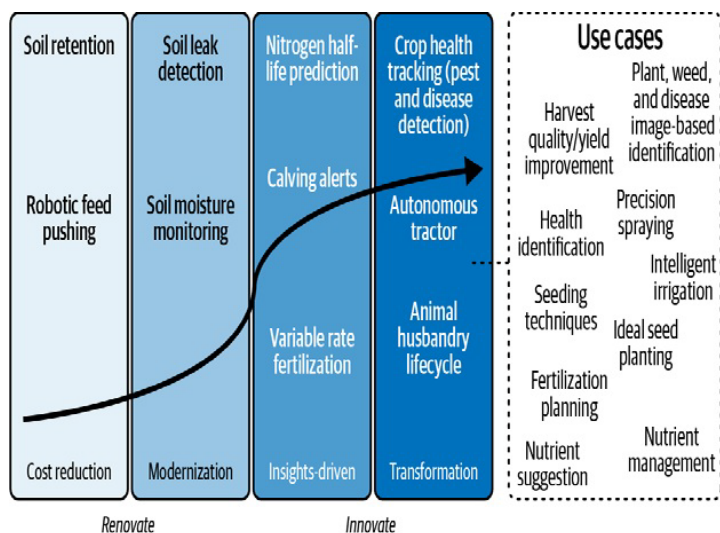


Рис. 1.5. Результаты построения модели компетентности по данным и ИИ на примере аграрного клиента

Прежде чем двигаться дальше, хотим поделиться еще несколькими мыслями о том, как могли бы выглядеть ваши бизнес-проекты, если перенести их на рис. 1.5. Во-первых, если вы хотите начать с «самого безопасного» места — это автоматизация и трата денег с целью сэкономить (снижение издержек). Но предупредим: если на этом все закончится, долгосрочная отдача от инвестиций будет невелика. Не поймите неправильно — разумная экономия — отличная стратегия, которая может освободить инвестиционный бюджет и даже поддержать рост выручки, но это не должно быть финалом вашего диалога об ИИ, иначе вы упустите правую часть этой модели. Мы говорим это, несмотря на то что многим по-прежнему приходится бюджетировать так, будто разговор действительно идет только об экономии. Поэтому даем обещание: в этой книге вы нигде не услышите от нас «делайте больше с меньшими ресурсами» — это мышление из начала 2000-х.

Возьмем, к примеру, управление. Большинство организаций старается реализовать требования регулирования с минимальными усилиями — «сделать, чтобы не оштрафовали». Их цели и ключевые результаты (OKR) — избежать штрафов (снижение издержек). Однако при этом упускается возможность превратить комплаенс в источник дивидендов — например, использовать полученные в процессе соблюдения регуляторных требований данные, чтобы объяснять выводы LLM для аудиторов. Мы видели слишком много проектов по управлению данными, которые в итоге обесценили свою пользу для бизнеса. А в мире GenAI и агентных систем это будет стоить дорого.

Во-вторых, вернитесь к модели на рис. 1.4. По мере продвижения вправо создаваемая ценность растёт. Отчасти это потому, что вы начинаете делать вещи «по-другому» (вы — AI+, а не +AI), но также и потому, что вы уже научились использовать ИИ для сокращения затрат и теперь занимаетесь модернизацией процессов. После COVID были компании, которые сумели перейти в правую часть, а были и те, кто остался в левой, низкодоходной части — и до сих пор там. Например, один американский магазин товаров для творчества внедрил систему самовывоза за считанные дни после закрытия офлайн-магазинов, тогда как известнейший крупный ритейлер до сих пор не смог эффективно связать запасы в физических магазинах с онлайн-заказами. Пандемия заставила компании провести 5–10 лет модернизации за 1 год — кто-то справился, кто-то нет.

Посмотрите на инновационные кейсы на рис. 1.5. Они потребовали от клиента переосмысления бизнес-моделей и рабочих процессов, чтобы по-настоящему воспользоваться всеми возможностями, которые открылись благодаря инвестициям в инновации. Но предупреждаем: вы упустите огромный пласт преимуществ, если сочтёте конечной точкой внедрение ИИ в рамках существующих моделей и процессов по схеме +AI. Нужно думать и планировать, как сами процессы могут (и должны) измениться благодаря GenAI и агентам, чтобы стать AI+. Это действительно критически важно. Когда будете строить собственные модели компетентности, пересматривайте процессы с учетом новых возможностей — потому что компании, работающие по модели AI+, обойдут тех, кто останется в модели +AI.

И, наконец, вы можете использовать ИИ в любом квадранте ценности на рис. 1.5. Да, по мере движения вправо кривая становится круче (больше отдача), но технологию можно применять где угодно. Например, агентный ИИ можно использовать как для снижения затрат, так и для трансформации. Вернемся к примеру с агентами: представьте, что вы создаете команду агентов, которые:

- анализируют данные из тикетов службы поддержки,
- формируют предложения по улучшениям,
- структурируют данные в таблицы с логичными группировками,
- визуализируют тренды с помощью графиков,
- составляют итоговый отчет с такими метриками, как количество обработанных тикетов на одного сотрудника, среднее время решения, удовлетворенность клиентов, проблемные зоны и многое другое.

С чего начать? Вот наш совет

Вопрос о том, с чего начать проект по внедрению генеративного ИИ, скоро станет актуальным для каждой компании. И хотя мы уже предложили вам инструменты, позволяющие увязать ИИ-инвестиции с бизнес-результатами, на этот конкретный вопрос все же нужно ответить отдельно. В конечном счете выбор конкретного кейса зависит от среды, в которой работает ваша компания, отраслевой специфики, приоритетов топ-менеджмента и стратегических целей. Тем не менее у нас есть рекомендация, которая нашла отклик у многих клиентов, только начинающих работу с GenAI.

Если вас беспокоят вопросы безопасного, выберите внутренний, низкорисковый кейс, ориентированный на автоматизацию и реализуемый по схеме *инвестируй, чтобы сэкономить*. Начните с GenAI, прежде чем подключать агентов — или с самого начала планируйте связку. Так вы получите ценный опыт работы со всеми компонентами технологии, включая агентов.

Обратимся к аналогии из американского бейсбола. (Мы хотели использовать пример из крикета, чтобы быть ближе к глобальной аудитории, но, признаемся, нам тяжело понять игру, которая длится пять дней и даже после этого не выявить победителя. Хотя, согласны — в игре, где есть позиция с названием Silly Mid, точно есть шарм.)

В бейсболе есть команды, которые делают ставку на *home runs* — эффектные дальние удары. Безусловно, они яркие, попадают в хайлайты спортивных новостей, но нет никаких доказательств, что ставка на *долгие мячи* гарантирует победу. А вот *одиночный выход на базу* (single) всегда дает шанс заработать очко. В итоге, *home runs* — это зрелищно, но без надежной подачи, защиты и системной работы команды они становятся пустыми калориями.

В мире ИИ — та же логика. Если вы только начинаете, и ставки высоки в случае неудачи, выберите внутренний кейс автоматизации. Просто *выйдите на первую базу*.

Пример? Автоматизация запроса отпусков. Вместо того чтобы заставлять сотрудников логиниться в перегруженную HR-систему, можно использовать чат-интерфейс, в котором достаточно ввести что-то вроде: «Хочу взять отпуск с 24 по 29 марта, по полдня». Или, если пойти дальше — система может учитывать школьные каникулы, корпоративные выходные и праздники и предлагать оптимальный график отпусков, позволяющий провести максимум времени с семьей при минимуме потраченных дней. А теперь подключите агентов — и ваш кейс выйдет на новый уровень. Что бы вы ни выбрали — это то, где ошибки легко исправить, и главное, они не окажутся в фокусе общественного внимания (в отличие от случаев, когда LLM-помощники публично генерировали странные ответы на жалобы клиентов, как это уже бывало у некоторых крупных компаний).

Станьте «шустрым» бизнесом: сначала сдвигайтесь влево, а потом сможете и вправо!

В разработке программного обеспечения и производстве широко применяется концепция shift left — «сдвиг влево». Смысл ее в том, что ошибки, выявленные на ранних этапах, обходятся значительно дешевле, чем те, что обнаруживаются уже «на выходе» — особенно когда продукт попадает в руки клиента.

Подумайте об этом: современные автомобили работают на примерно 100 миллионах строк кода. Для сравнения — у Boeing 787 Dreamliner их всего около 14 миллионов. (Да, нас это тоже удивило.) Очевидно, что физический дефект в автомобиле требует отзыва партии, но и ошибки в коде обходятся крайне дорого — особенно в автоиндустрии. Производитель должен устранить баг (выпустить патч), зачастую приглашая клиента в сервис, хотя тот может даже не подозревать, что в машине что-то не так. Да, есть возможность обновления «по воздуху» (OTA), но вместе с ней появляется огромная потенциальная уязвимость: через систему подключения авто можно организовать фишинговые атаки, в которых пользователя просят скачать якобы «исправление» — на деле, вредоносное ПО. Одним словом, это дорого.

И вот где в игру вступают генеративный ИИ и агентные технологии. Они позволяют бизнесу переосмыслить концепцию сдвига влево, сократить цепочку выполнения задач (делать правильно с первого раза, быстрее или вообще автоматизировать) и, соответственно, сжать издержки. Это делает компанию суперпродуктивной — а это, как вы узнаете в главе 3, станет ключевым ускорителем будущего роста и может раскрыть неиспользованный потенциал в вашей организации.

А как выглядят «сдвиги влево» в других ситуациях? Все сводится к одной идее: тратить деньги, чтобы сэкономить. В случае с автомобилем все просто — найдите баг до того, как он окажется внутри машины. Но мы предлагаем вам расширить представление о том, что значит «сдвинуться влево».

Каждый день мы проходим мимо проблем, которые можно решить или улучшить с помощью технологий

Это мантра, о которой мы хотим, чтобы вы начали думать. Именно поэтому так важно прочитать эту книгу: мы не просто рассказываем вам об ИИ, о подводных камнях, возможностях GenAI и агентов или о том, как все это работает. Мы даем вам стратегическое руководство для думающего человека о том, как внедрить ИИ в ваш бизнес — и это выходит далеко за рамки технологий.

Мы вооружаем вас искусством возможного и пониманием того, как GenAI (благодаря промтам) демократизировал взаимодействие с ИИ для каждого в вашей компании. Теперь у всех есть суперспособности для повышения продуктивности. Проще говоря, ваши сотрудники больше не будут проходить мимо проблем, которые можно решить с помощью технологий.

Наша команда авторов обладает огромным опытом в бизнесе и технологиях (это кодовая фраза для «Некоторые из нас уже давно в этом деле»). Мы могли бы написать целую книгу об искусстве возможного в решении проблем, и хотя здесь у нас нет места для этого, мы приведем несколько впечатляющих примеров из здравоохранения, которые изменят ваш взгляд на концепцию «сдвига влево» (shifting left) и заставят вас по-новому — в духе AI+ — оценивать каждый аспект вашего бизнеса.

Когда мы говорим о «сдвиге влево», мы имеем в виду сокращение затрат, уменьшение количества ошибок, снижение травматизма и повышение безопасности, предотвращение заболеваний и спасение жизней. Вот несколько ярких примеров такого подхода.

Личная мобильность: фундаментальное право человека

Согласно Статье 20 Конвенции ООН о правах инвалидов (CRPD), личная мобильность признана базовым правом человека. В США сегодня около 3,3 миллиона пользователей инвалидных колясок, 45% из которых — люди старше 65 лет. Из-за старения населения эта цифра скоро вырастет на 40%, добавляя примерно 2 миллиона новых пользователей ежегодно.

Исследования единодушно подтверждают: ограниченная мобильность снижает качество жизни, приводя к утрате чувства собственного достоинства, депрессии, множеству других негативных последствий. Механизированные коляски — мощный инструмент улучшения жизни, но и они не идеальны. Их вес достигает 160 кг, превращая их в «тараны» на колесах. 20% пользователей как минимум раз в год участвовали в столкновении, а 11% из них попали в больницу. Ущерб от таких инцидентов исчисляется миллионами долларов (разрушение имущества, травмы, страховые выплаты).

Где здесь «сдвиг влево»? В автомобилях уже давно используются системы предотвращения столкновений (CAS) и датчики — даже в бюджетных моделях. Почему бы не внедрить их в коляски? AI-решения могли бы оповещать пользователей через звук, свет или вибрацию, снижать количество травм, аварий и страховых случаев, повышать уверенность людей с ограниченной мобильностью. Компании вроде Braze Mobility уже работают над этим, оснащая коляски «автомобильными» технологиями безопасности. Результат: меньше госпитализаций, снижение ущерба для здоровья и имущества, рост самостоятельности пользователей.

Вывод: То, что многие воспринимают как «неизбежную проблему», на самом деле можно решить — стоит лишь перестать «проходить мимо» и применить технологии. Это и есть AI+ в действии — не просто улучшение, а переосмысление возможностей.

Сегодня США тратят примерно 327 миллиардов долларов в год на лечение диабета, и около трети этих расходов приходится на лечение диабетических язв стопы. У пациентов с диабетом такие язвы могут появиться из-за самых простых вещей: ударили палец и повредили кожу, наступил на что-то или натер мозоль при ходьбе. Для большинства людей подобные мелкие бытовые травмы заживают сами, но для диабетиков все иначе.

Эти язвы — осложнения диабета, которые сами пациенты часто не замечают. Почему? У диабетиков могут быть сопутствующие проблемы: от нейропатии (онемение в области стоп, из-за которого они не чувствуют, что что-то не так) до ретинопатии (ухудшение зрения, затрудняющее осмотр собственных ног).

Распространенность диабетических язв стопы связана с расовой и социально-экономической принадлежностью. Например, у чернокожих диабетиков риск их возникновения в три раза выше, а у бедных — на 93% выше. Это не медицинские, а социальные факторы, которые выходят за рамки нашей книги.

Что происходит, если такая язва появляется? Вероятность ампутации стопы увеличивается в 30 раз, а риск госпитализации (по любой причине) — в 3 раза. Фактически, диабетические язвы стопы — основная причина ампутаций в США, которые происходят каждые 3 минуты.

Где здесь «сдвиг влево»? В раннем обнаружении! Лечение диабетика с язвой стопы обходится системе примерно в 58 000 долларов в год, а без язвы — около 17 000 долларов. Как можно выявлять язвы на ранней стадии? С помощью термометрии — измерения температуры, которая, по сути, отражает уровень кинетической энергии частиц. Оказывается, воспаление предшествует язве, а с воспалением повышается температура (которую пациент может не чувствовать из-за нейропатии). Выявление «горячих точек» на стопах для создания базовых показателей — часть потенциального «сдвига влево» в диабетологии. Можно ли создать умные коврики для ванной, которые отслеживают тепловые сигнатуры и генерируют предупреждения, анализируя, что является нормой, а что — признаком возможного развития язвы?

Как мы уже говорили, каждый день мы сталкиваемся с проблемами, которые можно решить (или улучшить) с помощью технологий. Представьте, что могло бы сделать общество для диабетиков, если бы 70% разницы в затратах (между лечением пациента с язвой и без нее) направлялись не на устранение последствий, а на инновации — например, поиск лекарств и новых методов профилактики!

Предыдущие два примера не могут в полной мере передать, насколько трансформирующим может быть подход "сдвига влево" для вашего бизнеса. Но теперь вы наверняка видите, что понимание технологий позволяет совершенно по-новому взглянуть на свою деятельность.

Вот еще несколько примеров такого подхода.

Министерство по делам ветеранов США (VA) имеет шестимесячный бэклог необработанных заявлений. Автоматизация рутинных задач в этом процессе позволяет ускорить оказание помощи тем, кто служил стране.

Шведское правительство разрабатывает собственные публичные GenAI-модели для множества задач "сдвига влево", чтобы лучше обслуживать граждан. Представьте государственный чат-интерфейс, который понимает культурные особенности разных регионов Швеции.

Страховым компаниям требуются месяцы для обработки сложных или крупных исков. Автоматизация проверок и рутинных операций ускоряет выплаты и улучшает клиентский опыт.

Фармацевтическая компания Amgen тратила до 18 месяцев на поиск участников клинических испытаний, при этом 80% исследований не достигали плановых показателей. AI помог сократить сроки подбора, анализируя эффективность врачей и клиник по набору пациентов.

Компания Bayer с помощью AI сократила на 9 месяцев сроки набора участников для испытаний препарата Asundexian, экономя миллионы долларов и ускоряя вывод лекарства, способного снижать риск инсультов.

Процесс закупки даже такой простой вещи, как офисное кресло, в крупных корпорациях может занимать месяцы из-за бюрократии. Автоматизация этих процессов – классический пример "сдвига влево".

Хроническое одиночество становится новой медицинской проблемой в Северной Америке, приводя к депрессиям, алкоголизму и другим социальным последствиям. Ответственное использование AI для временного восполнения социальных контактов могло бы смягчить эту проблему.

В IBM система перевода сотрудников между отделами ранее требовала 20 минут ручного ввода и имела высокий процент ошибок. После автоматизации ошибки практически исчезли, а компания сэкономила 4000 часов рабочего времени. Всего подобные инициативы принесли IBM экономию в 3 миллиарда долларов.

Как видите, неважно, идет ли речь о переводе сотрудника между отделами или улучшении чьей-то жизни через вспомогательные технологии – везде можно найти возможности для "сдвига влево".

Очевидно, что предстоит огромный объем работы. Какая технология может помочь со всем этим справиться? Если оглядеться, ответ будет очевиден – GenAI и агенты.

Теперь о "сдвиге вправо". После того как вы частично "сдвинули влево" свой бизнес, сэкономив время, деньги и даже спасая жизни, у вас появляется уверенность и опыт для "сдвига вправо" — инвестирования в трансформационные изменения, которые приносят доход. "Сдвиг вправо" — это создание новых бизнес-моделей, но он также может стать решающим шагом, когда на кону стоит судьба компании или целой отрасли.

Мы видели, что происходит, когда компании не "сдвигаются вправо". Взгляните на Kodak, пионера в области фотографии. Несмотря на былой успех (компания была одной из самых процветающих в мире), ее история закончилась громким провалом. Почему? Kodak считала себя производителем пленки и стала жертвой узкого мышления, пытаясь защитить этот бизнес — хотя именно Kodak изобрела цифровую камеру еще в 1975 году. На самом деле компания занималась созданием воспоминаний, и когда способ их сохранения изменился, бизнес стал цифровым.

Иногда "сдвиг вправо" — это ваша стратегия, а иногда технологические изменения вынуждают к нему. Этот сдвиг преобразовал целую отрасль и породил новые бизнес-модели: фотографии стало проще делиться, их можно было делать больше без затрат на печать неудачных кадров, хранить тысячи снимков на ноутбуке вместо полок с альбомами. Компании использовали технологии, чтобы "сдвинуть" фотографию вправо: любители стали профессионалами благодаря редакторам, облачным сервисам и системам распознавания лиц. А Kodak... вы знаете, чем это закончилось.

Теперь вспомните Garmin, которая прославилась автомобильными GPS-навигаторами. Хотя этот продукт до сих пор существует (некоторые ценят его за надежность), компания осознала, что навигация изменилась — и что она занимается не автомобилями, а отслеживанием и картографией. Garmin "сдвинулась вправо", выйдя на рынки морской навигации, туризма и даже гольфа, предлагая статистику и рекомендации для игроков. Сообщества пользователей начали обмениваться маршрутами, улучшая карты через краудсорсинг. И теперь представьте, как GenAI и агенты могут продвинуть этот бизнес еще дальше.

Иногда "сдвиг вправо" — не просто переосмысление существующих технологий, а создание совершенно новых бизнес-моделей (как на крайней правой части рис. 1.4). Например, Airbnb появилась, когда ее основатели, неспособные оплатить аренду, заметили, что все отели Сан-Франциско забиты во время конференции. Они "сдвинули вправо" всю индустрию жилья!

Задумайтесь о ваших отношениях со страховой компанией — они построены почти что на конфронтации, не так ли? Вы оформляете полис, не подаете исков — а на следующий год взносы растут. Или попадаете в аварию, получаете выплату — и снова платите больше. Никакого партнерства. Но представьте, как GenAI и агенты могли бы "сдвинуть вправо" эту отрасль, превратив страхование в персонального помощника, который предупреждает риски, а не наказывает за них. Возможности безграничны — вопрос в том, увидите ли вы их раньше конкурентов.

Теперь рассмотрим андеррайтинг рисков на стройплощадке. Обычно работодатель стремится соблюдать нормы безопасности, чтобы избежать травм сотрудников. Страховая компания, принимающая на себя риски, тоже не хочет оплачивать лечение пострадавших и заинтересована в предотвращении несчастных случаев. GenAI и агенты могут полностью изменить подход к страхованию, трансформируя процесс оценки рисков на стройплощадке. Представьте систему, где AI отслеживает соблюдение мер безопасности: каски, спецодежда, перчатки, расстояние до электрооборудования и т.д. Камеры, установленные в разных зонах, используют AI-модели для выявления нарушений.

Например, в зоне сварки система контролирует противопожарные нормы, а в зоне работы крана — безопасность подъема балок. Компьютерное зрение фиксирует нарушения, а агенты автоматически обрабатывают видео, выделяют проблемы и формируют отчеты (ежечасно, ежедневно или еженедельно). Если происходит инцидент, страховой тариф может быть оперативно пересмотрен только для конкретной зоны, а не всей площадки. При этом заказчик получает возможность снизить премию, выполняя рекомендации из AI-отчетов, которые в реальном времени обновляют данные для страховщика. Такой подход открывает множество вариантов: тарификация по зонам, дням или месяцам, что представляет собой явный "сдвиг вправо" — переход к новой бизнес-модели и принципам взаимодействия между сторонами.

Советы по использованию базовых моделей и GenAI для вашего бизнеса

Мы долго спорили, куда поместить этот раздел — здесь, в главе 1, или в конце книги. В итоге решили дать эти советы сразу, чтобы вы получили наши ключевые рекомендации, даже если не станете читать дальше. А кто знает? Возможно, это даже побудит вас уделить время изучению всей книги, чтобы реализовать все наши идеи. В любом случае, если вы соотнесете свою бизнес-стратегию с этими рекомендациями, вы окажетесь в идеальной позиции для достижения потрясающих результатов с GenAI и агентами.

Совет 1: Действуйте безотлагательно

Это переломный момент в технологиях, поэтому будьте смелыми и используйте его! Вам необходимо разработать план обучения (глава 6), понять эту технологию и начать действовать. Именно для этого и написана эта книга — чтобы дать вам инструменты для быстрых, но продуманных действий, а не панической спешки. Это как тренировка пожарной эвакуации в отличие от реального пожара: быстрые, отработанные действия вместо хаотичной беготни.

Совет 2: Станьте создателем ценности с помощью ИИ, а не просто пользователем

Бизнесы должны включать в свою ИИ-стратегию планы по дообучению моделей на своих данных. Эти модели будут под их контролем — они станут их собственностью, возможно, самым ценным активом компании. Барри Меланкон (недавно ушедший в отставку CEO Ассоциации международных сертифицированных профессиональных бухгалтеров, наиболее влиятельной организации бухгалтеров с более чем 650 000 членов) сказал нам, что, по его мнению, наступит день, когда модели будут

учитываться как активы в финансовой отчетности. Мы подробно обсуждаем эту тему в главе 2, где вы узнаете разницу между создателем ценности с помощью ИИ и простым пользователем. Спойлер: не отдавайте свои данные и контроль на аутсорсинг и не сводите свою ИИ-стратегию к непрозрачным API-вызовам. Вам не стоит использовать чужую модель, если вы не знаете, как она обучалась, как управляется и как будут использоваться ваши данные.

Совет 3. Один алгоритм не будет править всем — делайте ставку на сообщество

Вам нужно сочетать передовые технологии с передовым сообществом. Открытое сообщество, развивающее ИИ (не путать с OpenAI), постоянно усиливает ту ценность, которую компании смогут извлекать из генеративного ИИ и агентных рабочих процессов. Ставьте на открытое ИИ-сообщество. (Подсказка: понять, действительно ли компания открытая, — это больше, чем просто посмотреть на ее название.)

Хотя многие слышали или пользовались только одной языковой моделью (например, ChatGPT), мы готовы поставить на карту свою репутацию и заявить: одна модель не сможет управлять всем, и ваша компания получит серьезные преимущества от новшеств и моделей, создаваемых в открытых сообществах. Прямо сейчас вы уже это наблюдаете: мощные open source-модели, такие как Granite, Llama и DeepSeek, постоянно вызывают резонанс. Именно поэтому мы советуем компаниям при формировании своей ИИ-платформы выбирать гибкие компоненты, чтобы можно было использовать как открытые, так и проприетарные модели.

Именно по этой причине компании вроде Meta и IBM выстраивают прочные партнерства с Hugging Face. Hugging Face — это социальная платформа мира ИИ. Она посвящена созданию и распространению сообщества ИИ-моделей. Мы любим думать о ней как о свахе между передовыми научными исследованиями в области ИИ и разработчиками, которые хотят создавать с их помощью крутые решения. У Hugging Face милый логотип, который вызывает теплые чувства, а за ним скрывается огромная база моделей, обучающих материалов и бенчмарков — все для того, чтобы сделать ИИ доступным каждому.

История Hugging Face действительно интересна и выходит далеко за рамки забавного названия и логотипа. Компания три года занималась разработкой разговорного ИИ, но, как это иногда бывает со стартапами, итоговая платформа и технологии оказались полезнее, чем сам конечный продукт.

Забавный факт: платформа Slack выросла из неудавшейся игровой компании Tiny Speck, разработавшей игру Glitch, в которой можно было жить внутри воображения древних гигантов. Основатель Tiny Speck, Стюарт Баттерфилд, сказал: «Если мы продолжим в том же духе, то просто сожжем все деньги через пару месяцев и ничего не добьемся. Но если остановимся сейчас — сможем использовать эти средства, чтобы создать что-то другое». Так появился коммуникационный инструмент, ставший Slack. Как и Hugging Face, Slack родился из совершенно другой идеи — и теперь посмотрите, как оба проекта изменили наш мир.

Когда Hugging Face начал публиковать части своей работы на GitHub, к нему начали подключаться участники open source-сообщества, а практики — делиться своими моделями.

Вы можете подумать: «Все говорят об открытых технологиях, созданных сообществом, но есть тысячи бесперспективных open source-проектов или тех, что просто дублируют друг друга». И вы будете правы. Мы часто сталкиваемся с клиентами, которые исследуют якобы открытые технологии, но выясняется, что 98% всего кода поддерживается и разрабатывается одним вендором — и это сводит на нет суть open source. С нашей точки зрения, настоящий open source — это когда сообщество действительно участвует и развивает продукт, а не просто доступ к коду открыт для просмотра.

Возможно, вам интересно, насколько велика вовлеченность и активность вокруг Hugging Face. Ответ прост: энергия — зашкаливает. Есть даже анекдот: как-то Клеман Деленг (один из основателей Hugging Face и по совместительству CEO) написал в Twitter приглашение на спонтанную встречу по теме GenAI в Сан-Франциско во время командировки — рассчитывал на 20 человек. Организаторам пришлось трижды менять площадку — в итоге пришло 5000 человек! Мероприятие прозвали «Вудстоком ИИ» — так что, как мы и говорили, энергия там действительно безумная.

Сегодня платформой Hugging Face пользуются десятки тысяч компаний, включая таких гигантов, как Google, IBM, Meta, MIT и Bloomberg, а также множество малых организаций. Вместе они делятся открытыми моделями, датасетами и демо. На момент финальной подготовки книги на платформе

было почти полтора миллиона моделей — и мы просто перестали считать. Так что мы подтверждаем: одна модель точно не будет править всеми (к слову, так называется и одна из глав — седьмая). И чтобы понять масштаб происходящего: только open source-модель Llama от Meta была скачана более 650 миллионов раз. Интерес и энергия вокруг открытого ИИ — более чем реальны.

Совет 4: Запускайте везде — эффективно

Мир уже прошел стадию Интернета вещей (IoT), где все было связано между собой, и вступает в эру Интернета всего, где все действительно взаимодействует со всем (у одного из авторов даже не получилось пропылесосить квартиру, потому что отключился Wi-Fi). Скоро это перерастет в «интеллект всего» — AI на границе (на edge-устройствах). Поэтому уже сейчас стоит оптимизировать проекты ИИ под производительность, задержки и стоимость, используя открытые гибридные технологии, а также обеспечив запуск моделей там, где это необходимо — будь то в пределах одного GPU, на устройстве без GPU вообще, или даже в «умном» коврик в ванной, измеряющем температуру ступней. То есть небольшие модели будут иметь все большее значение.

Совет 5: Действуйте ответственно, потому что доверие — это главный допуск к рынку

И это, пожалуй, самый важный совет. Все вышеописанное теряет смысл, если ИИ не разрабатывается ответственно и прозрачно, с надлежащим управлением на всех этапах жизненного цикла. Бизнесу следует постоянно управлять своими данными (а значит — иметь архитектуру данных, IA) и развивать решения совместно с проверенными партнерами. Доверие — это и есть ваша лицензия на деятельность. В мире данных легко найти как примеры добросовестных и недобросовестных участников, так и тех, кто проявил активную позицию или остался в стороне. Именно поэтому так важно сотрудничать с партнерами, которым вы полностью доверяете. Потеря доверия происходит ведрами, а восстановление — каплями. Ищите тех, у кого ведро уже полное — как в их действиях, так и в продуктах и подходах к построению LLM: все должно начинаться с доверия и прозрачности.

Следите за тем, чтобы ваш «резервуар доверия» всегда был наполнен. Это значит, что принципы справедливости, интерпретируемости, прозрачности и этики должны быть не запоздалыми мыслями, а изначальными требованиями. Это разница между подготовленной тренировкой по пожарной безопасности и настоящим пожаром, к которому вы совершенно не готовы.

А теперь — сосредоточимся на части, связанной с ИИ

В этой главе мы рассмотрели множество аспектов, касающихся бизнес-стороны искусственного интеллекта. Мы говорили о переходе от +AI к AI+, и мы уверены, что через год каждый, кто прочитал эту книгу, уже будет иметь опыт работы с генеративным ИИ. Вопрос лишь в том, будете ли вы (и компании, с которыми вы связаны) двигаться быстрее своих конкурентов. Подниметесь ли вы по нашей обновленной лестнице ИИ и начнете ли использовать архитектуру информации (IA) и искусственный интеллект (AI) для трансформации в полноценный «AI+ бизнес»?

Мы надеемся, что приведенные примеры вдохновили вас на размышления о том, что возможно (еще больше примеров вы найдете в главе 4). И мы надеемся, что вы воспользуетесь возможностью поэкспериментировать с технологиями, продолжите вдохновляться их потенциалом и перестанете проходить мимо проблем, которые можно решить с помощью ИИ.