

95% AI-пилотов дают ноль. Что общего у выживших 5%

2026-06-01

95% AI-пилотов дают ноль. Что общего у выживших 5%

В августе 2025 года исследовательская инициатива MIT NANDA опубликовала отчёт «The GenAI Divide: State of AI in Business 2025», и одна цифра из него за неделю обошла все деловые издания: несмотря на 30–40 млрд долларов, влиятельными корпорациями в генеративный AI, 95% организаций не получили от своих пилотов измеримого возврата (отчёт MIT NANDA, Fortune). Цифра прозвучала достаточно громко, чтобы на неё отреагировал рынок: Axios прямо связал августовскую коррекцию AI-акций с этим исследованием — инвесторы впервые увидели подтверждённый разрыв между расходами на AI и отдачей (Axios). Дальше начинается интересное: вывод, который сделали из этой цифры почти все, был неверным.

Стандартное прочтение звучит так: модели пока сырые, бюджеты маленькие, рынок незрелый — подождём год, всё наладится. Это удобная версия, потому что она ни к чему не обязывает. Отчёт NANDA говорит другое. Провал не в моделях и не в деньгах — в том, **куда** именно компания ставит AI. И у тех самых 5%, которые получили отдачу, есть общий структурный признак — он лежит не в технологии, а в процессе.

Почему дело не в модели, а в том, куда её поставили?

Сначала уберём с дороги две привычные объяснительные версии, потому что данные их не подтверждают.

Версия «модели слабые» не выдерживает простой проверки: те же самые модели от тех же поставщиков одновременно дают и провальные, и успешные внедрения. NANDA фиксирует, что разрыв проходит не по качеству модели, а по способу её встраивания: пилоты, привязанные к конкретной рабочей операции с обратной связью, переживают полгодакратно чаще, чем универсальные ассистенты «для всех сразу». Сам факт, что 95% не окупилась при 30–40 млрд расходов, исключает версию «мало вложили» — деньги были, отдачи не было. Любопытно, что инструменты для личной продуктивности при этом приживаются: сотрудники охотно пользуются универсальными чат-ассистентами в личном контуре. Но именно поэтому корпоративный пилот и проваливается — компания платит за систему, которую сотрудник дублирует бесплатным инструментом, не получая взамен ничего, что меняло бы сам процесс.

Версия «рынок незрелый, подождём» опровергается поведением самих компаний. По данным S&P Global Market Intelligence, к середине 2025 года 42%

компаний свернули или существенно сократили большинство своих AI-инициатив — против 17% годом ранее (S&P Global Market Intelligence). Это не «ждут зрелости» — это откатывают то, что уже запустили и что не сработало. Рост отказов в 2,5 раза за год — сигнал не о том, что AI рано, а о том, что большинство строит его неправильно.

Что же объединяет выживших? Один признак: у успешного пилота AI вшит в **один конкретный доходный процесс** — узкую операцию с измеримым входом, измеримым выходом и понятной ценой ошибки. У провального — AI размазан горизонтально: «дадим всем сотрудникам ассистента и посмотрим». Первое окупается, потому что отдачу можно посчитать в рублях на одной операции; второе — нет, потому что польза рассеяна по тысяче рабочих мест и не собирается ни в одну метрику, которую увидит финансовый директор.

Купить вертикальное решение надёжнее, чем построить своё

Второй сильный вывод NANDA касается развилки, на которой стоит почти каждая компания: купить готовое отраслевое решение у поставщика или построить своё силами внутренней команды. Данные отвечают однозначно. Покупка вертикального решения у внешнего поставщика приводит к успешному внедрению заметно чаще, чем внутренняя разработка — соотношение в отчёте составляет порядка двух к одному в пользу покупки.

Внутренняя команда строит AI как технологический проект: выбирает модель, поднимает инфраструктуру, пишет интеграции — и упирается в то, что у неё нет накопленного опыта именно этой операции в десятках других компаний. Внешний поставщик вертикального решения приходит не с моделью, а с **готовой обвязкой** (agentic harness) вокруг неё: размеченными сценариями, типовыми исключениями, кодифицированными регламентами отрасли, которые он собрал, обслуживая похожие процессы у других клиентов. Модель внутри у всех примерно одна. Разница — в том, что построено вокруг неё.

То, что создаёт ценность, — это не сама большая языковая модель, а **детерминированная оболочка** вокруг неё: слой, который превращает вероятностный текст модели в предсказуемое действие в рамках конкретной операции. Поставщики фронтальных моделей это понимают и сами двигают границу: Anthropic в работе «Building effective agents» описывает, что большинство успешных продакшн-агентов — это не «умная модель в свободном полёте», а композиция простых, проверяемых шагов с жёсткими ограничениями. AWS вынес эту логику в отдельный продукт. AWS Bedrock AgentCore — это управляемая среда исполнения для AI-агентов с контролем памяти, инструментов и границ действий (Bedrock AgentCore). Сама эта обвязка стремительно становится массовой и доступной. А значит, защита внедрения уезжает не в модель и не в стандартную оболочку, а в то, насколько точно оболочка пригнана к одной конкретной операции и насколько долго она накапливала следы реальных решений.

Это объясняет, почему «купить» обыгрывает «построить»: поставщик продаёт не технологию, а накопленную точность пригонки. Внутренняя команда начинает её копить с нуля — и в большинстве случаев не доживает до того момента, когда точность станет достаточной для окупаемости. Оговорка важна: вывод не означает «никогда не стройте». Он означает, что планка для решения «строим сами» выше, чем кажется на старте, и что инженерный бюджет в случае внутренней разработки должен уходить в обвязку, а не в бесконечное сравнение моделей.

Правило 10/20/70: деньги уходят не туда, куда смотрят

Третий признак выживших — структура бюджета. И здесь сходятся независимые наблюдения консультантов: успешные внедрения подчиняются правилу, которое короче всего формулируется как **10/20/70**. Десять процентов усилий и денег уходит на саму модель и алгоритмы, двадцать — на данные и инфраструктуру, и семьдесят — на людей, процессы и изменение способа работы. BCG, разбирая, где именно в AI-проектах возникает ценность, приходит к той же пропорции: технология — это меньшая часть работы, а основная масса усилий лежит в перестройке процессов и управлении изменениями (BCG).

Провальный пилот переворачивает эту пропорцию. Семьдесят процентов внимания уходит на выбор модели и сравнение поставщиков, двадцать — на инфраструктуру, и почти ничего — на то, чтобы перестроить процесс вокруг новой возможности и довести людей до того, чтобы они этим пользовались. Получается технически работающий пилот, которым никто не пользуется, потому что он навешен поверх старого процесса, а не встроен в новый. NANDA называет это разрывом обучения: системы не запоминают контекст, не адаптируются к рабочему процессу и не улучшаются от использования — и сотрудники тихо возвращаются к привычным инструментам.

Публичная история Klarna — иллюстрация обеих ошибок сразу. Компания громко заявила о замене 700 сотрудников поддержки AI-ассистентом (OpenAI), но через полтора года начала нанимать людей обратно — на верхний слой, к сложным и эмоционально нагруженным случаям. Внедрение само по себе работало; провалилась попытка горизонтально вычистить целый слой людей вместо того, чтобы перерисовать границу эскалации. Семьдесят процентов работы — про людей и процессы — были недоделаны, и реальность это вернула.

Конкретные тесты для трёх ролей

Из трёх признаков выживших следует три разных набора проверок — для трёх разных людей, принимающих решение.

Собственнику нетехнологичного B2B. Не спрашивайте «какой AI нам внедрить». Спросите иначе: какая одна операция в бизнесе имеет измеримый вход, измеримый выход и цену ошибки, которую владелец держит в голове без таблицы. Если такой операции не нашлось — пилот рано запускать, какая бы модель ни стояла внутри. Если нашлась — требуйте, чтобы отдача считалась

именно на ней, в рублях, а не «в сэкономленных часах вообще». Один практический признак готовности: если владелец может назвать сумму, которую теряет на этой операции каждый месяц, проект имеет шанс окупиться; если сумма не называется, считать отдачу будет не от чего.

Директору по операциям или цифре. Проверьте структуру плана внедрения на соответствие правилу 10/20/70. Если в смете 70% — это лицензии, модель и инфраструктура, а строчка «перестройка процесса и работа с людьми» либо отсутствует, либо стоит последней по бюджету — план перевернут, и его лучше переписать до старта, а не после отката. Второй тест — на буквальную обратную связь: умеет ли система запоминать исправления, которые вносит человек, и становиться от них точнее; если нет, в первый же месяц упрётся в потолок.

Тех-лиду на развилке «купить или построить». По умолчанию считайте, что покупка вертикального решения вероятнее доведёт до окупаемости, и стройте своё, только если у вас есть честный ответ на вопрос: «что в нашей операции настолько уникально, что готового поставщика для неё не существует». Если же строите — тратьте инженерное время не на саму модель (она у всех одинаковая), а на детерминированную оболочку: формализацию сущностей, разметку исключений, накопление следов решений. Это и есть актив, который через год отделит вас от того, кто «просто запускает» модель поверх процесса.

Сигналы, по которым видно разворот

За чем стоит следить в 2026 году, чтобы понять, что разрыв между 5% и 95% начинает закрываться. Первый сигнал — снижение доли свёрнутых проектов: если показатель S&P Global в районе 42% начнёт падать, это будет означать, что рынок усвоил урок про вертикальную привязку. Если он продолжит расти — компании по-прежнему ставят AI горизонтально и продолжают платить за это откатами.

Второй сигнал — смещение языка вендорских предложений с «нашей модели» на «нашу пригонку к процессу». Когда поставщики перестанут продавать характеристики модели и начнут продавать накопленную точность под конкретную операцию, это будет означать, что граница ценности окончательно переехала из модели в оболочку. Третий — появление в корпоративных бюджетах отдельной защищённой строки на перестройку процессов и работу с людьми, сопоставимой по размеру с расходами на технологию: пока её нет, правило 10/20/70 будет нарушаться, а 95% — оставаться 95%.

Главное

- Провал 95% AI-пилотов — не про слабые модели и не про маленькие бюджеты: при 30–40 млрд долларов вложений отдачи всё равно нет, значит, причина структурная.
- Выживших 5% объединяет один признак: AI вшит в один измеримый доходный процесс, а не размазан горизонтально по всей компании.

- Покупка готового вертикального решения доводит до окупаемости заметно чаще внутренней разработки — потому что поставщик продаёт не модель, а накопленную точность пригонки к процессу.
- Успешные бюджеты подчиняются правилу 10/20/70: 10% на модель, 20% на инфраструктуру, 70% на людей и перестройку процессов. Провальные переворачивают эту пропорцию.
- Ценность создаёт не умная модель, а детерминированная оболочка вокруг неё; модель у всех примерно одинаковая, разница — в том, что построено вокруг.

FAQ

Почему именно 95%, а не другая цифра? Это показатель из отчёта MIT NANDA «The GenAI Divide: State of AI in Business 2025»: 95% обследованных организаций не получили измеримого возврата от своих генеративных AI-пилотов в течение полугода. Цифра подтверждена реакцией рынка и независимыми обзорами деловой прессы.

Значит ли это, что AI в бизнесе не работает? Нет. Работают те 5%, у которых внедрение привязано к одной конкретной доходной операции. Вывод не «AI не работает», а «горизонтальное размазывание AI не работает, вертикальная привязка работает».

Если покупать надёжнее, чем строить, зачем вообще строить своё? Строить стоит только там, где операция действительно уникальна и готового поставщика не существует, и где инженерное время вкладывается в оболочку вокруг модели, а не в саму модель. В большинстве случаев операция типовая, и покупка быстрее доводит до окупаемости.

Что такое правило 10/20/70 на практике? Это пропорция распределения усилий в успешном AI-проекте: примерно 10% — модель и алгоритмы, 20% — данные и инфраструктура, 70% — люди, процессы и управление изменениями. Если в плане внедрения технология занимает большую часть бюджета, а перестройка процессов — меньшую, план почти наверняка приведёт в те самые 95%.