

## Кейсы AI-агентов измеряют сэкономленные часы. Маржа уходит выше

2026-05-08

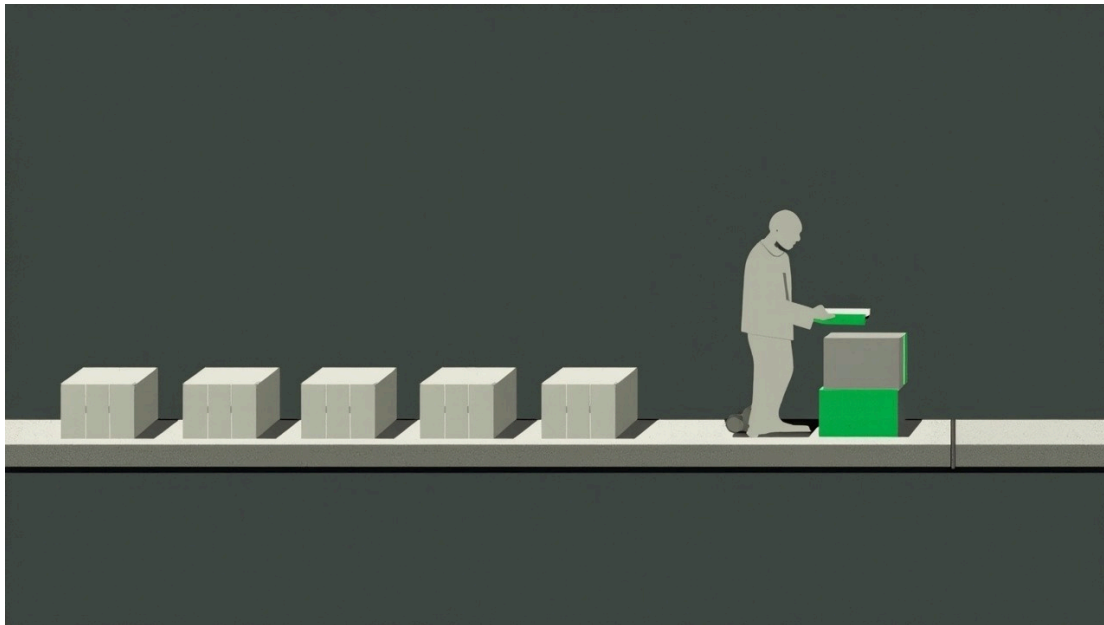


Рис. 1. Сборочная линия с одинаковыми силуэтами и одной выделенной фигурой — метафора уходящей маржи

### Кейсы AI-агентов измеряют сэкономленные часы. Маржа уходит выше

**TL;DR.** Публичные ROI-кейсы AI-агентов в 2026 году измеряют сэкономленные часы и сокращение цикла. Это потолок горизонтальной автоматизации. Реальная защита и маржа уходят в операционный слой бизнеса — онтологию предметной области, исполнимую политику решений и следы решений, которые организация накапливает изнутри.

#### Что показывают публичные ROI-кейсы и что они скрывают?

За последние девять месяцев индустрия получила первый качественный набор измеримых результатов от платформ для агентной автоматизации. На странице клиентов Activepieces Alan заявляет о более чем 6 300 сохранённых часах в год и трёхстах с лишним рабочих сценариях; Funding Societies — о ста с лишним сценариях в восьми департаментах. На странице кейсов n8n Flow AI рапортует о сокращении исходящих касаний с 3–5 часов до менее минуты, Field Aerospace переводит подготовку коммерческого предложения из двух недель работы команды в 25 минут, Formula Bot уменьшает срок добавления

нового коннектора с недели до полутора дней. Это не маркетинговые лозунги — реальные операторы согласились публиковать цифры с именами и должностями.

Если смотреть на эти кейсы не отдельно, а как на карту рынка, проступает другой узор. Все они решают одну задачу — ускоряют существующие горизонтальные процессы. И все выстраиваются в один и тот же ценовой и архитектурный класс. Потолок этого класса виден.

## **Где заканчивается стандартная среда исполнения и начинается отраслевой продукт**

Различим три рабочих понятия, без которых разговор про слои размывается.

Стандартная среда исполнения агента — слой с унифицированными интерфейсами: цикл сессии, маршрутизация инструментов, память, изолированная песочница, разворачивание. К маю 2026 года он перестал быть инженерным проектом и стал готовым продуктом у всех крупных провайдеров. Anthropic в инженерном посте про управляемые агенты формулирует это прямо: предположения, зашитые в обвязку агента, устаревают вместе с моделями, и поэтому компания упаковывает оркестрацию в стабильные интерфейсы поверх сменяемого исполнения. The New Stack в разборе релиза описывает сдвиг от модели к модели плюс готовая оркестрация как продукту провайдера.

Отраслевой SaaS-продукт — другой полюс: продукт со своими экранами, отчётами, ролями, встроенными требованиями регулятора и многолетним операционным контекстом. Veeva для фармы, Procore для стройки, ServiceTitan для сервисного бизнеса. Veeva в продуктовой развёртке Vault AI описывает, как встраивает агентные сценарии прямо в отраслевые модули — это сигнал, что вертикальные игроки видят угрозу снизу и начинают двигаться навстречу.

Между этими двумя полюсами есть промежуточный слой, который в публичных кейсах почти не виден. Это уровень, на котором агентная система перестаёт быть «лучше Zapier» и становится способом, которым организация фактически работает: операционный слой бизнеса (operating layer). На карте маржи именно туда смещается ценность по мере того, как стандартная среда становится массовой.

## **Что Alan, Flow AI и Field Aerospace на самом деле сделали**

Если разобрать публичные кейсы по слоям, получается следующая картина.

Alan развернул на Activepieces платформу для внутреннего обеспечения сотрудников. Цифры впечатляют, но вся работа происходит в горизонтальной плоскости: автоматизация HR-онбординга, синхронизация задач между корпоративными инструментами, эскалации в чатах. Это не отраслевая модель медицинского страхования — это родовая сантехника, которая ускоряет существующие процессы, но не переописывает их. Сама компания подчёркивает логику обеспечения: «300+ сценариев, 200+ внутренних авторов» означает,

что цель была дать сотрудникам инструмент, а не построить специализированную операционную систему страхового продукта.

Flow AI — стартап в дистрибуции страховых продуктов. Их голосовые исходящие касания действительно дают радикальное сжатие времени: 3–5 часов превращаются в 60 секунд. Но если посмотреть, на чём построены касания, это связка из ElevenLabs, веб-хуков, Postgres, n8n и SMS/email-провайдеров. Качественная горизонтальная сборка. Слой, на котором Flow AI становится незаменим для брокеров — отраслевая модель оценки рисков, история взаимодействия с конкретными типами клиентов, политики комплаенса для разных юрисдикций — в публичной презентации не виден. Не потому что его нет, а потому что не он там измеряется.

Field Aerospace сократил подготовку коммерческого предложения с двух недель до 25 минут и снял около 30 000 долларов годовых затрат на программное обеспечение. Цифры серьёзные. Но это автоматизация документального процесса — извлечение, форматирование, сборка. Под этим слоем лежит специфика аэрокосмических контрактов: регуляторные требования, история ценообразования, специфическая структура спецификаций, партнёрские договорённости. Эта специфика в кейсе не видна. Видна только верхняя горизонтальная аппликация.

Formula Bot ускорил добавление нового коннектора с недели до полутора дней. Это операционная победа платформенной команды, не отраслевая защита. Если завтра OpenAI Connector Registry выпустит официальный SAP-коннектор с лучшей надёжностью, ценность собственной интеграционной полки начнёт оседать.

Все четыре кейса — честная работа с измеримым результатом. Но они отвечают на один вопрос: как ускорить существующие процессы. И не отвечают на другой: что становится защищённым активом за горизонтом 18 месяцев, когда стандартная среда исполнения станет массовой.

### **Три уровня защиты и куда смещается маржа**

Полезно ввести явное разделение слоёв. На каждом из них живёт разный класс продукта и разный класс конкуренции.

Слой	Что это	Горизонт защиты	Кто здесь играет
Стандартная среда исполнения	цикл сессии, маршрутизация инструментов, память, песочница	6–18 месяцев до выхода управляемого аналога	n8n, Activepieces, OpenHands, Anthropic Managed Agents, AWS AgentCore
Операционный слой	онтология предметной области, политика принятия решений, следы решений конкретной организации	годы; невозможно воссоздать ретроспективно	специализированные интеграторы и вертикальные AI-продукты
Отраслевой SaaS	полный продукт отрасли с продавцами, комплаенсом, экосистемой	десятилетия	Veeva, Procore, ServiceTitan и их аналоги

Публичные кейсы живут в первом слое. Их экономия впечатляет, потому что сравнение идёт с ручным трудом и устаревшими инструментами; в этой системе координат любой грамотно собранный сценарий выглядит как прорыв. Но горизонт жизни этого преимущества короткий. Anthropic уже явно пишет, что предположения обвязки устаревают вместе с моделями, и сама же выпускает управляемую альтернативу. n8n и Activepieces встраивают MCP-серверы и конструктор агентов в ядро. То, что год назад требовало месяца сборки, через год будет ставиться из конфига.

Отраслевой SaaS — третий слой, куда нельзя зайти из горизонтального проекта без отраслевой команды, многолетнего опыта продаж и собственного продуктового цикла.

Операционный слой — средний слой, куда публичные кейсы не заходят: это требует отказа от логики «универсальное решение для всех» и согласия на меньший потенциальный рынок ради более глубокого контроля над одним конкретным процессом.

### **Откуда берётся накопительное преимущество**

Операционный слой держится на трёх вещах, которые не упаковываются в управляемый продукт.

Первое — модель предметной области (domain world model). Это онтология с конкретными сущностями, их состояниями, валидными переходами и исключениями. Не «у вас есть документы про X», а «в этой организации сущность A может перейти в состояние B только если выполнены условия C, D, E, и сотрудник X не может одобрить переход без подтверждения от Y». Поисковые

системы вроде Glean знают, что в компании есть документы. Они не знают, что партнёры определённого типа требуют иного порядка согласования, что обращения определённой категории клиентов требуют немедленной эскалации, а не стандартной очереди. Эта разница между корпусом документов и операционной моделью — фундаментальная.

Второе — институциональное знание как исполнимая политика. Это превращение неформализованных знаний в граф решений: когда запускать повторное касание, правила скидки до определённого порога без эскалации, триггеры эскалации по типу клиента. Все эти решения, которые в традиционной компании живут «в головах сотрудников» и в разрозненных документах, в операционном слое становятся явным, исполнимым уровнем. Это не данные и не код, это кодифицированная логика принятия решений, которая отделяет компанию от конкурентов внутри той же отрасли. Она объясняет, почему две компании в одной нише с одинаковой обвязкой и одинаковой моделью получают разный операционный результат.

Третье — следы решений (trajectory data). Закрытый цикл «решение → исход» в конкретном контексте. Не «лог действий», а связка между принятым решением и тем, что произошло через N дней: оплатил клиент или нет, прибыло вовремя или нет, привело к сделке или нет. Эту связку нельзя восстановить ретроспективно из логов: контекст момента уже потерян. Подробнее этот компонент мы разбирали в отдельной заметке про данные траекторий.

Все три компонента объединяет одно: их нельзя купить или упаковать в SDK — их можно только накопить изнутри конкретной операционной среды. Управляемый продукт по определению не может конкурировать на этом слое. О том, почему именно представление, а не модель, оказывается главным барьером переключения, мы писали в заметке про слой представления и вертикальную схему.

## **Почему публичные кейсы туда не идут**

Структура стимулов объясняет это лучше, чем намерения игроков.

Activepieces и n8n — горизонтальные платформы по бизнес-модели. Им нужны кейсы, демонстрирующие универсальность: чем шире применимость, тем сильнее их позиционирование для расширения базы пользователей. Вертикальный операционный слой — анти-пример для такого позиционирования. Он показывает, что максимум ценности достигается, когда сборка делается под одну конкретную организацию или одну конкретную отрасль, а не как переиспользуемый шаблон.

Alan, Flow AI, Field Aerospace, Formula Bot оптимизируют под скорость отдачи: быстрые цифры для инвестиционного цикла. Горизонтальная автоматизация даёт этот срез; операционный слой требует месяцев работы внутри домена без яркого «до и после».

Наконец, провайдерская сторона рынка — Anthropic, OpenAI, AWS, Google — активно толкает управляемую обвязку как готовый продукт. К маю 2026 года стандартная среда исполнения стала массовым товаром у всех четырёх гиперскейлеров. Это значит, что публично видимая часть рынка дальше будет смещаться к ещё более горизонтальным кейсам, потому что именно там провайдерам выгодно показывать свои продукты. Операционный слой останется в тени публичной коммуникации, хотя именно туда уходит маржа. Эту разделительную линию между обвязкой и операционным слоем мы подробно разбирали в заметке про переход обвязки в массовый товар.

## **Как проверить, в каком слое живёт ваш продукт**

Полезные тесты для двух разных аудиторий.

Для основателей: если обвязку можно заменить на Anthropic Managed Agents или OpenAI AgentKit за один спринт без потери ценности — продукт в массовом слое. Если при смене среды исполнения разрушается накопленная модель предметной области, политика решений и следы решений — продукт уже в операционном слое.

Второй тест: если завтра конкурент с большей командой и большим бюджетом захочет повторить продукт за 90 дней, что у него получится? Универсальный сценарий повторяется. Онтология предметной области конкретного клиента, накопленная за 12 месяцев работы внутри организации, не повторяется.

Для технических руководителей, которые оценивают AI-внедрения. Если поставщик показывает кейсы только класса «X часов сэкономлено в месяц», стоит уточнить, какой слой подрядчик построил поверх горизонтальной автоматизации. Часовая экономия — это эффект первого года; через 18 месяцев он начинает выравниваться по рынку. Глубина изменения процесса — это другая категория измерения. Temporal в документации по принципам исполнения процессов формулирует похожую мысль для инфраструктурного слоя: устойчивое состояние и долго живущие процессы становятся ценными там, где процесс действительно встроен в работу бизнеса, а не там, где автоматизирована отдельная задача.

Второй вопрос для оценки внедрения: фиксируются ли следы решений отдельным слоем, или система пишет только логи действий. Это инженерное решение, которое нужно принимать в начале, а не добавлять задним числом.

## **Сигналы 2026 года**

Три сигнала покажут, как развивается дифференциация между слоями.

Если управляемые обвязки от Anthropic, OpenAI и AWS дойдут до общедоступности с встроенными вертикальными шаблонами — это означает попытку провайдеров занять операционный слой сверху, через пресеты. Сценарий маловероятный в ближайшие 12 месяцев, потому что отраслевое знание не упаковывается в шаблон, но провайдерская сторона будет пробовать. На встречные движения вертикальных игроков указывает, например, продукто-

вая платформа Vault AI от Veeva — отраслевой SaaS не отдаёт операционный слой без боя.

Если на p8n и Activepieces появятся кейсы, где экономика измеряется не в часах, а в процентах улучшения отраслевого исхода — например, «точность решения X выросла на Y процентов за 6 месяцев» — это будет сигналом, что операционный слой начинает становиться публично видимым.

Если вертикальные SaaS-игроки уровня Veeva и Procore начнут публично анонсировать собственные среды исполнения агентов — это означает движение третьего слоя вниз, к операционному слою, и сжатие пространства для независимых вертикальных AI-продуктов. Похожую общую логику смещения ценности на стыке инфраструктуры и данных описывает материал a16z про большие технологические идеи 2025.

## Главное

- Публичные кейсы AI-агентов дают честную экономию, но измеряют только горизонтальную плоскость: сокращение времени на существующие процессы.
- Между стандартной средой исполнения и отраслевым SaaS существует средний слой — операционный слой бизнеса, который держится на онтологии предметной области, исполнимой политике решений и следах решений.
- Этот слой нельзя упаковать в управляемый продукт: его компоненты создаются внутри конкретной организации и не воспроизводятся ретроспективно.
- Маржа в категории смещается именно туда по мере того, как стандартная среда становится готовым продуктом у всех крупных провайдеров.
- Тест: если продукт можно за спринт перенести на чужую управляемую среду без потери ценности, он живёт в массовом слое.

## FAQ

### Чем операционный слой отличается от отраслевого SaaS?

Отраслевой SaaS — это полноценный продукт отрасли со своими экранами, отчётами, ролями, машиной продаж и десятилетним циклом накопления отраслевой экспертизы. Операционный слой — это уровень, на котором агентная система становится способом, которым конкретная организация работает: онтология предметной области, политика решений и следы решений одной компании или одной узкой ниши. Отраслевой SaaS требует капитала и команды другого порядка; операционный слой строится небольшими специализированными командами и держится за счёт глубины внедрения, а не масштаба распространения.

### Почему стандартная среда исполнения считается массовой, если стартапы вокруг неё растут?

Рост стартапов на этом слое идёт за счёт расширения базы пользователей и быстрой отдачи, а не за счёт долгосрочной защиты. Anthropic в материале про управляемые агенты фиксирует тезис о том, что предположения обвязки уста-

ревают вместе с моделями. К маю 2026 года четыре крупнейших провайдера выпустили управляемые аналоги. Это не отменяет рост категории, но смещает источник маржи выше по стеку.

### **Когда операционный слой не имеет смысла строить?**

Когда задача действительно горизонтальная — общие исходящие касания, обработка документов без отраслевой специфики, синхронизация задач между корпоративными инструментами. В таких сценариях глубина операционного слоя избыточна, и горизонтальная сборка даёт лучшую экономику. Операционный слой оправдан там, где есть устойчивая предметная область с собственной онтологией и регуляторными или операционными ограничениями.

### **Сколько времени занимает построение операционного слоя?**

По открытым отраслевым материалам первая устойчивая версия онтологии и явной политики решений появляется обычно в горизонте от полугода до года активной работы внутри домена. Следы решений — самостоятельная категория, требующая отдельной инженерной работы; их накопительный эффект проявляется после нескольких месяцев замкнутого цикла «решение → исход».

### **Как отличить кейс операционного слоя от хорошо упакованной горизонтальной сборки?**

Главный маркер — что измеряется в результатах. Горизонтальная сборка измеряет сэкономленные часы и сокращение цикла существующего процесса. Операционный слой измеряет качество принимаемых решений в домене и устойчивость этого качества при смене модели или среды исполнения. Если кейс не показывает второй тип метрик, скорее всего, он живёт в горизонтальной плоскости.