AI Dev Live · 11 ноября 2026

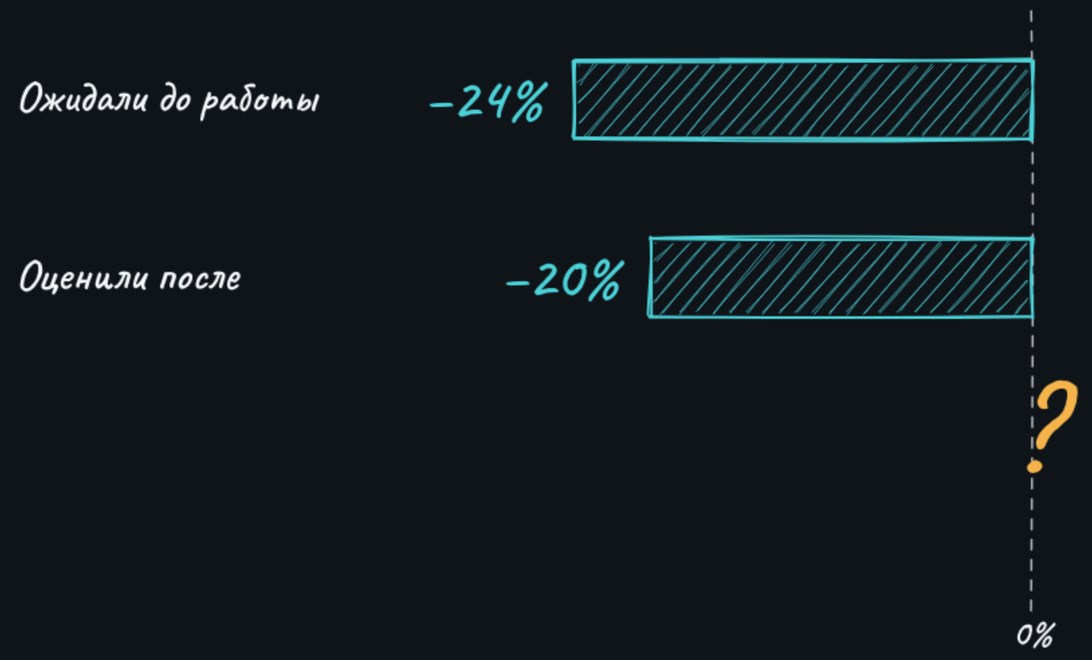
AI ускорил разработку



Как понять, что это правда

Что показало измерение?

METR · изменение времени выполнения

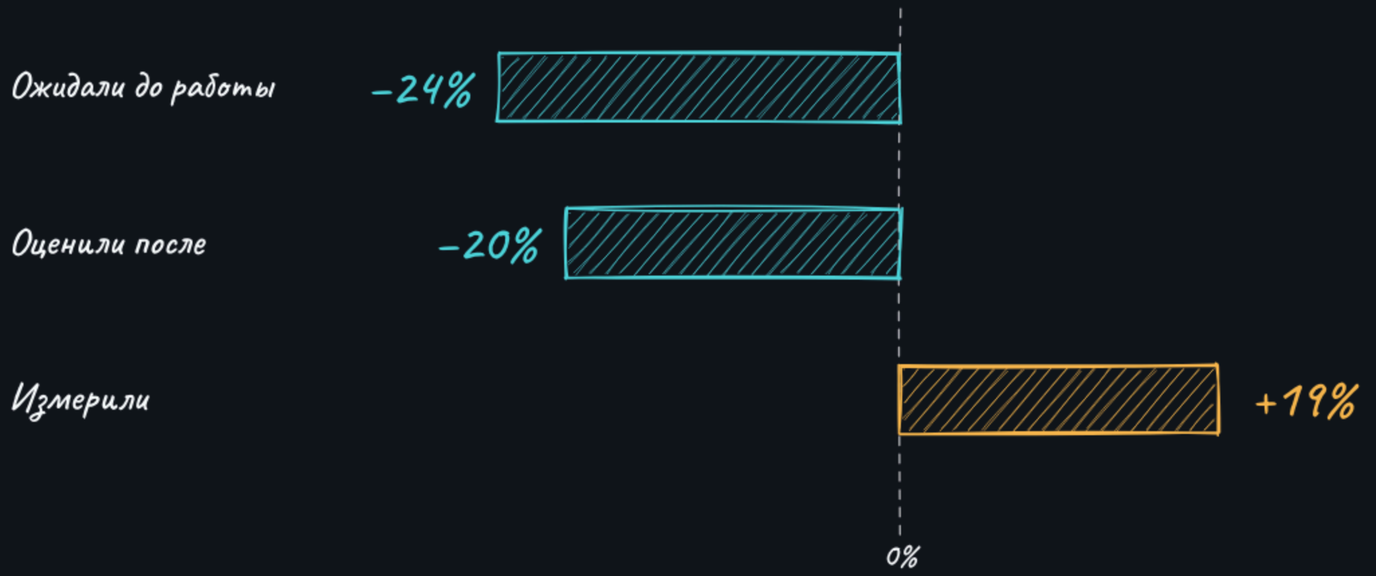


Ожидание и самоотчёт (self-report). Измеренный результат — дальше.

METR · Becker et al. (2025)

Ощущение ускорения требует отдельной проверки

METR · изменение времени выполнения



Все три числа — про время; отрицательное означает быстрее

METR · Becker et al. (2025)

Результат METR получен в знакомых репозиториях

METR · февраль–июнь 2025

+19%

время выполнения

95% ДИ: +2…+39%

16

опытных разработчиков

246

реальных задач

Cursor Pro · Claude 3.5/3.7 Sonnet · знакомые проекты

METR · Becker et al. (2025)

CLI-агенты: больше слитых PR

Microsoft · январь–апрель 2026

Число PR, а не время и качество

+24%

слитых PR на инженера в день

95% интервал: +14,5…+33,7%

−1,1%

на ложной дате запуска

интервал: −10,6…+8,6%

115

дней после запуска

Claude Code и Copilot CLI · ранние пользователи

Microsoft · Murphy-Hill et al. (2026)

В Anthropic стало больше коммитов на инженера

Anthropic · октябрь 2026



Каждый результат относится к своим условиям

Сравниваем условия, а не заголовки

METR	Microsoft	Anthropic
Время на задачу	Слитые PR в день	Объём работы на человека
Свой репозиторий	PR из Azure DevOps	Разные функции
Cursor · 2025	CLI-агенты · 2026	Агенты · 2026

Переносим проверку условий, а не процент

METR · Becker et al. (2025) Microsoft · Murphy-Hill et al. (2026) Anthropic · отчёт о занятости (2026)

Архитектура · команды · AI4SDLC

Александр Поломодов

Технический лидер с опытом работы в ролях
CTO и Technical Fellow в крупном финтехе



polomodov.tech

Research Insights Made Simple

Разбираю исследования и связываю их с инженерной практикой.

Фокус: Продуктивность · архитектура · AI4SDLC

Два масштаба измерения

Discovery + delivery

уровень компании

ценность для клиента

Только delivery

уровень компании

охват, влияние, затраты

Экономика сценариев

ревью, миграции,

прототипы

Сначала выбираем границу измерения

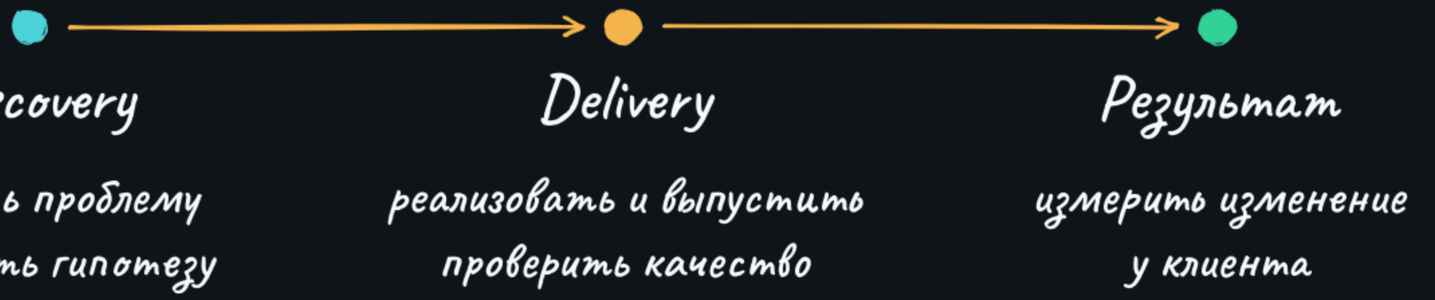
Измерение на уровне компании

1A

Discovery + delivery · путь до результата для клиента

От гипотезы до ценности для клиента

Компания · discovery + delivery



Связываем гипотезу, выпуск и эффект по одной инициативе

Как измерить сквозной результат

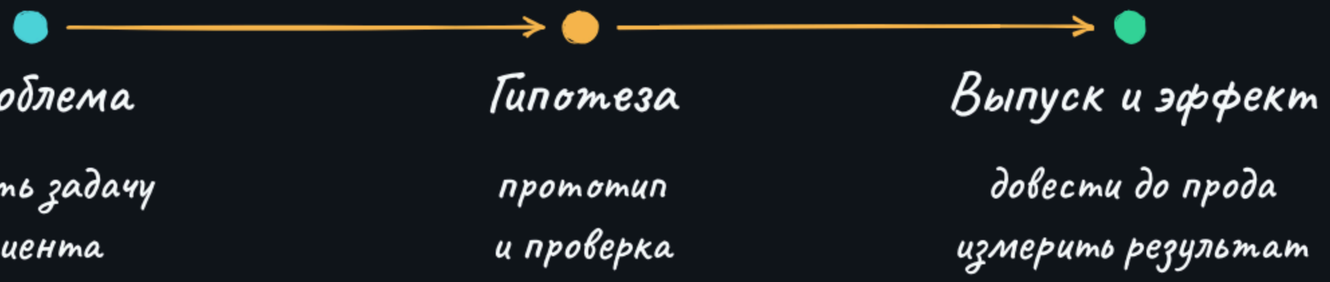
Компания · условный пример активации

Активация	Сквозной срок	Полные затраты
доля новых клиентов первый успех за 7 дней	от гипотезы до проверки результата	работа команд и AI включая неудачные попытки

Сравниваем сопоставимые инициативы и когорты клиентов

Product engineer ведёт задачу end-to-end

Компания · discovery + delivery



Возврат к основам продуктовой разработки

Code of Leadership №76 · Глеб Михеев Мой доклад · Эволюция разработки

Tiny team отвечает за весь цикл

Передачи

число передач
и время ожидания

Время до результата

от постановки задачи
до эффекта у клиента

Устойчивость

инциденты в проде
и нагрузка команды

Размер команды зависит от продукта, риска и поддержки платформы

Мой доклад · ВШЭ Мой доклад · Эволюция разработки

AI Dev Live · 2026

Delivery на уровне компании

1Б

Measuring AI code assistants and agents · DX и RIMS №11

Как DX измеряет AI в delivery

Measuring AI code assistants
and agents · RIMS №11



Смотрим AI-метрики вместе с общей продуктивностью команд

DX · Measuring AI code assistants and agents RIMS #11 Мой доклад · DotNext 2026

Считайте внедрение в контексте работы

Delivery · использование

Кто	Где	Как
команда · опыт	работа · репозиторий	частота · режим

Лицензия не описывает рабочую практику

Мой доклад · DotNext 2026

Проценту внедрения нужен знаменатель

Кого и что считаем

Люди

использовали / имели доступ

Задачи

с AI / подходили для AI

Окно наблюдения и правила включения фиксируются заранее

Режим помощи меняет работу

Ассистент

человек ведёт задачу
AI помогает в операциях

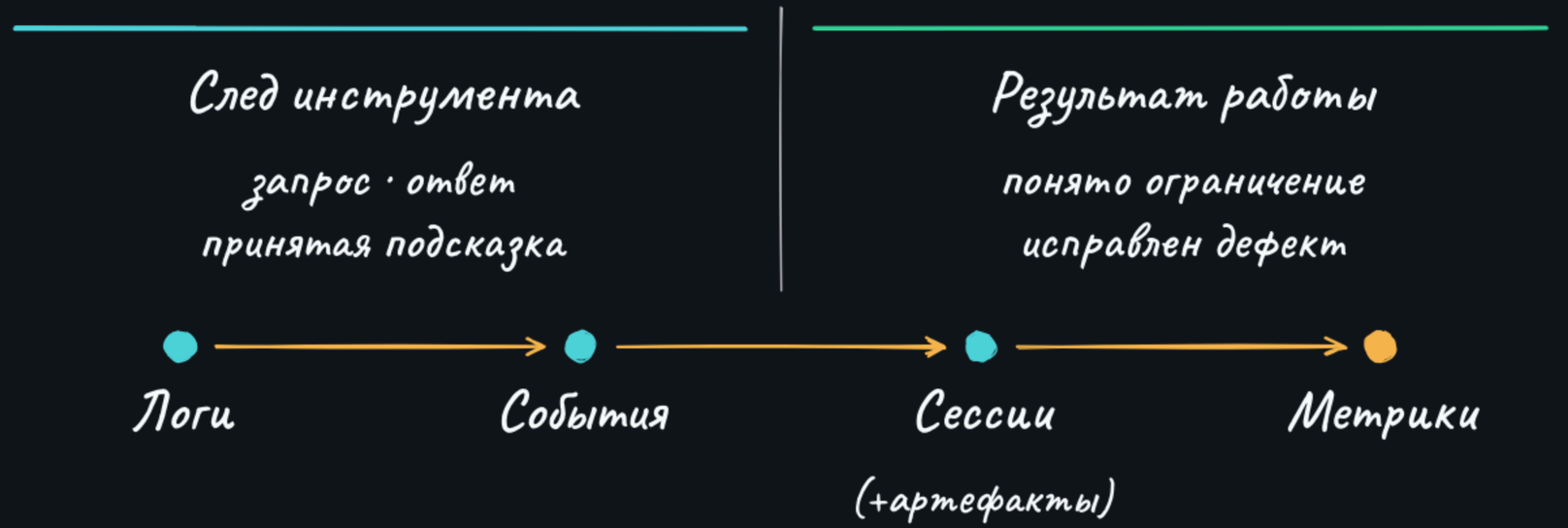
Делегирование

агент исполняет задачу
человек задаёт и принимает

Параллельность — отдельная характеристика обоих режимов

Событие инструмента ещё не результат

Что именно наблюдаем



Результат проверяем по критерию задачи

InSession (2020) Google Research · Measuring Developer Goals Мой разбор серии · RIMS №30

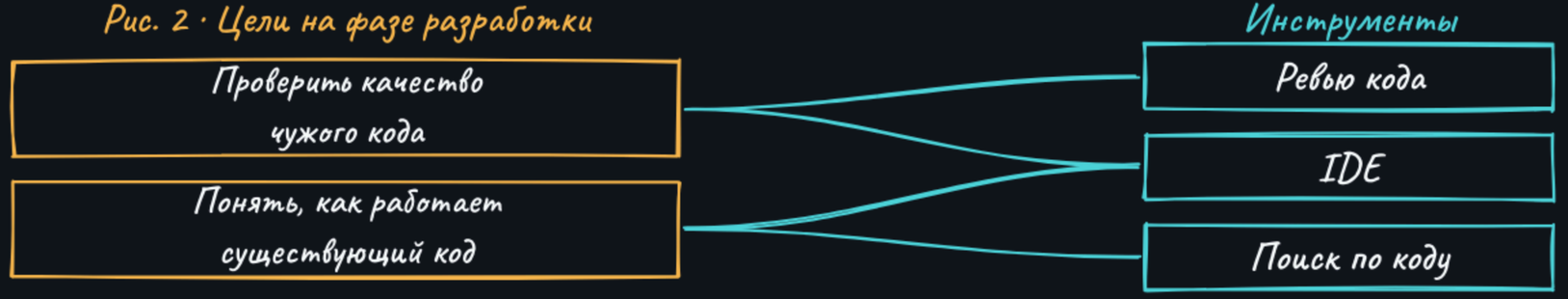
Google · Measuring Developer Goals

IEEE Software · 2024

Рис. 1 · Фазы SDLC и примеры целей



Рис. 2 · Цели на фазе разработки

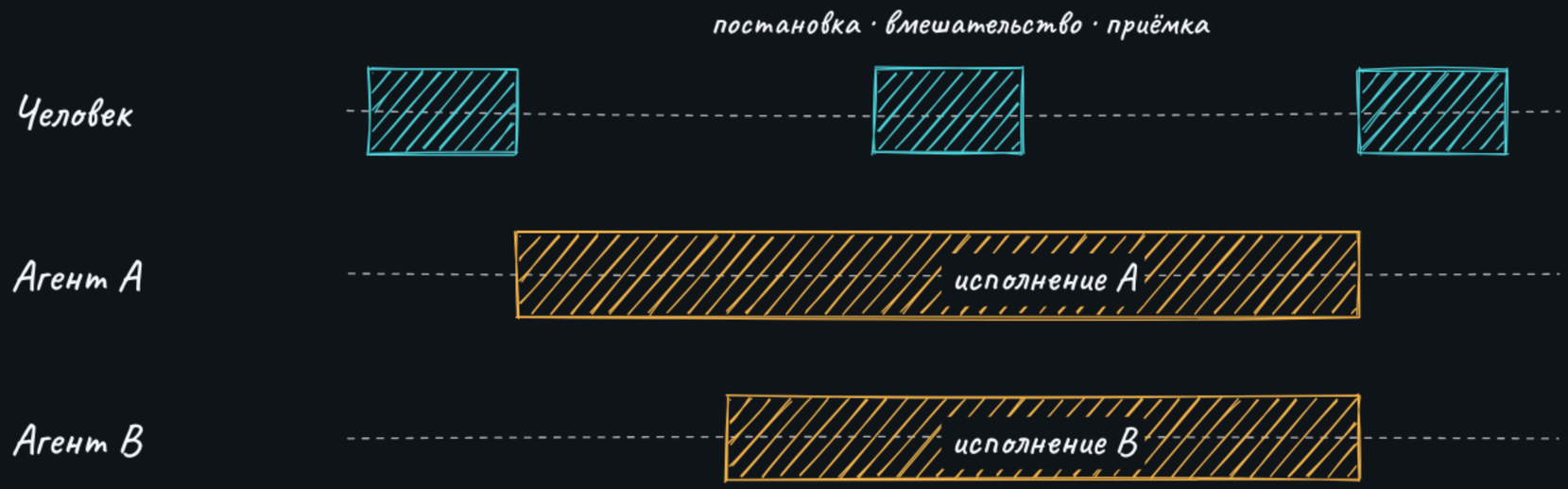


Пример приёмки · PR отвечает требованиям, блокирующие замечания устранены

Google Research · Measuring Developer Goals

Параллельные сессии нельзя складывать как трудозатраты

Схема без масштаба
· человек и агенты

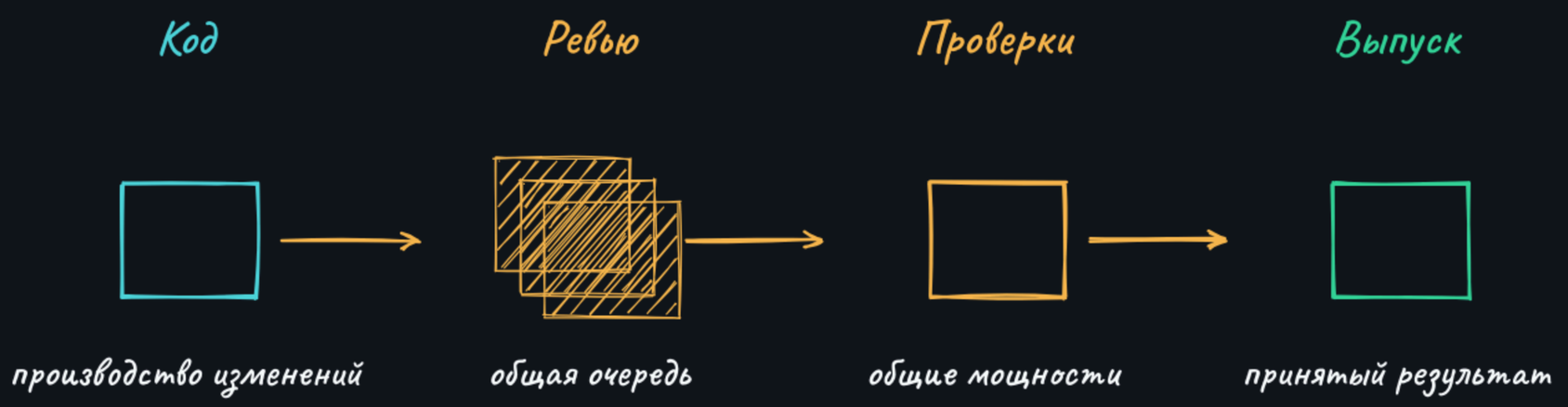


Три учёта: усилия человека · календарный срок · машинный расход

METR · пересмотр эксперимента (2026)

Локальное ускорение меняет очередь дальше

Схема потока · без масштаба



Смотрим ожидание, повторы и размер изменений

Мой доклад · Turbo ML Conf 2026

Логи и опрос складываются в одну картину

Anthropic · логи Claude Code и опрос · 2025

Агент делает больше шагов сам

логи: серия действий 9,8 → 21,2
реплик человека 6,2 → 4,1

Время на категорию задач — чуть меньше

опрос: зато заметно вырос объём сделанной работы

Выигрыш — в объёме, и это самоотчёт

Anthropic · как AI меняет работу (2025)

Три метода объясняют разные стороны

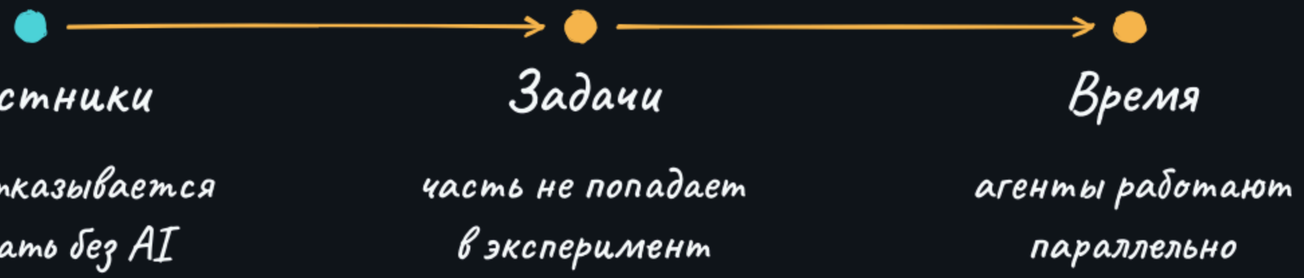
Логи	Опрос	Интервью
что происходило	как это переживали	почему возникло расхождение

Расхождение — повод разобраться

Google Research · Measuring Productivity Мой разбор серии · RIMS №30

Новый инструмент меняет само измерение

METR · обновление февраля 2026



Оставшаяся выборка может не представлять весь поток

METR · пересмотр эксперимента (2026)

Сравнение определяет силу вывода

Случайное назначение

сильнее контроль альтернативных причин

Поэтапное внедрение

нужны сравнимые группы и учёт времени

До и после

полезный сигнал слабее причинный вывод

METR · Becker et al. (2025) Microsoft · Murphy-Hill et al. (2026) Anthropic · отчёт о занятости (2026)

Экономика конкретных применений

02

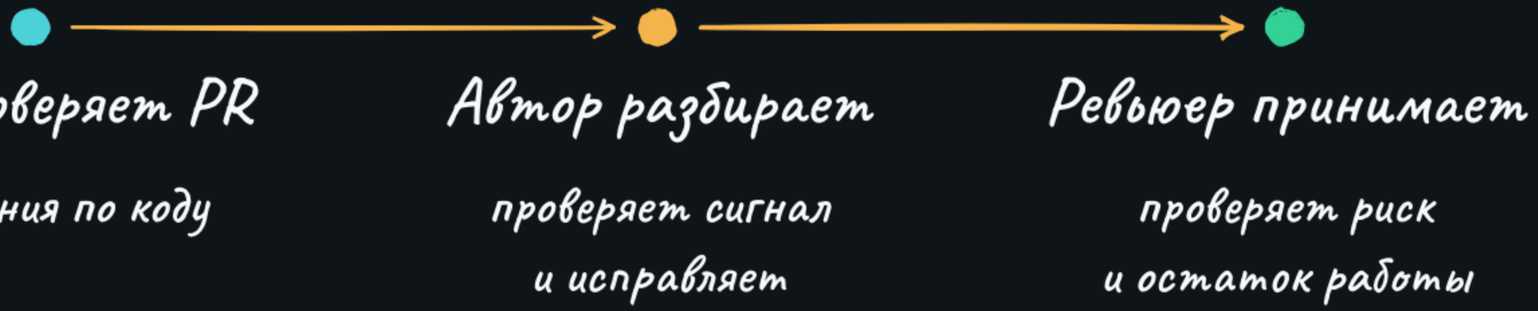
AI-ревью · миграции · прототипы интерфейса

У каждого сценария своя единица результата

AI-ревью	Миграция	Прототип
проверенный PR и качество после мержа	принятая партия с проверкой в проде	решение по гипотезе после проверки

В затратах учитываем людей, AI, проверки и переделки

AI-ревью уже включено



Точка сравнения — текущая конфигурация

GitHub · Automatic Copilot reviews

У полезного комментария есть цена

Польза	Труд и AI	Качество
подтверждённые важные находки	разбор, правки, повтор и стоимость модели	пропущенные дефекты после мержа

Сравниваем PR сходного размера, риска и типа

Developer Productivity for Humans #7 · Software Quality Мой разбор · Книжный куб DORA 2025

Настройка ревью меняет стоимость

AI-ревью · условный расчёт

$510	$320	$190
Текущие правила	Точнее правила	Разница
8 ч × $60 + $30 AI	5 ч × $60 + $20 AI	на 100 сопоставимых PR

Сохраняем критерий качества при сравнении

Мой разбор · Экономика AI в разработке

Настраиваем ревью по риску PR

Что меняем

правила, модель и повторные запуски

Как сравниваем

текущая настройка и новая настройка

Когда расширяем

меньше работы при прежнем качестве

Проверяем пропуски на ручной выборке

Миграции окупаются на повторениях

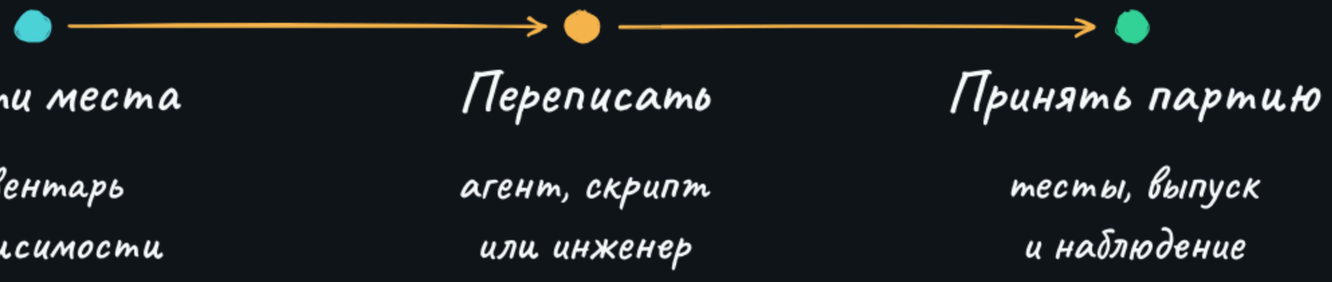
Google · Migrating Code At Scale

39	595	≈50%
Миграций	Изменений	Меньше времени
за 12 месяцев	отправлено в кодовую базу	оценка инженеров

Сравнение с прежними ручными миграциями

Google · Migrating Code At Scale

Миграция заканчивается приёмкой



Считаем весь путь, включая повторы и ручные исключения

Google · Migrating Code At Scale

Подготовка окупается с пятой партии

$720	$150	5 партий
Подготовка	Экономия на партии	Безубыточность
12 ч × $60	3 ч × $60 − $30 AI	$720 ÷ $150 ≈ 4,8

Условный пример · 6 ч вручную, 3 ч с автоматизацией и приёмкой

Хвост миграции влияет на решение

Покрытие

что ещё осталось на старом API

Надёжность

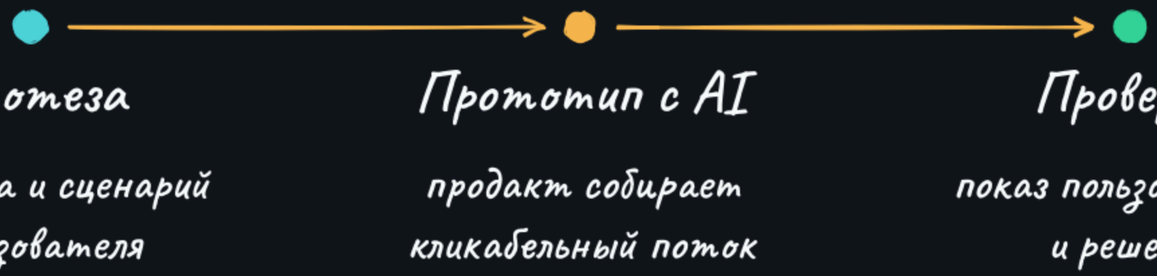
регрессии, откаты и сопровождение

Дальше

расширить автоматизацию или закрыть хвост вручную

Эффект от отключения старого контура считаем отдельно

Продакт проверяет интерфейс раньше



Результат — проверенная гипотеза и следующий шаг

Anthropic · Claude Design

Считаем стоимость решения по гипотезе

Время	Затраты	Результат
от вопроса до теста с людьми	продакт, дизайн, AI и переделки	проверить сценарий и выбрать следующий шаг

Сравниваем вопросы и качество проверки

Anthropic · Claude Design

Один тестовый сценарий дешевле на $160

Вайбкодинг · условный расчёт

$360	$200	$160
Обычный процесс	Прототип с AI	Разница
6 ч × $60	3 ч × $60 + $20 AI	один сценарий для той же проверки

Одинаковый пользовательский тест в обеих оценках

Прототип передаёт решение в delivery

Из discovery

сценарий, наблюдения и принятое решение

В delivery

состояния, доступность и интеграция

После выпуска

активация, ошибки и обратная связь

Код прототипа оцениваем по требованиям к продакшену

Anthropic · Claude Design Мой доклад · Эволюция разработки

Выигрыш бывает в часах, деньгах и риске

Время команды

полезная работа на высвободившихся часах

Денежный эффект

изменившиеся расходы или выручка

Риск и качество

меньше дефектов и потерь у клиентов

Считаем ROI для выбранного результата и периода

Что менять после расчёта

AI-ревью

настроить правила по риску PR

Миграции

сверить объём с точкой окупаемости

Прототипы

быстрее проверить продуктовый вопрос

Итог связываем с результатом клиента и работой команды

Продолжим разговор в подкастах

01 *Продуктивность для людей*

RIMS · цели инженера и три вида данных

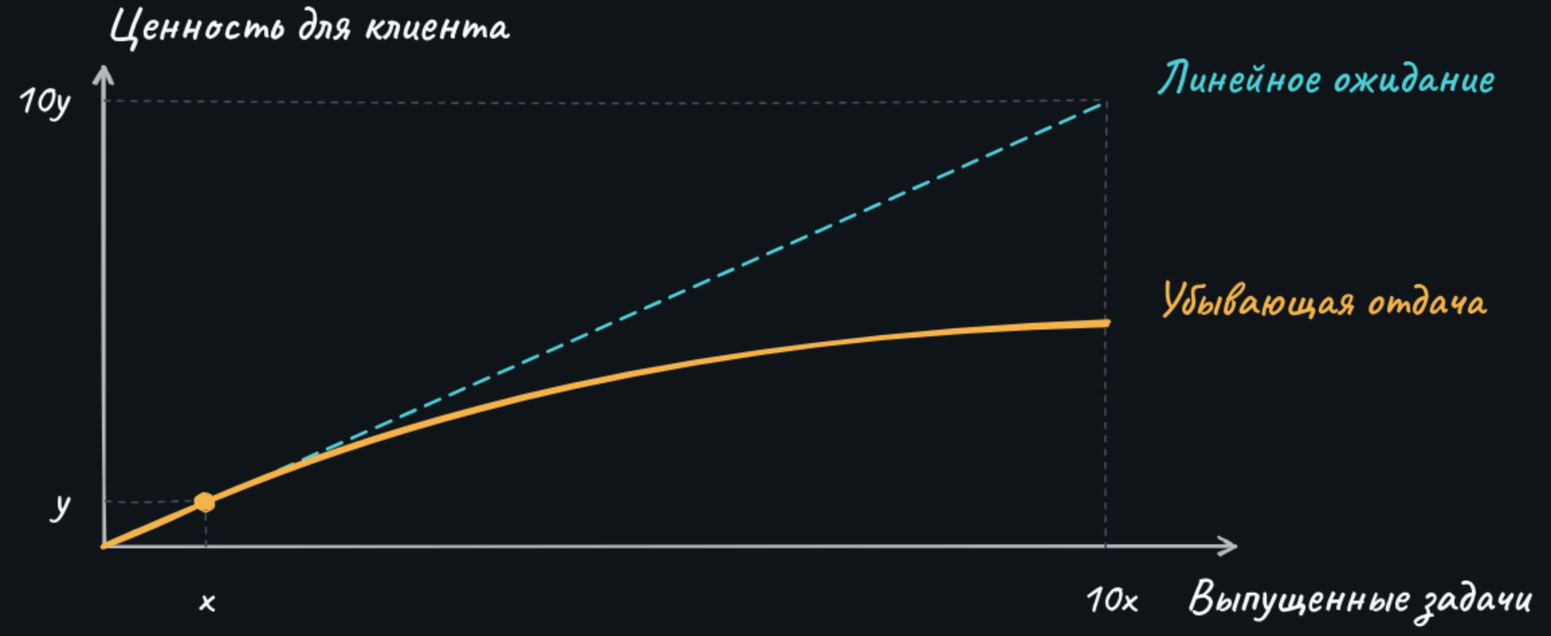
02 *AI и продуктивность разработчиков*

Код лидерства · разговор с Владимиром Долговым

03 *Экономика AI в разработке*

RIMS · стоимость принятой работы

Следующие задачи могут приносить меньше ценности



Условный пример убывающей предельной полезности

Code of Leadership №76 · Глеб Михеев OpenStax · микроэкономика

AI Dev Live · 11.11.2026

Измеряйте изменение работы

Discovery + delivery · экономика применений

Какое решение вы сможете принять после измерения?

Александр Поломодов · Технический лидер

polomodov.tech

Книжный куб



t.me/book_cube