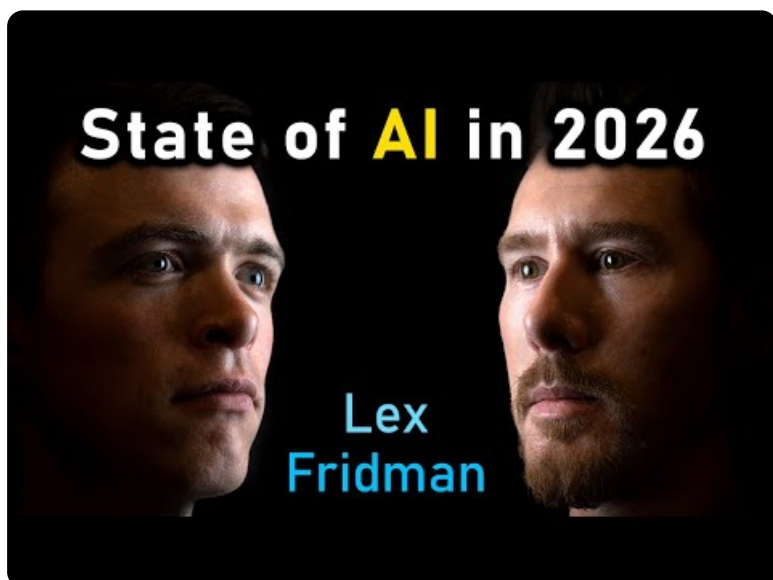


State of AI in 2026: LLMs, Coding, Scaling Laws, China, Agents, GPUs, AGI | Lex Fridman Podcast #490

@lexfridman · 2026-01-31 · <https://www.youtube.com/watch?v=EV7WhVT270Q>



Выпуск записан в формате визионерского анализа из начала 2026 года. Лекс Фридман обсуждает с исследователями машинного обучения Себастьяном Рашкой и Натаном Ламбертом тектонические сдвиги, произошедшие в индустрии за прошедший 2025 год.

Главная мысль эпизода: экстенсивное масштабирование претрейна уступает место алгоритмам RLVR (обучение с подкреплением на проверяемых наградах) и масштабированию на этапе инференса. Параллельно разворачивается жесткая геополитическая гонка, где китайские open-weight модели захватывают глобальное влияние, вынуждая США пересматривать свою стратегию закрытых API, а профессия программиста необратимо трансформируется от написания синтаксиса к управлению ИИ-агентами.

Ландшафт ИИ в 2026 году: США против Китая и битва моделей

Конец технологической монополии 03:09: Ни одна компания больше не обладает эксклюзивным доступом к секретным технологиям из-за постоянной ротации исследователей между лабораториями. Единственными реальными дифференцирующими факторами остаются размер бюджета и физические ограничения оборудования.

Смена лидеров хайп-цикла 04:12: Claude Opus 4.5 полностью доминирует в информационном поле, особенно в задачах написания кода. Модель Gemini 3,

несмотря на выдающиеся технические характеристики и более ранний релиз, потеряла маркетинговый импульс на фоне агрессивного продвижения Anthropic.

Фактчек: Проект «ADAM» (American Truly Open Models) — ▲: Ошибка транскрипции — инициатива Натана Ламберта называется ATOM Project, запущена летом 2025. Суть (ответ на доминирование китайских open-weight) передана верно.

Китайская экспансия open-weight 05:48: Релиз DeepSeek R1 в январе 2025 года запустил массовую волну открытых моделей в Китае. Компании Zhipu AI (линейка GLM), MiniMax и Moonshot (Kimi) выпускают флагманские модели, активно оспаривая статус DeepSeek как главного поставщика open-source ИИ.

Фактчек: Претрейн стоит 2-5 млн аренды кластера — ▲: Цифра верна лишь для средних open-моделей (OLMo; DeepSeek-V3 — ~5.6 млн только на финальный прогон без экспериментов); фронтальные претрейны — от десятков до сотен миллионов. Общий тезис «инференс для миллионов пользователей дороже обучения» — □, но GPT добавляет: универсального превосходства инференс-масштабирования над претрейном нет, зависит от задачи и задержки.

Стратегия влияния через открытый код 06:51: Китайские компании делают ставку на open-weight, так как западный корпоративный сектор отказывается покупать доступ к китайским API из соображений безопасности. Бесплатные веса позволяют Китаю интегрировать свои технологии в глобальную инфраструктуру и формировать международные стандарты.

Фактчек: gpt-oss-120b — первая открытая модель, заточенная под инструменты — ▲: Рецензенты разошлись (Gemini □, Claude ▲, GPT □). Нативный tool calling раньше был у Qwen, GLM, Kimi K2, Hermes 2 Pro, Gorilla; уникальность gpt-oss — браузер и Python встроены в само reasoning-обучение. Инструменты снижают, но не устраняют галлюцинации: модель может выбрать не тот источник или исказить результат.

Экономика маршрутизации запросов 11:13: Внедрение роутера в GPT-5 экономит колоссальные вычислительные ресурсы. Система автоматически перенаправляет простые запросы на дешевые модели, избавляя OpenAI от необходимости тратить дорогие GPU-часы флагманской сети на базовые задачи пользователей.

Структурное преимущество Google 12:19: Google обладает уникальным преимуществом благодаря собственной инфраструктуре TPU. Отсутствие необходимости платить гигантскую маржу NVIDIA за чипы позволяет компании масштабировать Gemini с гораздо меньшими капитальными затратами по сравнению с конкурентами.

Разделение сценариев использования 14:29: Пользователи активно комбинируют модели в зависимости от задачи. Быстрые сети (Gemini или GPT-5 non-thinking) применяются для мгновенной генерации bash-скриптов, тогда как

«Pro»-модели запускаются в фоновом режиме на десятки минут для глубокого анализа и верификации сложных документов.

Фактчек: Claude Opus 4.5 доминирует в коде, Gemini 3 потеряла импульс — [1]: Opus 4.5 первым пробил 80% на SWE-bench Verified и признан разработчиками лидером для кода; Gemini 3 Pro вышла почти одновременно, но проиграла информационно.

Хрупкость пользовательской лояльности 18:13: Приверженность конкретной LLM работает до первой критической ошибки. Пользователь эксплуатирует любимую модель до тех пор, пока она не выдаст откровенно глупый ответ (эффект порога), после чего моментально переключается на конкурента (Claude, ChatGPT или Gemini).

Фактчек: Компенсация в OpenAI превышает 1 млн/год — [2]: По данным WSJ/The Information медианная компенсация технического персонала OpenAI в 2025 — ~1.5 млн, в основном акциями.

Превосходство закрытых моделей США 20:51: Американские проприетарные модели остаются умнее и быстрее китайских аналогов. Из-за экспортных ограничений китайские сети часто разворачиваются с меньшим количеством GPU на реплику, что снижает скорость инференса и заставляет профессионалов продолжать платить за API из США.

Архитектура трансформеров и новые законы масштабирования

Эволюция открытых архитектур 30:54: Китайские open-weight модели исторически тяготеют к массивным архитектурам Mixture of Experts (MoE) для достижения пиковой производительности. Западные разработчики (Mistral, Gemma, OLMo) изначально фокусировались на компактных размерах, но сейчас также переходят к масштабам 400B+ параметров (например, грядущий Nemotron).

Интеграция инструментов как стандарт 32:31: Модель gpt-oss-120b стала первой открытой сетью, архитектурно заточенной под использование внешних инструментов (веб-поиск, интерпретатор Python). Вызов API — наиболее эффективный метод борьбы с галлюцинациями, так как модель вычисляет или ищет факты вместо попыток извлечь их из собственных весов.

Отсутствие фундаментальных сдвигов 38:50: Базовая архитектура не менялась со времен GPT-2 (авторегрессионный трансформер). Все современные прорывы базируются на точечных модификациях: Group Query Attention (GQA), Sliding Window Attention, Multi-head Latent Attention (MLA) для сжатия KV-кэша и замене слоев нормализации (RMSNorm).

Механика Mixture of Experts (MoE) 41:26: Архитектура MoE заменяет плотный полносвязный слой на множество независимых «экспертов» (например, 256 нейросетей). Специальный роутер направляет каждый токен только к нужным

экспертам, что радикально увеличивает общую емкость знаний модели без роста вычислительных затрат на один прямой проход.

Фактчек: МоЕ увеличивает ёмкость без роста вычислений на прямой проход —

□: Активируется лишь часть экспертов на токен. Но «без роста затрат» — преувеличение: все веса надо хранить в памяти, маршрутизация и обмен между GPU дают накладные расходы.

Системная инженерия важнее архитектуры 46:12: Переход на форматы вычислений FP8 и FP4 дает огромный прирост скорости. Снижение объема памяти на один параметр уменьшает издержки на коммуникацию между GPU и увеличивает пропускную способность кластера с 10 000 до 13 000 токенов в секунду на видеокарту.

Трехмерные законы масштабирования 48:51: Scaling Laws теперь работают в трех плоскостях: претрейн (размер модели и объем данных), обучение с подкреплением (длительность этапа проб и ошибок) и масштабирование на этапе инференса (генерация скрытых цепочек рассуждений перед выдачей финального ответа).

Экономика претрейна против инференса 52:32: Претрейн стоит относительно дешево (аренда кластера обходится в \$2–5 млн) по сравнению с миллиардными затратами на инференс для миллионов пользователей. Лаборатории стремятся делать модели как можно меньше, чтобы радикально снизить постоянные операционные расходы на обслуживание.

Фактчек: Qwen 3: с 15% до 50% за 50 шагов → RLVR не создаёт новых знаний —

△: Направление верное — работы 2025 (one-shot RLVR, «Spurious Rewards») подтверждают быструю разблокировку латентных навыков претрейна, но эффект во многом специфичен для Qwen (насыщен мат-данными), на Llama слабее. Конкретные 15%→50% не верифицированы, а число шагов само по себе не аргумент — один шаг может обрабатывать огромный батч. Обобщение «RLVR никогда не учит новому» — открытый исследовательский вопрос.

Гетерогенные вычисления для RL 59:23: Масштабирование обучения с подкреплением требует сложной сетевой топологии. Узлы-«акторы» генерируют ответы, а узлы-«ученики» параллельно вычисляют градиенты (алгоритмы PPO или GRPO), что создает совершенно иные, асинхронные нагрузки на кластер по сравнению с плотным синхронным претрейном.

Триумф масштабирования инференса 1:02:29: Парадигма o1 доказала, что сейчас выгоднее тратить вычислительные мощности на генерацию длинных рассуждений в компактной модели во время запроса, чем вкладывать сотни миллионов в претрейн гигантской сети, которая устареет через полгода.

Данные для обучения: качество, синтетика и копирайт

Трехэтапный пайплайн обучения 1:04:33: Современное обучение делится на три фазы: претрейн (впитывание сырых знаний через предсказание токенов), мид-трейн (специализация на длинных контекстах и высококачественных данных для предотвращения катастрофического забывания) и пост-трейн (разблокировка навыков через RL и SFT).

Синтетика как стандарт претрейна 1:08:48: Лаборатории массово применяют OCR-модели для извлечения триллионов токенов из нераспознанных PDF-файлов. Также сети обучаются на лучших ответах ChatGPT, чтобы с первых шагов привить модели идеальную грамматику и структуру, избегая «грязных» данных из интернета.

Алгоритмический подбор датасетов 1:11:59: Оптимальный состав обучающей выборки вычисляется с помощью линейной регрессии. Исследователи обучают маленькие модели на разных пропорциях данных (GitHub, Reddit, Wikipedia) и корректируют микс в зависимости от целевых метрик (например, резко увеличивают долю кода для моделей рассуждения).

Проприетарные данные как следующий ров 1:16:15: Следующий рубеж масштабирования — закрытые отраслевые данные (результаты клинических испытаний, закрытая юриспруденция, финансы). Корпорации будут нанимать инженеров для создания внутренних моделей на данных, к которым у создателей ChatGPT нет и не будет легального доступа.

Юридические риски парсинга 1:17:18: Anthropic проиграла судебный иск на \$1.5 млрд в пользу авторов за использование торрентов с пиратскими книгами для обучения моделей. Этот прецедент подчеркивает колоссальные финансовые риски нелегального сбора данных и подталкивает индустрию к легальным схемам компенсации (по аналогии со Spotify).

Фактчек: Anthropic проиграла иск на \$1.5 млрд за пиратские книги — ⚠: Все три рецензента сходятся: это мировое соглашение (Bartz v. Anthropic, сентябрь 2025), а не проигрыш; суд признал обучение на легально купленных книгах fair use, ответственность возникла только за пиратскую библиотеку (LibGen/PiLiMi). Gemini указал дату «июль 2026» — ошибка, на момент записи соглашение уже состоялось.

Влияние ИИ на разработку и человеческий опыт

Фильтрация ИИ-кода человеком 1:19:24: Open-source мейнтейнеры выгорают от потока пулл-реквестов, сгенерированных ИИ. Однако при наличии человека-верификатора этот процесс превращается в мощный фильтр, создающий высококачественные данные, аналогичные дорогостоящей разметке для RLHF.

Утрата авторского голоса 1:23:07: LLM теряют глубину мысли из-за алгоритмов RLHF. Обучение на основе усредненных человеческих предпочтений

сглаживает углы, делая тексты безопасными, но лишенными резкости, оригинальности и сырой эмоции, присущих настоящим исследователям.

Санитаризация моделей 1:25:47: Фронтирные лаборатории жестко цензурируют модели, чтобы избежать PR-катастроф (например, обвинений ИИ в доведении пользователей до самоубийства). Эта фильтрация делает сети стерильными и блокирует сложные, провокационные диалоги, необходимые для интеллектуального роста человека.

Сеньоры доверяют ИИ больше джунов 1:31:09: Опрос 791 разработчика показал, что senior-программисты чаще junior-ов отправляют в продакшен код, более чем на 50% написанный ИИ. Опытные специалисты лучше понимают, как формулировать архитектурные задачи и верифицировать сгенерированный результат.

Опасность отсутствия трудностей 1:34:21: Использование ИИ для отладки лишает программистов опыта преодоления трудностей. Без самостоятельной многочасовой борьбы с багами junior-разработчики рискуют никогда не развить глубокую интуицию и не перейти на уровень экспертов.

Post-training: Триумф RLVR и ограничения RLHF

Механика RLVR 1:38:29: Алгоритм RLVR (Reinforcement Learning with Verifiable Rewards), популяризованный DeepSeek R1, масштабирует обучение за счет проверяемых задач (математика, код). Точность ответа служит объективным сигналом награды, полностью устраняя необходимость в дорогостоящей человеческой разметке.

Спонтанное возникновение рассуждений 1:42:11: В процессе RLVR модели самостоятельно учатся писать пошаговые объяснения и исправлять собственные ошибки («aha moments»). Эти паттерны возникают естественным образом, так как длинные цепочки рассуждений статистически повышают итоговую точность ответа.

RLVR не создает новых знаний 1:44:48: RLVR лишь разблокирует навыки, полученные на этапе претрейна. Базовая модель Qwen 3 повысила точность решения математических задач с 15% до 50% всего за 50 шагов обновления градиента, что физически невозможно для усвоения принципиально новой информации.

Проблема вырождения сигнала в GRPO 1:49:08: Алгоритм GRPO оценивает ответ относительно других попыток решения той же задачи. Если задача слишком простая и все ответы сети верны, сигнал награды исчезает. Это вынуждает лаборатории постоянно усложнять тренировочные датасеты, переходя к сложным научным доменам.

Пределы масштабирования RLHF 1:53:22: RLHF не подчиняется законам масштабирования: избыток вычислений приводит к переоптимизации стиля (over-optimization). В отличие от него, RLVR дает линейный прирост производительности при логарифмическом увеличении вычислительных мощностей, забирая на себя основные бюджеты пост-трейна.

Образование, исследования и культура выгорания в Big Tech

Обучение через создание с нуля 1:59:15: Написание LLM с нуля на одном GPU — лучший способ понять архитектуру. Промышленные библиотеки вроде Hugging Face Transformers содержат сотни тысяч строк кода для поддержки всех возможных юзкейсов, что делает их абсолютно непригодными для фундаментального обучения.

Нишевые исследования для независимых ученых 2:04:01: Независимые исследователи могут влиять на индустрию, фокусируясь на узких нишах (например, «character training» для настройки сарказма модели) или создавая новые бенчмарки, выявляющие слабые места будущих флагманских сетей.

Альтернативные издержки PhD 2:15:13: Финансовые потери от выбора академической карьеры огромны: средняя компенсация в OpenAI превышает \$1 млн в год в виде акций. Однако работа в закрытых лабораториях превращает ученых в безымянных «винтиков системы», лишая их возможности публиковать статьи.

Культура 9-9-6 в ИИ-индустрии 2:21:07: В топовых лабораториях укоренилась культура работы с 9 утра до 9 вечера, 6 дней в неделю. Жесткая конкуренция приводит к массовому выгоранию, вынуждая компании вводить экстренные протоколы «спасения браков» для принудительной отправки инженеров домой.

Пузырь Кремниевой долины 2:25:48: Кремниевая долина функционирует в «поле искажения реальности». Этот информационный пузырь стимулирует сверхпродуктивность и веру в скорый AGI, но полностью отрывает разработчиков от понимания реальных потребностей и опыта остального человечества.

Альтернативные архитектуры, использование инструментов и память

Текстовые диффузионные модели 2:29:36: Диффузионные сети (например, Gemini Diffusion) генерируют токены параллельно через итеративное шумоподавление. Они работают значительно быстрее авторегрессионных моделей при выдаче огромных массивов кода (diff-ов), хотя пока уступают в задачах сложного логического вывода.

Рекурсивные языковые модели 2:35:55: Рекурсивные LLM разбивают задачи с длинным контекстом на подзадачи. Модель автономно вызывает собственные копии для решения каждого фрагмента, что экономит память и повышает точность по сравнению с попытками обработать весь контекст за один проход.

Преимущество закрытых моделей в интеграции 2:37:31: Проприетарные модели имеют огромное преимущество в использовании инструментов благодаря глубокой интеграции в облачную инфраструктуру (например, Claude Code напрямую правит репозитории). Open-source модели вынуждены оставаться универсальными, перекладывая настройку окружения на пользователя.

In-context learning вместо обновления весов 2:39:40: Непрерывное обучение (continual learning) с обновлением весов на лету пока экономически нецелесообразно для массового рынка. Вместо этого лаборатории полагаются на in-context learning, загружая историю пользователя в огромные контекстные окна.

Проблема забывания в LoRA 2:43:53: Использование LoRA-адаптеров для персонализации моделей сталкивается с проблемой «бесплатного сыра не бывает». Чем больше новых специфических знаний модель усваивает через адаптеры, тем больше базовой информации она катастрофически забывает.

RL-управление контекстом 2:48:34: Будущие агентные модели будут использовать RL для самостоятельного управления контекстом. Модель научится сжимать свою историю из 100 000 токенов в минимальный маркированный список, сохраняя при этом максимальную точность ответов для экономии вычислений.

Робототехника, AGI и автоматизация программирования

Мировые модели (World Models) в LLM 2:51:55: Внедрение мировых моделей заставит сети симулировать физическую среду и промежуточные переменные, а не просто предсказывать следующий токен. Это аналогично переходу AlphaFold к явному моделированию физических взаимодействий молекул.

Open-source экосистема для роботов 2:55:07: Робототехника переживает бум благодаря инфраструктуре, созданной для LLM. Ожидается появление экосистемы открытых робототехнических моделей на Hugging Face, которые пользователи смогут скачивать и дообучать под конкретные физические задачи.

Барьер sim-to-real в домашних роботах 2:58:17: Разрыв между симуляцией и реальностью и экстремальные требования к безопасности делают появление универсальных домашних роботов делом далекого будущего. Робот в доме человека не имеет права на ошибку даже в 1% случаев, в отличие от текстовых LLM.

Неравномерность навыков ИИ-кодера 3:01:28: Отчет «AI 2027»

предсказывает появление сверхчеловеческого ИИ-программиста к 2031 году. Однако возможности моделей крайне неравномерны: они блестяще пишут frontend, но проваливаются в распределенном ML из-за критического недостатка обучающих данных в этой узкой сфере.

Фактчек: Отчёт AI 2027 предсказывает сверхчеловеческого программиста к 2031 — Δ : Сам сценарий ставил superhuman coder на 2027; к концу 2025 авторы (Кокотайло, Лифланд) сдвинули медианы на 2029–2031. 2031 — обновлённая оценка авторов, а не содержание отчёта.

Трансформация профессии программиста 3:06:50: По мере автоматизации написания синтаксиса роль инженеров сведется к проектированию систем. Разработчики будут управлять дашбордами, координируя работу множества ИИ-агентов, которые автономно реализуют фичи и исправляют баги.

AGI через узкие домены 3:13:46: Путь к AGI лежит через узкоспециализированные научные домены. Стартапы с сотнями миллионов долларов финансирования уже интегрируют LLM с реальными мокрыми лабораториями и математическими средами (Lean) для автоматической генерации и проверки гипотез.

Провал универсального computer use 3:18:32: Обучение ИИ управлению компьютером (имитация кликов и ввода) продвигается медленно. Моделям гораздо сложнее взаимодействовать с визуальным интерфейсом ОС и нечеткими намерениями пользователя, чем работать через строго заданные backend API.

Экономика ИИ: реклама, слияния и проект ADAM

Истинный экономический эффект ИИ 3:29:08: Главный эффект ИИ на ВВП — не внезапная сингулярность, а тихая революция доступности знаний. Возможность любого ребенка или работника в мире получить персонализированный ответ на любой вопрос окажет колоссальное, компаундирующее влияние на глобальную экономику.

Неизбежность ИИ-рекламы 3:32:58: ИИ-чатботы неизбежно внедрят нативную рекламу для субсидирования гигантских затрат на инференс. Тот, кто первым запустит этот маховик (по аналогии с YouTube), получит неоспоримое финансовое преимущество для инвестиций в R&D новых моделей.

Разрушение стартап-экосистемы 3:37:14: Экосистема стартапов разрушается из-за «лицензионных сделок» (псевдо-поглощений корпорациями для обхода антимонопольных законов). Такие сделки забирают лучших талантов, но оставляют рядовых сотрудников без выплат по акциям, уничтожая мотивацию работать в стартапах.

Крах проекта Llama 3:44:42: Проект Llama от Meta пострадал из-за внутренних политических конфликтов и одержимости бенчмарками. Руководство сфокусировалось на гигантских моделях для пиара, проигнорировав потребность open-source сообщества в компактных и удобных для локального запуска сетях.

Фактчек: Крах Llama из-за внутренней политики и одержимости бенчмарками —

□: Инсайдерская оценка закрытых корпоративных процессов, независимых подтверждений нет.

Проект ADAM против гегемонии Китая 3:50:33: Проект ADAM (American Truly Open Models) запущен в ответ на доминирование китайских open-weight моделей. Китайские компании захватывают глобальное влияние, предоставляя бесплатные веса, в то время как США рискуют остаться только с закрытыми платными API.

Невозможность запрета open-weight 3:56:24: Запрет на публикацию открытых моделей в США под предлогом безопасности невозможен без создания аналога «Великого китайского файрвола». Обучение моделей за \$100 млн доступно многим игрокам в мире, поэтому США обязаны лидировать в open-source для удержания талантов.

Железо, историческая перспектива и будущее человечества

Уязвимость монополии NVIDIA 4:01:13: Главный ров NVIDIA — это не кремний, а 15-летняя экосистема CUDA. Однако LLM уже способны писать код, абстрагирующий работу с железом, что в будущем может снизить зависимость индустрии от CUDA и открыть дорогу альтернативным чипам.

Роль личности в истории технологий 4:04:25: Выдающиеся личности, такие как Дженсен Хуанг, радикально ускоряют ход истории. Без его фанатичной веры в GPU революция глубокого обучения могла бы задержаться на десятилетия, направив технологическое развитие цивилизации по совершенно иному пути.

Забвение архитектуры трансформера 4:11:51: Через 100 лет историки назовут главными драйверами нашей эпохи «вычислительные мощности» и «глубокое обучение». Конкретные архитектуры, такие как трансформер, скорее всего, будут полностью вытеснены новыми парадигмами и забыты.

Премия за физическую реальность 4:17:44: По мере того как интернет будет тонуть в цифровом ИИ-мусоре (slop), общество начнет придавать огромную ценность физическим объектам. Настоящие картины, живые мероприятия и верифицированный человеческий труд получают беспрецедентную культурную и финансовую премию.

Гонка аппаратных водяных знаков 4:20:26: Будущее контента — это гонка вооружений в сфере верификации. Устройства будут на аппаратном уровне

криптографически подписывать фотографии и тексты, созданные человеком, чтобы отличать их от бесконечного потока ИИ-генераций.

ИИ как инструмент, а не субъект 4:23:06: ИИ не захватит мир, так как лишен сознания и собственной воли (agency). Модели остаются исключительно инструментами, которые требуют человеческого намерения и целеполагания для совершения любых действий в физическом или цифровом мире.

Фактчек: ИИ не захватит мир, потому что лишён сознания и воли — □: Логически слабый аргумент — сознание не требуется для автономного выполнения задач, ошибок и ущерба через инструменты. Агентность задаётся устройством системы (цели, цикл действий, память, разрешения), а не субъективными переживаниями.

[4:24:11-4:25:13] смолток, взаимные благодарности и завершение подкаста.

ЛИЧНО ДЛЯ ТЕБЯ

Раз в неделю чини сложный баг руками: Тезис «сеньоры доверяют AI-коду, потому что умеют верифицировать» — про тебя. Режим «AI пишет, ты читаешь» работает ровно до тех пор, пока навык отладки не атрофировался; страховка — перед чтением AI-объяснения самому предсказать поведение кода и граничные случаи.

Маршрутизируй модели и меряй время до проверенного результата:

Быстрая модель (Gemini/GPT non-thinking) — для одноразовых скриптов; Pro/deep в фоне на десятки минут — для анализа документов (анализы крови, юридические бумаги по стройке); для фактов — только модель с поиском/кодом. Не храни лояльность — переключайся после первой глупой ошибки. Сравни Claude, Codex и Gemini на одинаковых задачах по минутам на постановку+ревью+правки, а не по объёму кода.

Агентам — песочница и минимальные права: Bash, удаление файлов, доступ к секретам и публикации требуют подтверждения необратимых операций независимо от рейтинга модели. Совпадение ответов двух моделей — не независимое доказательство; тест или первоисточник надёжнее.

Винил и дом — ставка на актив, который ИИ не обесценивает: Прогноз «интернет тонет в slop, физические объекты и верифицированный человеческий труд дорожают» совпадает с твоими вложениями в аудиотехнику, пластинки, живые концерты и стройку. Не считай это иррациональной тратой.

КОНСЕНСУС РЕЦЕНЗЕНТОВ

Все трое сошлись: дело Anthropic — мировое соглашение, а не проигрыш; «\$2–5 млн за претрейн» верно только для средних open-моделей; ключевой личный вывод — сохранять навык ручной верификации и отладки. Разошлись по gpt-oss-120b (Gemini □ vs GPT □ — были ли раньше tool-native открытые модели) и по

тезису «RLVR не создаёт знаний» (Claude считает подтверждённым литературой, GPT — недоказанным); Gemini допустил ошибку в дате соглашения Anthropic («июль 2026»).