

Terence Tao - How the world's top mathematician uses AI

@DwarkeshPatel · 2026-03-20 · <https://www.youtube.com/watch?v=Q8Fkpi18QXU>



Дваркеш Патель беседует с Теренсом Тао, одним из величайших современных математиков и лауреатом Филдсовской премии. Главная тема выпуска — как искусственный интеллект меняет научный метод, процесс математических исследований и повседневную работу ученых.

Тао объясняет, почему ИИ обнулil стоимость генерации гипотез, сделав главным узким местом науки их верификацию. Подробно разбираются пределы возможностей текущих моделей, разница между «искусственной изобретательностью» и подлинным интеллектом, роль формальных языков доказательств (Lean) и критическая важность сохранения случайности (серендипности) в жестко оптимизированном академическом мире.

Исторические параллели: Кеплер, данные и генерация гипотез

Кеплер открыл законы движения планет благодаря доступу к уникальному массиву данных Тихо Браге 00:00: Изначально Кеплер пытался вписать орбиты шести известных планет в пять платоновых тел (куб, тетраэдр и т.д.), считая это математическим воплощением божественного замысла. Эта красивая теория оказалась ошибочной: данные многолетних наблюдений Браге расходились с ней на 10%. Потратив годы на подгонку данных, Кеплер эмпирически вывел, что орбиты являются эллипсами, а квадрат периода обращения пропорционален кубу расстояния до Солнца. Он не имел физического объяснения этим законам — лишь век спустя Ньютон вывел их из закона всемирного тяготения.

Фактчек: Кеплер отбросил модель платоновых тел из-за расхождения с данными Браге на 10% — $\triangle$: Расхождение рецензентов. Gemini подтверждает цифру, Claude уточняет: решающим было отклонение в 8 угловых минут положения Марса от круговой орбиты, а модель платоновых тел расходилась на несколько процентов, и Кеплер долго считал её верной. Канва бесспорна: эллипсы (1609), $T^2 \propto a^3$ (1619), вывод из тяготения — Ньютон (1687).

Кеплер действовал подобно LLM с высокой температурой 04:22: Он десятилетиями генерировал случайные, порой абсурдные гипотезы (вплоть до поиска связи орбит с музыкальными нотами и астрологией). Этот метод сработал исключительно благодаря наличию верифицируемой базы данных. Если есть надежный способ проверки, даже случайный перебор сотен закономерностей способен двигать фундаментальную науку вперед.

В истории науки произошел сдвиг парадигм от теории к большим данным 07:44: Классический научный метод подразумевал формулирование гипотезы и последующий сбор данных для ее проверки. В XX веке появились компьютерные симуляции, а затем наступила эпоха больших данных. Теперь процесс развернулся: сначала собираются гигантские массивы информации, а затем статистика и машинное обучение извлекают из них скрытые закономерности, формулируя гипотезы постфактум.

Эмпирический поиск закономерностей на малых данных часто приводит к ложным выводам 10:30: Астроном Иоганн Боде, вдохновившись Кеплером, вывел формулу расстояний между планетами на основе всего шести точек. Открытие Урана и Цереры идеально совпало с его предсказаниями, что вызвало восторг в научном мире. Однако открытие Нептуна полностью разрушило правило Боде, доказав, что это было лишь числовое совпадение.

Фактчек: Правило Боде подтвердилось Ураном и Церерой, разрушено Нептуном — $\square$: Предсказание $\sim 19,6$ а.е. для Урана (факт 19,2) и $\sim 2,8$ для Цереры (2,77), но $\sim 38,8$ для Нептуна при реальных 30,1. Классическая подгонка под шесть точек.

Фактчек: ИИ решил ~ 50 из ~ 1100 задач Эрдёша, дальше плато — $\triangle$: Gemini подтверждает по новостям начала 2026, Claude — с оговоркой, GPT не считает проверяемым. Это счёт erdosproblems.com, где значительная часть «решений» — найденные ИИ старые результаты в литературе; автономно решённых заметно меньше.

ИИ и кризис верификации в современной науке

ИИ снизил стоимость генерации научных идей почти до нуля 12:08: Это похоже на то, как интернет обнулil стоимость коммуникации. Однако само по себе это не создает изобилия полезных открытий. Главным узким местом науки стала оценка и верификация теорий. Традиционная система рецензирования (peer review) захлебывается от потока сгенерированных ИИ статей, не

справляясь с отделением реальных прорывов от наукообразного мусора в промышленных масштабах.

Оценить истинную ценность научной идеи в моменте практически

невозможно 15:21: Успех концепции зависит от будущего контекста и социальных факторов. Архитектура Transformer стала стандартом для LLM, но на ее месте могла оказаться другая технология, если бы общество стандартизировало ее первой. Десятичная система счисления победила римскую не из-за уникальных математических свойств числа 10, а из-за социальной инерции и адаптации под нее всей вычислительной инфраструктуры.

Исторически верные теории на начальном этапе часто выглядят хуже

ошибочных 17:33: Гелиоцентризм Аристарха отвергался античными греками из-за отсутствия видимого параллакса звезд. Теория Коперника изначально давала менее точные предсказания, чем геоцентрическая модель Птолемея, которая дорабатывалась тысячелетие. Прогресс часто требует не добавления новых теорий, а удаления укоренившихся базовых допущений, вроде аристотелевской идеи о том, что объекты стремятся к состоянию покоя.

Фактчек: Гелиоцентризм Аристарха отвергли из-за отсутствия параллакса; Коперник предсказывал хуже Птолемея — $\triangle$: Про параллакс верно (измерен лишь в 1838). Про Коперника — спорно: его модель с круговыми орбитами была примерно *не точнее* птолемеевой, а не заметно хуже; преимущество дал только Кеплер.

Фактчек: Гаусс вручную нашёл плотность простых $\sim 1/\ln x$ — $\square$: Делал это подростком (1792–93) по таблицам блоками по тысяче; позже стало теоремой о распределении простых.

Успех научной теории зависит от искусства коммуникации, а не только от

данных 23:09: Дарвин убедил мир в теории эволюции во многом благодаря тому, что писал на простом английском языке без формул, создав убедительный нарратив, несмотря на отсутствие знаний о механизмах наследственности. Ньютон, напротив, писал на латыни, изобретал новый математический аппарат и скрывал свои идеи от конкурентов, из-за чего его труды стали понятны массам лишь спустя десятилетия благодаря популяризаторам.

Фактчек: Вся криптография базируется на псевдослучайности простых — $\triangle$: Claude — преувеличение, GPT — $\square$. RSA стоит на вычислительной трудности факторизации, ECC — на дискретном логарифме, есть решётки и хеши; эвристика Крамера полезна теории чисел, но стойкость шифров из неё не следует.

Астрономия исторически лидирует в извлечении максимума информации

из скудных данных 26:30: Сбор космических наблюдений крайне сложен, поэтому астрономы стали экспертами в поиске скрытых сигналов. Этот навык

делает астрофизиков с PhD одними из самых востребованных кандидатов для найма в квантовые хедж-фонды.

Фактчек: Сейчас формализованы только доказательства в аксиоматике ZFC — □: Lean основан на зависимой теории типов, не на ZFC; существуют и другие формальные основания. Проверка Lean гарантирует выводимость записанного, но не то, что записано именно нужное утверждение.

[28:06-29:12] Рекламная интеграция Jane Street (задача по восстановлению слоев ResNet).

Социология науки может использовать нестандартные метрики для оценки качества исследований 29:12: Чтобы измерить, насколько внимательно ученые читают цитируемые ими работы, исследователи отслеживали распространение мелких опечаток в ссылках. Копирование ошибки из статьи в статью явно указывает на то, что автор вставлял ссылку вслепую, не читая первоисточник. Подобные следы в данных могут помочь автоматизировать оценку научного прогресса.

Фактчек: Математика исторически была чисто теоретической, эксперимент принесёт ИИ — △: Численные эксперименты — от таблиц простых до компьютерных исследований динамических систем — существовали задолго до ИИ. ИИ меняет масштаб, а не создаёт экспериментальную математику с нуля.

ИИ в математике: от решения задач к изменению парадигмы

ИИ-инструменты решили около 50 из 1100 открытых математических проблем Эрдёша, но прогресс замедлился 30:50: Наступило плато, так как ИИ собрал все «низковисящие плоды». Попытки натравить передовые модели на все оставшиеся задачи одновременно не принесли новых автономных решений. Оставшиеся проблемы требуют сложного многошагового взаимодействия между людьми и различными ИИ-агентами (генерация стратегии, критика, написание кода, поиск по литературе).

Фактчек: Реальная вероятность успеха ИИ на сложной задаче — 1-2% — □: Оценка Тао, не измерение. Универсальной цифры нет — зависит от набора задач, модели, бюджета попыток; сам тезис про ошибку выжившего в корпоративных кейсах обоснован.

Фактчек: Теорема о четырёх красках доказана перебором, элегантного доказательства нет — □: Appel и Haken (1976) — 1936 конфигураций, Робертсон и др. (1997) — 633, формализация в Coq (2005). Концептуального «человеческого» доказательства до сих пор нет.

Текущие модели работают как прыгающие роботы в темноте 32:27: Они могут подпрыгнуть на два метра и преодолеть низкие стены, недоступные человеку. Но они не способны зацепиться за уступ, закрепиться на нем и

продолжить восхождение. ИИ либо решает задачу с одной попытки, либо полностью проваливается. Модели крайне плохо справляются с созданием частичного прогресса или определением промежуточных этапов, на которых стоит сфокусироваться.

Фактчек: ИИ не способен на частичный прогресс и адаптацию, только перебирает с нуля — □: GPT ставит □: модели исправляют решение по обратной связи, агентные системы сохраняют промежуточные результаты. Это наблюдение Тао о текущей практике в математике, а не свойство архитектуры — категоричное «не способен» преувеличено.

ИИ превосходит человека в широте охвата, тогда как люди сильны в глубине 34:41: Как только ИИ достигает определенного уровня компетенции, он может мгновенно применить этот навык к миллионам доступных задач. Научный процесс должен быть перестроен: ИИ следует поручать массовое картирование новых областей и решение рутинных задач, чтобы выявлять «островки сложности» для глубокой проработки людьми-экспертами.

ИИ совершит революцию в экспериментальной части математики 39:42: Исторически математика была чисто теоретической наукой. Вместо того чтобы вручную решать отдельные задачи ради понимания процесса, математики смогут собирать масштабные данные о том, какие методы работают, а какие нет. Появится возможность тестировать тысячи гипотез одновременно, переходя к индустриальному масштабу исследований.

ИИ решает задачи путем комбинации существующих методов, но не создает новые 41:21: Большинство ИИ-решений проблем Эрдеша сводилось к применению малоизвестных техник к задачам, по которым почти не было литературы. ИИ отлично справляется с перебором стандартных подходов, часто допуская меньше ошибок, чем человек. Однако, когда стандартные методы не работают и требуется изобрести концептуально новый аппарат (что составляет суть статей в топовых журналах), ИИ оказывается бесполезен.

Реальная вероятность успеха ИИ на любой конкретной сложной задаче составляет всего 1-2% 44:40: Иллюзия невероятной эффективности создается за счет масштаба и ошибки выжившего: компании публикуют только успешные кейсы, игнорируя миллионы провалов. Для объективной оценки прогресса научному сообществу необходимо создать стандартизированные наборы сложных задач, независимые от корпоративного пиара.

Влияние ИИ на продуктивность и природу интеллекта

ИИ уже ускоряет написание математических статей в несколько раз 46:53: Прогноз Тао от 2023 года о том, что к 2026 году ИИ станет надежным соавтором, сбывается. Модели автоматизируют побочные задачи: глубокий поиск по литературе, генерацию сложного кода для численных проверок, создание

графиков и форматирование текста. При этом ядро работы — решение самой сложной части математической проблемы — по-прежнему выполняется человеком с ручкой и бумагой.

Существует фундаментальная разница между «искусственной изобретательностью» и подлинным интеллектом 49:40: Человеческий интеллект проявляется в адаптивности: при совместном решении задачи люди тестируют идею, видят ее недостатки, модифицируют подход и кумулятивно выстраивают путь к решению. ИИ на это не способен: он использует метод грубого перебора, повторяя попытки с нуля, и забывает весь полученный опыт при запуске новой сессии.

[51:49-53:00] Рекламная интеграция Labelbox.

Некоторые великие проблемы могут быть решены ИИ исключительно методом грубой силы 53:00: Теорема о четырех красках уже была доказана компьютером через перебор тысяч вариантов, и концептуально элегантного решения для нее до сих пор нет. Если гипотеза Римана будет доказана (или опровергнута) гигантским нечитаемым машинным вычислением, это станет большим разочарованием, так как математики ценят такие задачи именно за необходимость создания новых теорий, связывающих разные области науки.

Формальные языки доказательств (Lean) породят профессию математиков-рефакторов 56:17: Lean позволяет атомарно изучать каждый шаг сгенерированного ИИ решения. В будущем специалисты будут брать гигантские, нечитаемые машинные доказательства, проводить над ними абляционный анализ (удалять части для проверки их значимости), сжимать их и извлекать концептуально новые математические идеи, скрытые в машинном коде.

Математике остро не хватает полуформального языка для описания стратегий и оценки гипотез 59:32: Сейчас формализованы только строгие дедуктивные доказательства (аксиоматика ZFC). Если бы удалось создать систему, которая полуавтоматически оценивает вероятность истинности новой идеи на основе данных и эвристик (без возможности «взломать» метрику через обучение с подкреплением), это открыло бы путь для ИИ к генерации действительно глубоких концепций.

Теорема о распределении простых чисел — классический пример дата-ориентированной гипотезы 1:03:02: Гаусс вручную вычислил десятки тысяч простых чисел и заметил статистическую закономерность: их плотность падает обратно пропорционально натуральному логарифму. Это наблюдение породило аналитическую теорию чисел и эвристическую модель, в которой простые числа рассматриваются как псевдослучайное множество. На этой вере в отсутствие скрытых паттернов у простых чисел базируется вся современная криптография.

Фактчек: Опечатки в ссылках как метрика чтения первоисточников — []:

Исследование Симкина и Ройчоудхури «Read before you cite!» (2003): по копированию ошибок ~70–90% цитирующих оригинал не читали.

[1:08:39-1:09:42] Рекламная интеграция Mercury.

Обучение, серендипность и будущее профессии

Написание текстов — лучший способ закрепить знания при изучении

новых областей 1:09:42: Тао применяет подход «лисы» (знать понемногу обо всем), но с элементами обсессивности: если кто-то решил задачу неизвестным ему методом, он испытывает дискомфорт, пока не разберется в трюке. Формулирование мыслей в блоге помогает не забыть концепцию спустя полгода, а сам процесс написания служит приятным отдыхом от административной рутины.

Жесткая оптимизация расписания убивает серендипность, критически

важную для науки 1:13:29: Переход на удаленку во время COVID сохранил количество встреч, но уничтожил случайные разговоры в коридорах. Раньше, листая физический журнал в библиотеке, ученый мог случайно наткнуться на прорывную статью по соседству. Поисковики и ИИ выдают ровно то, что вы просите, лишая процесс полезного шума, отвлечений и незапланированных открытий.

В течение десятилетия ИИ полностью автоматизирует рутину

математиков 1:17:02: То, на что сегодня уходит основная часть времени студентов, будет делаться машинами. Это нормальный исторический процесс: сто лет назад математики вручную решали дифференциальные уравнения для физиков, а сегодня это делает софт за секунды. Профессия не исчезнет, она перейдет на новый уровень абстракции, как генетика перешла от ручного распутовывания хромосом к автоматизированному секвенированию геномов.

ИИ открывает нестандартные пути в науку, минуя долгие годы получения

PhD 1:21:07: Начинающим математикам следует готовиться к эпохе непредсказуемых изменений и сохранять максимальную адаптивность.

Традиционное образование остается важным, но благодаря инструментам вроде Lean и LLM, даже старшеклассники теперь могут вносить реальный вклад в передовые математические исследования.

ЛИЧНО ДЛЯ ТЕБЯ

Верификация дорожает, генерация обесценилась: твой режим «AI пишет, я валидирую» — ровно та роль, которая растёт в цене. Вкладывая время не в промпты, а в тесты, инварианты и граничные случаи, заданные независимо от модели — тесты, написанные ею же по своей реализации, закрепляют её

ошибки. Где есть надёжный чекер, даже тупой перебор двигает дело (Кеплер + данные Браге).

Не жди от модели частичного прогресса на большой задаче: режь её на «уступы», за которые модель может зацепиться с одной попытки, и связывай куски сам. Заведи свой набор из 10–20 повторяющихся задач и считай успехи, стоимость и время ручных правок — это объективнее пресс-релизов «ИИ решил X».

Оставляй в расписании шум: жёсткая оптимизация дня (только целевой поиск, только нужные встречи) убивает случайные находки. Хотя бы час в неделю бесцельного чтения вне своей области — физика, музыка, чужие статьи.

Записывай разобранное текстом: после каждой новой темы — короткое объяснение механизма, пример и границы применимости без подглядывания, затем сверка с источником. Иначе через полгода уходит; активное воспроизведение вскрывает дыры в понимании.

КОНСЕНСУС РЕЦЕНЗЕНТОВ

Все сошлись на правиле Боде, теореме о четырёх красках, Гауссе, параллаксе Аристарха и на практическом выводе «верификация важнее генерации». Расхождения: Gemini подтверждает «10%» Кеплера и число задач Эрдёша, Claude и GPT считают первое неточным, второе — непроверяемым; тезис о криптографии Claude смягчает до преувеличения, GPT отвергает; GPT единственный оспаривает утверждения о неспособности ИИ к частичному прогрессу, о ZFC и о «чисто теоретической» математике.