

How AI is changing math and science forever — Как ИИ навсегда меняет математику и науку

По ссылке · 2026-09-03 · https://www.youtube.com/watch?v=svl_1upFpQo



Теренс Тао, один из ведущих математиков современности и лауреат Филдсовской премии, анализирует трансформацию научного метода под влиянием искусственного интеллекта. ИИ переводит науку от ручного труда к автоматизированной генерации гипотез и доказательств.

Главная проблема современного этапа заключается не в способности ИИ решать задачи, а в отсутствии у него понимания физической реальности и логики первых принципов. Это создает феномен «несварения доказательств» (proof indigestion), когда научное сообщество не успевает проверять и осмыслять огромный объем генерируемых машиной данных. Будущее науки — в симбиозе, где ИИ выполняет рутинную генерацию и поиск паттернов, а человек выступает в роли куратора, придающего открытиям смысл и структуру.

Эволюция научных парадигм

От теории к автоматизации 01:37: Исторически наука опиралась на две парадигмы: теорию (например, законы Ньютона) и эксперимент. Математика была почти полностью теоретической, за редкими исключениями вроде Карла Гаусса, который вручную вычислил 100 000 простых чисел для создания теоремы о распределении простых чисел. Позже появились компьютерные симуляции, позволившие моделировать сложные явления (например, ураганы), и анализ больших данных (Big Data), где законы извлекаются из петабайтов информации без проведения прямых экспериментов.

Фактчек: Гаусс вручную вычислил 100 000 простых чисел и так создал теорему о распределении простых — Δ : все три рецензента сходятся, что формулировка искажена. Гаусс подростком работал с уже готовыми таблицами (до миллионов), подсчитывая простые в интервалах по 1000, и на этих данных *угадал* асимптотику $\pi(x) \approx x/\ln x$ — это была гипотеза; теорему доказали Адамар и Валле-Пуссен в 1896 году, через 40 лет после его смерти.

Масштабирование научного процесса 02:58: ИИ трансформирует все существующие методы. Если раньше ученые вручную проводили эксперименты, писали код для симуляций или анализировали данные, то теперь автоматизированные лаборатории, агенты-программисты и системы автоматического доказательства теорем делают это на порядки быстрее. ИИ способен генерировать и проверять гипотезы в масштабах, недоступных ни одному исследовательскому коллективу.

Ценность медленного процесса 03:58: Скорость ИИ имеет обратную сторону. Медленный, ручной процесс решения задачи — расчеты на бумаге, отладка кода, полевые эксперименты — часто приводит к неожиданным инсайтам. Ученый замечает новые феномены, видит скрытые связи с другими областями и формирует глубокое понимание проблемы. ИИ, выдающий готовый ответ, лишает исследователя этого побочного, но критически важного продукта научного поиска.

Механика и парадоксы современных ИИ-моделей

Предиктивный анализ вместо логики 05:25: Современный ИИ базируется на машинном обучении, суть которого — поиск паттернов в данных. Это усложненная версия регрессионного анализа: система получает множество входных и выходных данных и строит многомерную кривую, описывающую их зависимость. Большие языковые модели (LLM) работают по тому же принципу, просто предсказывая следующее слово в предложении на основе триллионов параметров.

Фактчек: LLM предсказывают следующее слово на основе триллионов параметров — Δ : путаница масштабов. Триллионы — это токены обучения; параметров у фронт-моделей сотни миллиардов, у MoE суммарно ~1–2 трлн при десятках миллиардов активных.

Фактчек: «Это просто усложнённая регрессия, значит, логики и понимания нет» — Δ : обучение действительно оптимизирует предсказание токенов, но вывод об отсутствии внутренней логики не следует автоматически — модели усваивают алгоритмические и структурные закономерности. Реальная проблема не в «отсутствии понимания», а в отсутствии гарантий корректности и плохой калибровке уверенности. Метафора «многомерная кривая» корректна как упрощение, но не объясняет in-context learning, цепочки рассуждений и RL-дообучение.

Усвоение скрытых структур 08:01: LLM успешно решают задачи не потому, что понимают их суть, а потому, что человеческий язык (включая язык математики) содержит глубокие неписанные правила. Подобно ребенку, усваивающему грамматику без заучивания правил, ИИ впитывает структуру математических доказательств из обучающей выборки. Это позволяет моделям выдавать правильные ответы на арифметические примеры и генерировать логичные шаги решений.

Отсутствие привязки к реальности 09:07: ИИ не опирается на физические законы или строгую логику первых принципов. Он лишь имитирует интеллектуальную дискуссию, опираясь на статистику. Из-за этого модели способны генерировать правдоподобные, но абсолютно ошибочные доказательства (галлюцинации). ИИ сейчас напоминает эрудированного, но слегка нетрезвого коллегу: он может выдать гениальную идею, но каждый его шаг требует жесткой верификации.

Фактчек: Модели генерируют правдоподобные, но неверные доказательства —

□: текстовая связность не равна доказательной корректности; надёжная проверка достигается формализацией в Lean/Coq/Isabelle, где ядро проверяет каждый шаг.

Риск переобучения (Overfitting) 16:44: При решении сложных задач ИИ склонен к переобучению. Модель может создать невероятно сложную конструкцию, которая идеально описывает конкретный набор тренировочных данных, но не имеет никакого отношения к реальным физическим или математическим законам и не способна экстраполироваться на новые данные.

Историческая аналогия: путь Кеплера к законам движения

Красивая, но ошибочная теория 13:43: Прогресс в науке редко бывает линейным. Иоганн Кеплер, пытаясь объяснить орбиты шести известных тогда планет, создал модель на основе пяти вложенных друг в друга платоновых тел (куб, тетраэдр и т.д.). Это была эстетически совершенная геометрическая теория, которая казалась идеальным объяснением устройства Солнечной системы.

Фактчек: Кеплер построил модель из пяти платоновых тел, столкнулся с данными Тихо Браге и отказался от неё в пользу эллипсов — △: канва верна (Mysterium Cosmographicum 1596, орбита Марса, Astronomia Nova 1609), но ключевая деталь искажена — Кеплер никогда не отбросил модель платоновых тел и развивал её до Harmonices Mundi (1619) параллельно с законами движения. Мораль «жесткие данные убили красивую теорию» звучит лучше, чем работает исторически.

Фактчек: Overfitting как объяснение ошибок ИИ в доказательствах — □: метафора, а не механизм. Доказательство либо валидно, либо нет; «переобучение» буквально относится к подгонке численных моделей, а не к выводу теорем.

Столкновение с данными 14:29: Когда Кеплер попытался совместить свою элегантную теорию с высокоточными астрономическими наблюдениями Тихо Браге, модель рассыпалась. Данные неумолимо показывали, что орбиты планет (в частности, Марса) не могут быть идеальными кругами, вписанными в правильные многогранники.

Итеративный поиск истины 14:56: Отказавшись от изначальной идеи, Кеплер потратил годы на перебор других форм и в итоге открыл, что планеты движутся по эллипсам. Этот хаотичный процесс — когда структурированная, но неверная теория корректируется жесткими данными — демонстрирует реальный путь науки. ИИ сегодня находится на схожем этапе: он генерирует множество несовершенных моделей, которые требуют столкновения с реальностью для кристаллизации истинного знания.

Текущие возможности ИИ в математике

Стремительный рост компетенций 12:23: За последние несколько лет ИИ прошел путь от решения задач средней школы до уровня математических олимпиад. Недавно модели начали успешно справляться с задачами квалификационных экзаменов в аспирантуру и даже решать старые открытые проблемы (например, опровержение гипотезы Пауля Эрдёша о планарных графах).

Фактчек: ИИ опроверг гипотезу Эрдёша о планарных графах — $\triangle$: **главное расхождение рецензентов.** Gemini уверенно подтверждает («май 2026, внутренняя модель OpenAI, гипотеза о единичных расстояниях»), но датировка из будущего сама по себе делает это подтверждение ненадёжным. Claude и GPT указывают на другое: осенью 2025 громкий эпизод «GPT-5 решил 10 проблем Эрдёша» оказался пиар-провалом — модель нашла *уже существующие* публикации, а не новые доказательства, и Тао публично это разбирал. Без точного названия задачи и статуса рецензирования считать факт установленным нельзя.

Бенчмарк First Proof 13:48: Для оценки реальных возможностей ИИ был создан тест First Proof. Ведущим моделям предложили 10 свежих, еще не опубликованных математических проблем исследовательского уровня. Лучшие системы смогли решить 5-6 задач из 10. Однако их решения часто неструктурированы: ИИ может исписать страницы, доказывая тривиальные вещи, и при этом проигнорировать или скомкать действительно сложный концептуальный шаг доказательства.

Фактчек: Бенчмарк First Proof: 10 свежих неопубликованных задач, лучшие модели решили 5-6 — $\triangle$: Gemini подтверждает (запуск февраль 2026, результат 5-7), Claude и GPT не подтверждают конкретное название и долю 50-60%. Жанр реален (FrontierMath от Epoch AI, задачи с erdosproblems.com, Humanity's Last Exam), но на FrontierMath топ-модели держались заметно ниже. Требуется проверки по

первоисточнику, включая критерий «решила»: набросок, формально проверяемое доказательство или принятая экспертами работа — это три разные вещи.

Неспособность к оценке сложности 18:30: ИИ пока не умеет отличать рутинные вычисления от концептуальных прорывов. Для машины любой шаг — это просто генерация следующего токена. Человек, сталкиваясь с трудным этапом доказательства, останавливается и тратит на него максимум времени и текста. ИИ же распределяет усилия равномерно, из-за чего его доказательства тяжело читать и проверять.

Фактчек: Равномерное распределение усилий: ИИ не отличает тривиальный шаг от концептуального прорыва — □: это узнаваемый и воспроизводимый failure mode — страницы на очевидное и скомканный критический шаг, поданный тем же уверенным тоном.

Будущее науки: симбиоз человека и машины

Кризис «несварения доказательств» 20:11: Математическое сообщество столкнулось с феноменом proof indigestion. ИИ генерирует огромное количество потенциальных решений и доказательств. Раньше появление решения сложной проблемы было редким событием, на проверку которого мобилизовались все эксперты. Теперь поток сгенерированных текстов превышает физические возможности ученых по их верификации (peer review).

Фактчек: «Proof indigestion» — кризис, с которым столкнулось математическое сообщество — □: Gemini атрибутирует термин Тао (середина 2026), GPT считает это сильной метафорой, а не зафиксированным измеримым кризисом. Существенно то, что пересказ упустил главный ответ Тао на эту проблему — формальную верификацию в Lean/mathlib и AlphaProof-подобные системы: машинно-проверяемое доказательство снимает нагрузку с рецензентов и может, наоборот, удешевить верификацию.

Смена роли ученого 17:42: В будущем главной задачей математиков станет не столько решение проблем с нуля, сколько кураторство. Ученые будут выступать в роли редакторов: брать сырые, неструктурированные инсайты ИИ, отсеивать галлюцинации, выстраивать логику и переводить машинный результат в форму, понятную человеческому разуму и пригодную для обучения следующих поколений.

Ограничения в коллаборации 18:36: Текущие ИИ-инструменты не способны на полноценный диалог. Опытные соавторы-люди понимают друг друга с полуслова, обладают общим контекстом и могут подхватить мысль на лету. ИИ пока лишен такой ментальной синхронизации и памяти. Тем не менее, использование ИИ для рутинных задач (поиск по литературе, проверка гипотез, написание базового кода) уже сейчас ускоряет работу программистов и математиков в разы, освобождая время для глубокой концептуальной работы.

Фактчек: Использование ИИ уже ускоряет программистов и математиков в разы
— ⚠: для программистов эмпирика не подтверждает. RCT METR (июль 2025) на опытных OSS-разработчиках в знакомых репозиториях показал **замедление на 19%** при субъективном ощущении ускорения на 20%. Выигрыш реален на незнакомом коде, бойлерплейте и прототипах.

ЛИЧНО ДЛЯ ТЕБЯ

Узкое место — верификация, а не генерация: твой режим «AI пишет — я валидирую» и есть описанное кураторство, и упирается он в твою личную пропускную способность ревьюера. Инвестируй не в скорость генерации, а в дешёвую проверку: тесты, типы, property-based tests, инварианты, детерминированные прогоны, реплей входных данных, независимая модель-ревьюер. Это прикладной аналог Lean для твоего домена.

Замерь своё ускорение вместо того, чтобы его ощущать: данные METR показывают, что на знакомом коде опытный разработчик с ИИ может замедлиться, продолжая чувствовать ускорение. Один раз измерь время до рабочего результата с ИИ и без на своих типовых задачах — и планируй сроки по замеру, а не по ощущению.

Blindspot и медицинский контур: не полагайся на «правдоподобный ответ»: храни транскрипт, источники, версию промпта и модели; требуй цитат из исходного контекста; жёстко разделяй извлечение фактов и генерацию рекомендаций; добавь детерминированный слой правил/справочника поверх LLM и human approval перед любым действием.

Размечай сложные места сам: раз агент тратит одинаково токенов на тривиальное и на критическое, явно указывай в промпте, где концептуально трудный шаг, и требуй остановки с вопросом вместо гладкого текста. И не отдавай ИИ концептуальную часть — «побочные инсайты» медленного ручного процесса (продуктовые связи, ложные предпосылки, выбор верной задачи) остаются человеческими; ИИ сильнее всего в расширении пространства вариантов: edge cases, альтернативные реализации, тестовые сценарии.

КОНСЕНСУС РЕЦЕНЗЕНТОВ

Все три модели сошлись на том, что тезис про Гаусса упрощён до искажения (гипотеза, а не теорема), а концептуальная рамка — галлюцинации при формально связном тексте, неумение ИИ ранжировать сложность шагов, сдвиг роли учёного к кураторству — верна и практически применима. Главное расхождение — в фактуре свежих новостей: Gemini уверенно подтверждает и бенчмарк First Proof, и опровержение гипотезы Эрдёша, ссылаясь на события 2026 года, тогда как Claude и GPT эти конкретные цифры и формулировки не подтверждают и напоминают о разоблачённом эпизоде с «решёнными

проблемами Эрдёша»; дополнительно Claude и GPT считают, что пересказ упустил ключевое — формальную верификацию в Lean как выход из «несварения доказательств», и оспаривают тезис об ускорении разработчиков «в разы».