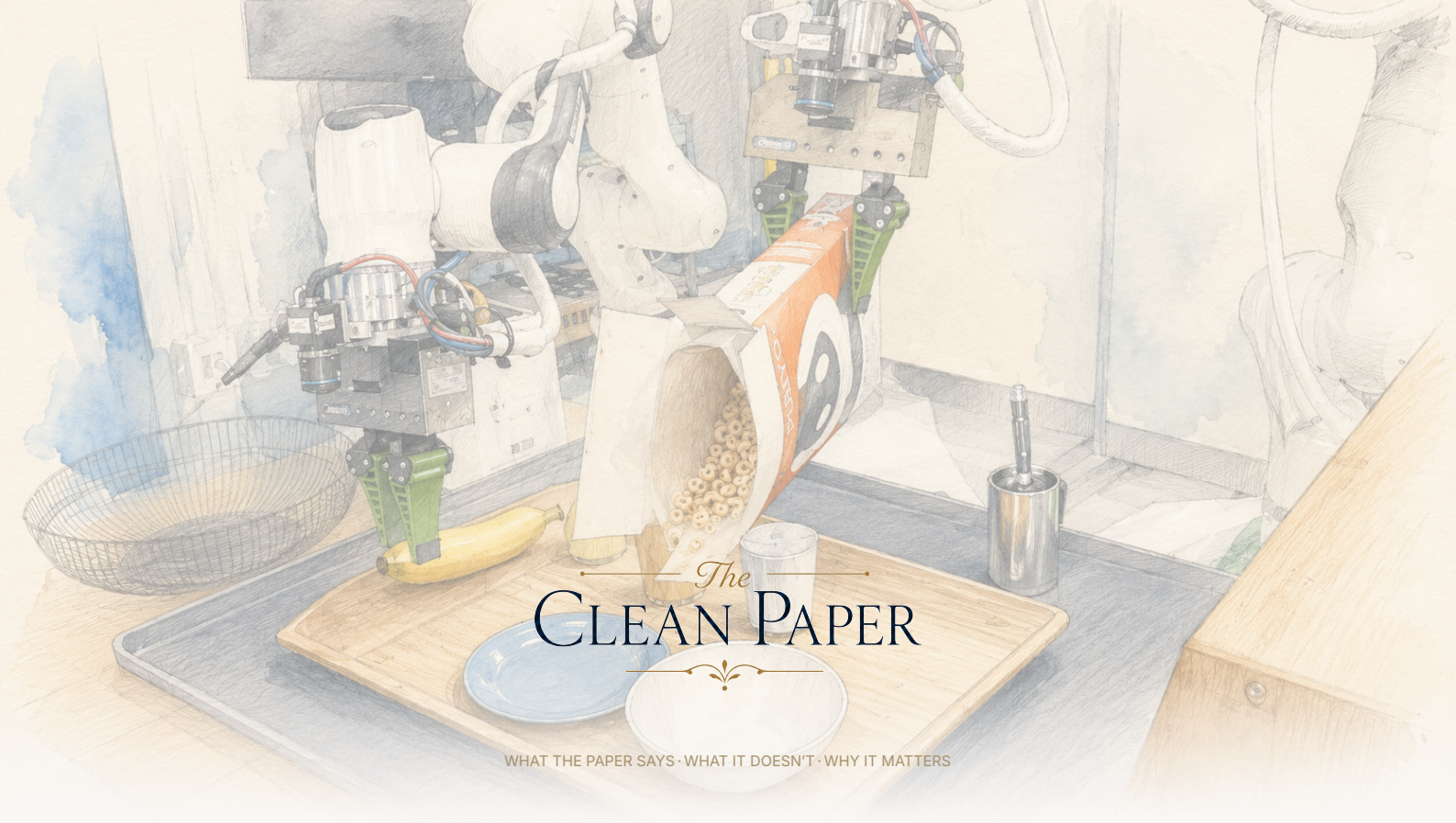


The Clean Paper

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

# Действительно ли большие «фундаментальные модели» роботов работают лучше? Осторожный ответ: умеренно да — и большинство исследований не способны надёжно это увидеть

*Toyota Research Institute обучил «большие модели поведения» — роботизированные политики с предварительным обучением примерно на 1 700 часах разнообразных манипуляций — и сравнил их с моделями одной задачи, обученными с нуля, с необычной строгостью: слепые рандомизированные испытания большого масштаба (~1 800 реальных и 47 000+ симуляционных) с настоящей статистикой. После дообучения на конкретной задаче большие модели в среднем работали лучше, требовали примерно в 3–5 раз меньше специфичных данных и были устойчивее к изменению условий; производительность плавно росла с объёмом pretraining-данных. Но без дообучения они не имели стабильного преимущества над моделями одной задачи, некоторые эффекты были настолько малы, что становились видимыми только в больших выборках, а обычный выбор нормализации данных значил больше архитектуры. Это сдержанная поддержка направления фундаментальных моделей для роботов — не универсальный робот, не zero-shot универсал и не «эмерджентный скачок» — плюс острое предупреждение, что значительная часть робототехники может измерять шум.*

---

Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

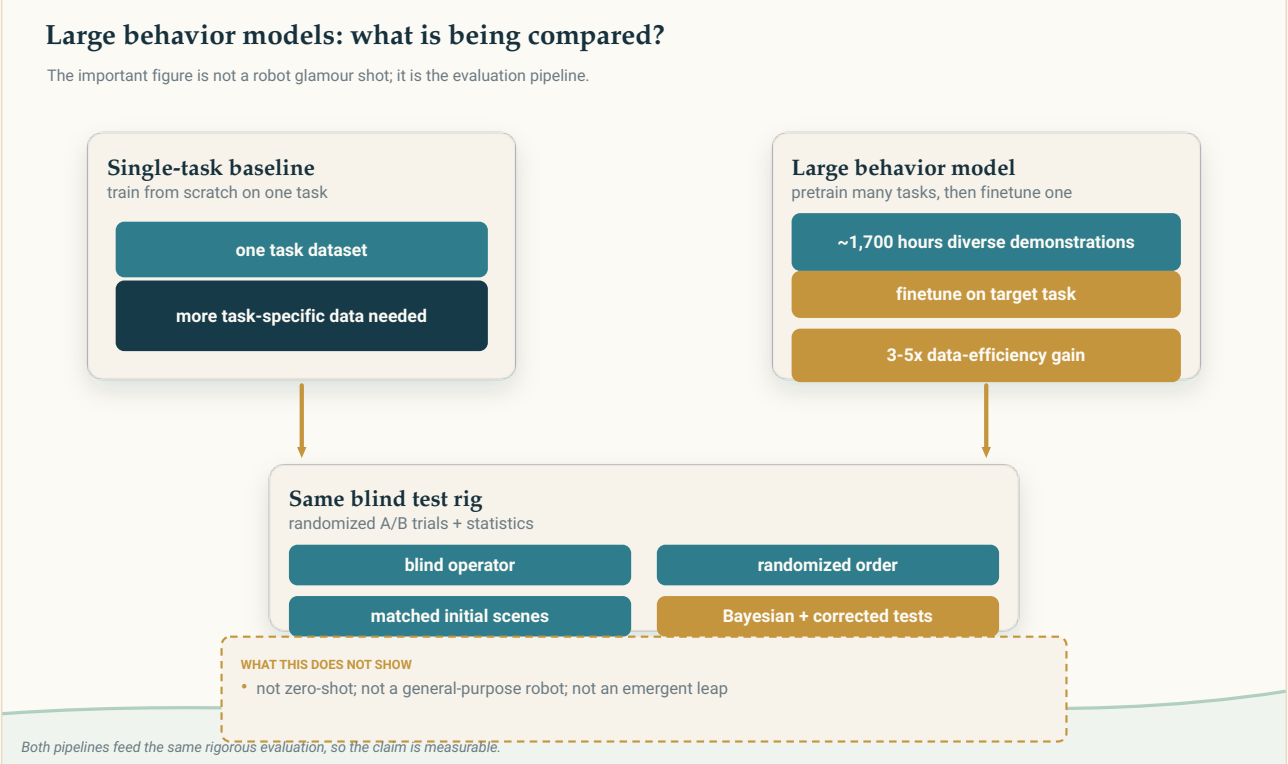


Paper: *A Careful Examination of Large Behavior Models for Multitask Dexterous Manipulation*, Toyota Research Institute Large Behavior Model Team — J. Barreiros, A. Beaulieu, et al.; senior authors incl. R. Ambrus, B. Burchfiel, S. Feng, H. Kress-Gazit (Cornell), R. Tedrake, Science Robotics (2026); preprint arXiv:2507.05331 · DOI: 10.1126/scirobotics.aea6201

# Статья о роботах, настоящая тема которой — честность

Робототехника переживает бум «фундаментальных моделей». Идея, заимствованная из языкового и визуального AI, очень привлекательна: вместо того чтобы обучать робота отдельно для каждой задачи, обучить одну большую **«большую модель поведения» (large behavior model, LBM)** на огромной и разнообразной коллекции демонстраций и получить систему, которая умеет многое и быстро адаптируется. Энтузиазм — и инвестиции — огромны. Заголовок пишется сам: *универсальные мозги для роботов уже здесь*.

Эта статья Toyota Research Institute интересна именно тем, что отказывается писать такой заголовок. Её настоящий вклад — не более эффектный робот. Это жёсткий взгляд на обманчиво простой вопрос — *действительно ли подход с большой моделью работает лучше и как мы вообще можем это знать?* — с уровнем статистической аккуратности, который, по словам самих авторов, необычен для отрасли. Результат — действительно полезное «да, но» и предупреждение, что значительная часть робототехники, возможно, измеряет шум.




Оба подхода проходят одинаковое слепое рандомизированное оценивание. Утверждение статьи не в том, что фундаментальные модели роботов уже являются zero-shot универсалами, а в том, что предварительное обучение может повышать эффективность данных и устойчивость, если измерять это надлежащим образом.

# Что сделали авторы

Они построили LBM в конкретном смысле: **диффузионные визуально-моторные политики** — Diffusion Transformer, который читает изображения с камер, короткую языковую инструкцию и положения суставов самого робота, а затем с частотой 10 Гц выдаёт короткие последовательности моторных команд. Эти модели **предварительно обучали примерно на 1 700 часах** демонстраций роботов — более 500 разных задач, собранных внутри лаборатории, плюс публичные наборы данных, — а затем **дообучали** на отдельных задачах. Во всех сравнениях базовой системой является **политика одной задачи, обученная с нуля** только на данных этой задачи.

Однако сердце статьи — именно *оценивание*, которое авторы представляют как главный результат. Чтобы не обмануть самих себя, они использовали:

- **Слепое рандомизированное A/B-тестирование** в реальном мире: человек, запускавший робота, не знал, какую политику тестируют, а порядок был случайным.
- **Контролируемые и воспроизводимые начальные условия**: перед каждой попыткой операторы выравнивали сцену по наложенному эталонному изображению.
- **Большое число попыток**: 50 реальных запусков на каждую задачу, политику и условие; 200 на задачу в симуляции. Всего — около **1 800 слепых реальных запусков и более 47 000 симуляционных**.
- **Надлежащую статистику**: байесовские оценки вероятности успеха и попарные проверки гипотез с поправками на множественные сравнения вместо визуального сопоставления перекрывающихся погрешностей. Они даже провели контроль качества на четверти попыток, оценённых людьми, чтобы измерить ошибки самого оценивания.

[video pending]

Сравнение моделей, сервирующих стол для завтрака: слева — базовая политика одной задачи, справа — LBM. Оба видео воспроизводятся со скоростью 1×. Это одна оцениваемая задача, а не доказательство универсальной автономности.

Вся эта машина оценивания и есть суть. Статья в целом утверждает: без неё невозможно отличить настоящее улучшение от удачи.

# Что они обнаружили

**Большие модели после дообучения в среднем превосходят модели одной задачи, обученные с нуля.** В совокупности задач LBM, предварительно обученная, а затем дообученная на конкретной задаче, надёжно опережала политику, обученную с нуля на той же задаче, и в симуляции, и в реальном мире; разница была статистически значимой. На отдельных задачах дообученная LBM была статистически не хуже или

лучше почти всегда: 3/3 реальных задач и 15/16 симуляционных.

**Самый большой и чистый выигрыш — эффективность данных.** Дообученная LBM достигала производительности модели, обученной с нуля, используя примерно **в 3–5 раз меньше специфичных для задачи данных**. В одной реальной задаче — сервировке стола для завтрака — LBM, дообученная лишь на **15%** демонстраций, превзошла политику, обученную с нуля на **100%**.

**Предварительное обучение больше всего помогает, когда условия меняются.** Когда тестовую среду намеренно отклоняли от тренировочных условий («distribution shift»), преимущество дообученной LBM *увеличивалось*. В одном наборе симуляций она статистически превосходила модель с нуля на 3 из 16 задач в обычных условиях, но на 10 из 16 при сдвиге распределения. Поскольку реальные среды всегда отличаются от тренировочных, эта устойчивость, вероятно, является практически самым важным результатом.

**Больше данных предварительного обучения помогало плавно.** Производительность стабильно росла с увеличением данных предобучения — **без внезапного скачка или «эмерджентного» перехода** в исследованном диапазоне. Полезно, предсказуемо, не драматично.

**А вот история про универсала без дообучения не подтвердилась.** Предварительно обученная LBM в режиме *zero-shot* — без специфичного дообучения на задаче — *не* превосходила политики одной задачи последовательно. Одна сеть могла выполнять много задач, но мечта «просто скажи ей, что делать» здесь не сбылась; отчасти авторы связывают это с хрупкостью небольшого языкового энкодера.

**И выигрыш был достаточно мал, чтобы его легко не заметить — или случайно нарисовать из шума.** Многие эффекты стали видимыми лишь благодаря большему, чем обычно, числу попыток и тщательным тестам. Авторы прямо пишут, что с учётом размера эффектов и шума **существует значительный риск, что многие статьи по робототехнике измеряют статистический шум**. Они также обнаружили, что будничное решение — как именно *нормализовать* данные — влияло на результат сильнее архитектурных изменений, а **ошибку** нормализации во время предобучения обнаружили только *после* завершения оцениваний.

## Что это, вероятно, означает

Защищённое прочтение: масштабное предварительное обучение на разнообразных роботизированных данных — действительно полезный компонент. Оно уменьшает количество данных, нужных для новой задачи, и делает политики устойчивее, когда мир не совпадает с тренировочными условиями. Это реально поддерживает направление, на которое ставит отрасль. Но преимущества *умеренны и условны*: преимущественно проявляются после дообучения и наиболее отчётливы в агрегированных результатах и под стрессом, а не означают появления готового универсального робота.

Более тихий и, возможно, более важный смысл — методологический. Статья фактически предлагает мерную линейку: показывает, сколько доказательств на самом деле нужно, чтобы надёжно заявить об улучшении роботизированной политики, и намекает, что значительная часть опубликованного энтузиазма опирается на слишком маленькие выборки. Такой корректив отрасли нужнее ещё одной модели.

## Чего это *не* доказывает

- Это **не универсальный робот**. Преимущества показаны для конкретной архитектуры (диффузионных политик), дообученной отдельно для каждой задачи, в контролируемых условиях и на демонстрациях телеоперации — а не для автономного робота, выполняющего произвольные новые задания по команде.
- Это **не подтверждает zero-shot использование**. Без дообучения большая модель не имела стабильного преимущества над базовыми моделями одной задачи.
- Это **не свидетельство «эмерджентного скачка»**. Масштабирование давало плавные улучшения; нет никакого разрыва, который поддерживал бы нарратив «и вдруг она стала способной».
- Числа **относительны и привязаны к лаборатории**. Абсолютные показатели успеха намеренно настраивали примерно к 50%, чтобы сделать сравнения чувствительнее; это не оценка реальной надёжности. И это одна архитектура из одной лаборатории.
- Работа **не объясняет, почему** конкретная политика успешна или неудачна; несколько отдельных задач, где большая модель работала *хуже*, сообщены, но не объяснены.
- Она **ничего не говорит о безопасности, автономности или развёртывании** вне тестовой установки.

## Насколько сильны доказательства?

Для центральных сравнительных утверждений — что дообученные LBM в агрегате превосходят модели с нуля, требуют в несколько раз меньше данных и лучше выдерживают сдвиг распределения — доказательства сильны *и* необычно хорошо контролируются: слепой дизайн, рандомизация, большие выборки, статистические тесты и проверка качества человеческого оценивания. Это редкий случай, когда методология достаточно надёжна, чтобы главные выводы можно было воспринимать почти буквально.

Честные оговорки авторы выдвигают сами. Их интервалы неопределённости учитывают случайность *оценивания*, но **не** случайность *обучения*: обучите ту же модель дважды — и можете получить существенно другую политику, а эта вариативность в статистику не входит. Реальные задачи имели по 50 попыток — достаточно для эффектов среднего размера, но мелкие могут остаться незамеченными. Языковое кондиционирование

использовало скромный энкодер, поэтому выводы о «просто скажите роботу, что делать» могут быть другими для более крупных систем. И есть откровенное сообщение о найденной постфактум ошибке нормализации. Ничто из этого не топит главные результаты, но это именно те вещи, которые, по аргументу статьи, отрасль часто заметает под ковёр.

И одно замечание об источнике в том же духе: этот материал опирается на препринт авторов. Нам не удалось получить опубликованную журнальную версию, поэтому мы не проверяли, есть ли между препринтом и финальным текстом изменения.

## Почему это важно

«Фундаментальные модели для роботов» — словосочетание, созданное для преувеличений, и такую работу легко неверно прочитать в обе стороны: как триумфальное *«работает!»* или как пренебрежительное *«всё раздуто»*. Точный ответ полезнее обоих: предварительное обучение на разнообразных данных даёт реальные, измеримые, но умеренные преимущества — прежде всего меньше данных на задачу и большую устойчивость — и масштабирование улучшает их предсказуемо.

Более глубокая причина важности в том, что статья направляет свою строгость на собственную отрасль. Показывая, что настоящие эффекты достаточно малы, чтобы исчезать при небрежном оценивании, а скучный выбор нормализации данных может значить больше хитрой новой архитектуры, она доказывает: значительная часть прогресса в обучении роботов требует более прочного измерения, прежде чем ему можно верить. Статья, которая тратит свой кредит доверия на охрану границы между результатом и желанием, делает нечто более редкое и ценное, чем возглавить таблицу лидеров.

## Краткий итог

Исследователи Toyota Research Institute обучили «большие модели поведения» — диффузионные политики для роботов, предварительно обученные примерно на 1 700 часах разнообразных данных манипуляций, — и сравнили их с политиками одной задачи, обученными с нуля, используя необычно строгий протокол: слепые рандомизированные испытания большого масштаба (≈1 800 реальных и более 47 000 симуляционных) с настоящей статистикой. После дообучения на конкретной задаче большие модели надёжно работали лучше в совокупности, требовали примерно **в 3–5 раз меньше специфичных данных** и были **устойчивее к изменению условий**, а производительность плавно росла с увеличением данных предобучения. Но **без дообучения** они не имели стабильного преимущества над моделями одной задачи; несколько эффектов были настолько малы, что стали видимыми лишь благодаря

большим выборкам, а обычный выбор нормализации данных влиял сильнее архитектуры. Это сильная и сдержанная поддержка направления роботизированных фундаментальных моделей — **не** универсальный робот, не zero-shot универсал и не «эмерджентный скачок» — плюс острое предупреждение, что значительная часть робототехники может измерять шум.

## Проверка без преувеличений

**Что показывает статья:** При строгом, слепом и статистически мощном оценивании (≈1 800 реальных + 47 000+ симуляционных запусков) диффузионные политики, предварительно обученные на многих задачах и затем дообученные (LBM), в агрегате превосходят политики одной задачи, обученные с нуля; достигают эквивалентной производительности примерно с в 3–5 раз меньшим количеством специфичных данных; лучше выдерживают сдвиг распределения; производительность плавно масштабируется с объёмом данных предобучения.

**Что правдоподобно, но не доказано:** Что эти преимущества перенесутся на значительно более крупные vision-language-action модели (их языковой энкодер был небольшим); что плавное масштабирование продолжится за пределами исследованного диапазона данных.

**Чего работа не показывает:** Универсального или zero-shot робота (без дообучения нет стабильного преимущества); какого-либо «эмерджентного» скачка возможностей; объяснения конкретных неудач на отдельных задачах; абсолютной реальной надёжности (успешность намеренно держали около 50% для чувствительности); ничего о безопасности или автономном развёртывании.

**Основные ограничения:** Статистика учитывает случайность оценивания, но не вариативность между отдельными запусками обучения; 50 реальных попыток на задачу могут не обнаруживать мелкие эффекты; одна архитектура и одна лаборатория; скромный языковой энкодер; ошибку нормализации данных нашли после оцениваний; анализ основан на препринте (опубликованную версию не сверяли).

**Сколько доверия стоит иметь обычному читателю?** Высокое в том, что многозадачное предварительное обучение плюс дообучение даёт реальные, умеренные преимущества — особенно в эффективности данных и устойчивости — и что их измерили необычно тщательно. Высокое в том, что это **не** универсальный или zero-shot робот и **не** эмерджентный скачок. Среднее относительно того, насколько эти преимущества будут масштабироваться на более крупные модели. И стоит всерьёз отнестись к предупреждению самих авторов: эффекты в отрасли достаточно малы, чтобы недостаточно мощные исследования могли сообщать шум. Уместная позиция: сдержанный оптимизм относительно подхода и здоровое недоверие к результатам robot-AI без такого статистического фундамента.

## Источники

arXiv:2507.05331
https://arxiv.org/abs/2507.05331

Peer-reviewed version (not retrieved) — Science Robotics, 10.1126/scirobotics.aea6201
https://doi.org/10.1126/scirobotics.aea6201

Project page
https://toyotaresearchinstitute.github.io/lbm1/