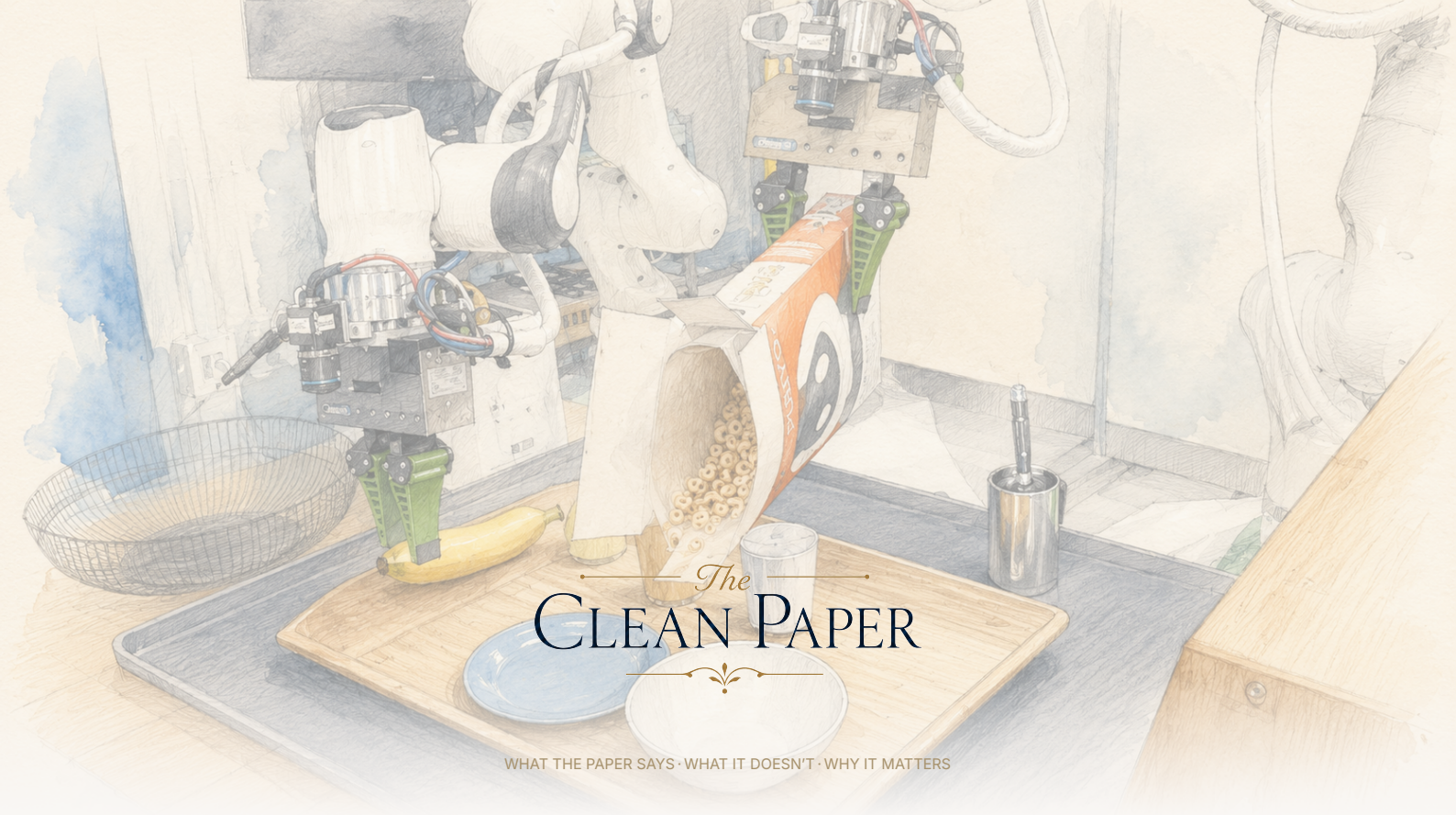


Работят ли големите роботни „foundation models“ по-добре? Внимателният отговор е: умерено да — и повечето изследвания не могат да разберат

Toyota Research Institute обучи „large behavior models“ — роботни политики, предварително обучени върху ~1 700 часа разнообразни манипулационни данни — и ги сравни с еднозадачни политики, обучени от нулата, с необичайна строгост: слепи, рандомизирани, мащабни изпитвания (~1 800 реални и над 47 000 симулационни) с истинска статистика. След дообучаване за конкретна задача големите модели средно се справят по-добре, нуждаят се от около 3–5× по-малко специфични данни и са по-устойчиви при промяна на условията; производителността нараства плавно с повече предварителни данни. Но без дообучаване не превъзхождат последователно еднозадачните модели, някои ефекти са толкова малки, че се виждат само при големите извадки, а обикновен избор за нормализация на данните има по-голям ефект от архитектурата. Това е измерена подкрепа за посоката на robot foundation models — не общ робот, не zero-shot generalist и не „емергентен скок“ — плюс ясно предупреждение, че голяма част от роботиката може да измерва шум.

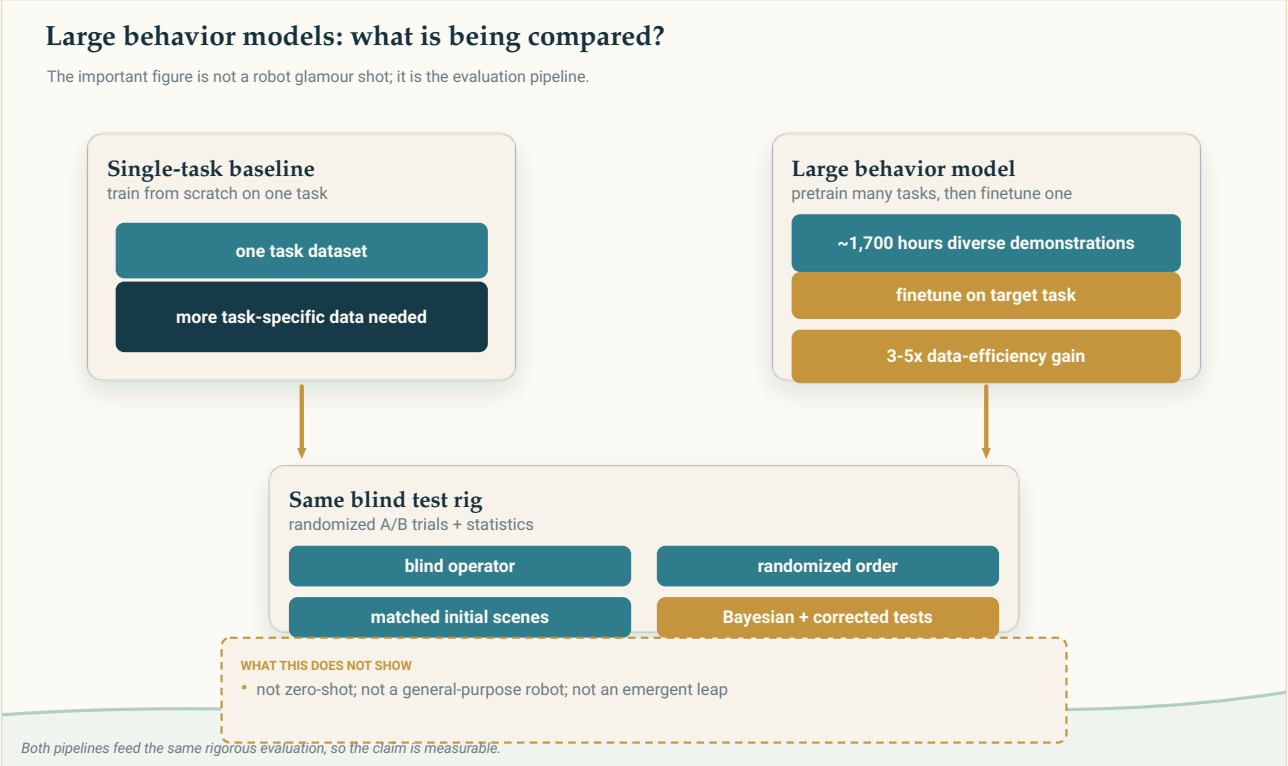
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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

Статия за роботи, чиято истинска тема е честността

Роботиката е в разгара на вълна от „фундаментални модели“. Идеята, заета от езиковите и визуалните AI системи, е примамлива: вместо роботът да се обучава по една задача наведнъж, обучавате един голям **„голям модел на поведение“ (large behavior model, LBM)** върху огромна и разнообразна купчина демонстрации и получавате система с широки способности и бърза адаптация. Ентусиазмът — и инвестициите — са огромни. Заглавието се пише само: *общите роботни мозъци вече са тук.*

Статията на Toyota Research Institute е интересна точно защото отказва да напише това заглавие. Истинският ѝ принос не е по-ефектен робот. Тя разглежда трудно един измамно прост въпрос — *големият модел действително ли работи по-добре и как въобще бихме разбрали?* — с ниво на статистическа грижа, което според самите автори е необичайно за областта. Резултатът е истински полезно „да, но“ и предупреждение, че значителна част от роботиката може да измерва шум.



И двата подхода преминават през една и съща сляпа, рандомизирана оценка. Твърдението на статията не е, че фундаменталните модели за роботи са универсални системи без дообучаване (zero-shot), а че предварителното обучение може да подобри ефективността на данните и устойчивостта, когато се измерва внимателно.

Какво направиха авторите

Те изграждат LBM в конкретен смисъл: **визуомоторни политики, базирани на дифузионни модели** — Diffusion Transformer, който чете изображения от камерата, кратка езикова инструкция и позициите на ставите на робота и извежда кратки поредици моторни команди при 10 Hz. Моделите са **предварително обучени върху приблизително 1 700 часа** роботни демонстрации — над 500 различни задачи, събрани вътрешно, плюс публични набори — и след това **дообучени** за отделните задачи. Сравнението навсякъде е с **еднозадачна политика, обучена от нулата** върху данните само за тази задача.

Сърцето на статията обаче е *оценяването*, което авторите третират почти като основен резултат. За да не се самозаблудят, те използват:

- **Сляпо, рандомизирано A/B тестване** в реалния свят — човекът, който управлява теста, не знае коя политика се изпитва, а редът е случаен.
- **Контролирани, възпроизводими начални условия** — операторите съпоставят сцената с изображение-подложка преди всеки опит.
- **Голям брой опити** — 50 реални изпълнения на задача, политика и условие; 200 на задача в симулация. Общо: около **1 800 слепи реални изпълнения и над 47 000 симулационни изпълнения**.
- **Истинска статистика** — байесови оценки на вероятността за успех и двойкови хипотезни тестове с корекции за множествени сравнения, вместо визуално сравнение на припокриващи се ленти на грешката. Те дори правят контрол на качеството върху една четвърт от оценените от хора опити, за да измерят грешката при оценяване.

[video pending]

Сравнение рамо до рамо на модели, които подреждат маса за закуска: вляво еднозадачната базова линия, вдясно LBM. И двата видеоклипа вървят на 1x скорост. Това е една оценявана задача, не доказателство за обща автономност.

Тази измервателна машина е самата идея. Цялата статия е аргумент, че без нея не можете да различите истинско подобрение от късмет.

Какво откриха

Дообучените големи модели превъзхождат еднозадачните модели от нулата — средно. Когато резултатите се обединят между задачите, LBM, който е предварително обучен и после дообучен за конкретна задача, надеждно превъзхожда политика, обучена от нулата върху същата задача, както в симулация, така и в реалния свят, и разликата е статистически значима. На ниво отделна задача дообученият LBM е

статистически поне толкова добър или по-добър почти във всеки случай (3/3 реални задачи, 15/16 симулационни).

Най-голямата и ясна победа е ефективността на данните. Дообучен LBM достига производителност, еквивалентна на модела от нулата, с приблизително **3–5× по-малко специфични за задачата данни**. В една реална задача — подреждане на маса за закуска — LBM, дообучен само с **15%** от демонстрациите, превъзхожда модел от нулата, обучен с **100%** от тях.

Предварителното обучение помага най-много, когато условията се променят. Когато тестовата среда е нарочно изместена спрямо условията на обучение („изместване на разпределението“), предимството на дообучения LBM *нараства*. В един симулационен набор той статистически превъзхожда модела от нулата на 3 от 16 задачи при нормални условия, но на 10 от 16 при изместване на разпределението. Тъй като реалните внедрявания винаги се отклоняват от тренировъчните условия, тази устойчивост може да е най-практично важната находка.

Повече предварителни данни помагат плавно. Производителността се повишава постепенно с добавянето на повече данни за предварително обучение — без внезапен скок или „емергентен“ пробив при тестваните мащаби. Полезно, предвидимо, недраматично.

Но историята за общ модел без дообучаване не се потвърждава. Предварително обучен LBM, използван *без дообучаване (zero-shot)* — не превъзхожда последователно еднозадачните политики. Една мрежа може да изпълнява много задачи, но мечтата „просто му кажи какво да прави“ не се реализира тук; авторите отдават част от проблема на крехкостта на малкия си езиков енкодер.

А печалбите са достатъчно малки, че лесно да бъдат пропуснати — или фалшиво „намерени“. Много ефекти стават видими едва с по-големите от обичайното извадки и внимателните тестове. Авторите казват директно, че при този размер на ефектите и този шум **има значителен риск много статии по роботика да измерват статистически шум**. Те установяват също, че един прозаичен избор — как данните се *нормализират* — влияе повече от промени в архитектурата, а **бъг** в нормализацията при предварителното обучение е открит едва *след* приключване на оценяванията.

Какво вероятно означава това

Защитимият прочит: мащабното предварително обучение върху разнообразни роботни данни е реално полезна съставка — нужни са по-малко данни за всяка нова задача и политиките са по-устойчиви, когато светът не съвпада с обучението. Това действително подкрепя посоката, на която областта залага. Но печалбите са *умерени и условни* — появяват се главно след дообучаване и са най-ясни в агрегат и при стрес — не пристигането на готов общ робот.

По-тихото и по-важно значение е методологично. Статията всъщност е мерителна

пръчка: показва колко доказателства са нужни, за да се направи надеждно твърдение за роботна политика, и подсказва, че много публикуван ентусиазъм се опира на твърде малко данни. Този коректив е по-нужен на областта от още един модел.

Какво *не* доказва това

- Това **не е общ робот**. Печалбите са показани за конкретна архитектура — дифузионни политики — дообучавана по задачи, в контролирани условия, от телеоперирани демонстрации, не за автономен робот, който изпълнява произволни нови задачи по команда.
- То **не** валидира **zero-shot** употреба. Без дообучаване големият модел не превъзхожда последователно еднозадачните базови линии.
- То **не** е доказателство за **„емергентен скок“**. Мащабирането подобрява резултатите плавно; няма прекъсване, което да подкрепи разказа „и изведнъж стана способен“.
- Числата са **относителни и лабораторни**. Абсолютните проценти на успех са нарочно настроени около ~50%, за да бъдат сравненията чувствителни; те не са мярка за надеждност в реалния свят, а работата е за една архитектура от една лаборатория.
- То **не** решава **защо** дадена политика успява или се проваля, а няколко конкретни задачи, на които големият модел се справя *по-зле*, са описани, но не обяснени.
- То не казва **нищо за безопасност, автономност или внедряване** извън тестовата установка.

Колко силни са доказателствата?

За централните сравнителни твърдения — че дообучените LBM превъзхождат базовите модели от нулата в агрегат, нуждаят се от няколко пъти по-малко данни и са по-устойчиви при изместване на разпределението — доказателствата са силни *и* необичайно добре контролирани: слепи, рандомизирани, с големи извадки, статистически тествани и с QA проверка на оценяването. Това е рядък случай, в който методологията е достатъчно здрава, за да приемем основните изводи почти буквално.

Честните уговорки са тези, които самите автори повдигат. Лентите на грешката улавят случайността на *оценяването*, но **не** случайността на *обучението* — обучете същия модел два пъти и може да получите осезаемо различна политика, а тази вариация не е в статистиката. Реалните задачи имат по 50 опита, достатъчно за средни ефекти, но малките могат да останат незабелязани. Езиковото условяване използва скромен енкодер, така че твърденията за „просто кажи на робота какво да прави“ може да изглеждат различно при по-големи системи. И има откровеното съобщение за бяг в нормализацията, открит постфактум. Нито едно от тези неща не потапя основния резултат, но са точно от онези детайли, които статията твърди, че областта често

замита.

Една бележка за източника, в същия дух: този текст е базиран на препринта на авторите. Не успяхме да получим публикуваната журнална версия, затова не сме проверили дали между препринта и окончателния текст има промени.

Защо това има значение

„Фундаментални модели за роботи“ е израз, създаден за преувеличение, а такава статия лесно може да бъде прочетена и в двете посоки — като триумфално *„работи!“* или пренебрежително *„прехвалено е“*. Точният прочит е по-полезен от двете: предварителното обучение върху разнообразни данни носи реални, измерими, но умерени ползи — най-вече по-малко данни на задача и повече устойчивост — и пътят се подобрява предвидимо с мащаба.

По-дълбоката причина да има значение е, че статията насочва строгостта си към собствената област. Като показва, че истинските ефекти са достатъчно малки, за да изчезнат при небрежно оценяване, и че скучен избор като нормализацията на данните може да надвие умна нова архитектура, тя аргументира, че голяма част от напредъка в обучението на роботи се нуждае от по-здраво измерване, преди да бъде вярван. Статия, която харчи доверието си за контрол на границата между резултат и желание, прави нещо по-рядко и по-ценно от това да оглави класация.

Кратко и ясно

Изследователи от Toyota Research Institute обучават „големи модели на поведение“ — роботни политики, базирани на дифузионни модели, предварително обучени върху ~1 700 часа разнообразни манипулационни данни — и ги сравняват с еднозадачни политики от нулата чрез необичайно строг протокол: сляп, рандомизиран, с големи извадки (≈1 800 реални и над 47 000 симулационни опита) и истинска статистика. След дообучаване за конкретна задача големите модели надеждно се справят по-добре в агрегат, нуждаят се от приблизително **3–5× по-малко специфични данни** и са **по-устойчиви при промяна на условията**, а производителността нараства плавно с повече предварителни данни. Но **без дообучаване** те *не* превъзхождат последователно еднозадачните модели, няколко ефекта са толкова малки, че само големите извадки ги разкриват, а един обикновен избор за нормализация на данните има по-голямо значение от архитектурата. Това е солидна, измерена подкрепа за посоката на фундаменталните модели за роботи — **не** общ робот, не универсална система без дообучаване и не „емергентен скок“ — плюс ясно предупреждение, че голяма част от роботиката може да измерва шум.

Проверка без излишен шум

Какво показва статията: При строг, сляп и статистически достатъчно мощен тест (≈1 800 реални + над 47 000 симулационни изпълнения) дифузионни политики, предварително обучени многозадачно и после дообучени (LBM), превъзхождат в агрегат еднозадачните политики от нулата, достигат еквивалентна производителност с ~3–5× по-малко специфични данни и са по-устойчиви при изместване на разпределението; резултатите се подобряват плавно с повече предварителни данни.

Какво е правдоподобно, но не е доказано: Че тези ползи се пренасят към много по-големи модели за зрение, език и действие (vision-language-action) (езиковият им енкодер е малък); че плавното мащабиране продължава отвъд тествания диапазон от данни.

Какво не показва: Общ или робот, работещ без дообучаване (без дообучаване → без последователно предимство); никакъв „емергентен“ скок в способностите; обяснение на конкретните провали по задачи; реална надеждност в абсолютни стойности (процентите на успех са настроени близо до 50% за чувствителност); нищо за безопасност или автономно внедряване.

Основни ограничения: Статистиката включва случайността на оценяването, но не на отделните тренировъчни пускания; 50 реални опита на задача могат да пропуснат малки ефекти; една архитектура и една лаборатория; скромен езиков енкодер; бъг в нормализацията на данните е открит след оценките; анализът е базиран на препринта (публикуваната версия не е проверена).

Колко уверен трябва да бъде общият читател? Висока увереност, че многозадачното предварително обучение плюс дообучаване дава реални, умерени ползи — особено за ефективността на данните и устойчивостта — и че те са измерени необичайно внимателно. Висока увереност, че това **не** е общ или робот, работещ без дообучаване и **не** е емергентен скок. Средна увереност колко далеч се мащабират ползите към по-големи модели. И си струва да се приеме сериозно собственото предупреждение на авторите: ефектите в областта са достатъчно малки, че недостатъчно мощни изследвания може просто да съобщават шум. Подходящата позиция е измерен оптимизъм към подхода и здравословен скептицизъм към резултати за роботизиран ИИ без такава статистическа опора.

Източници

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/