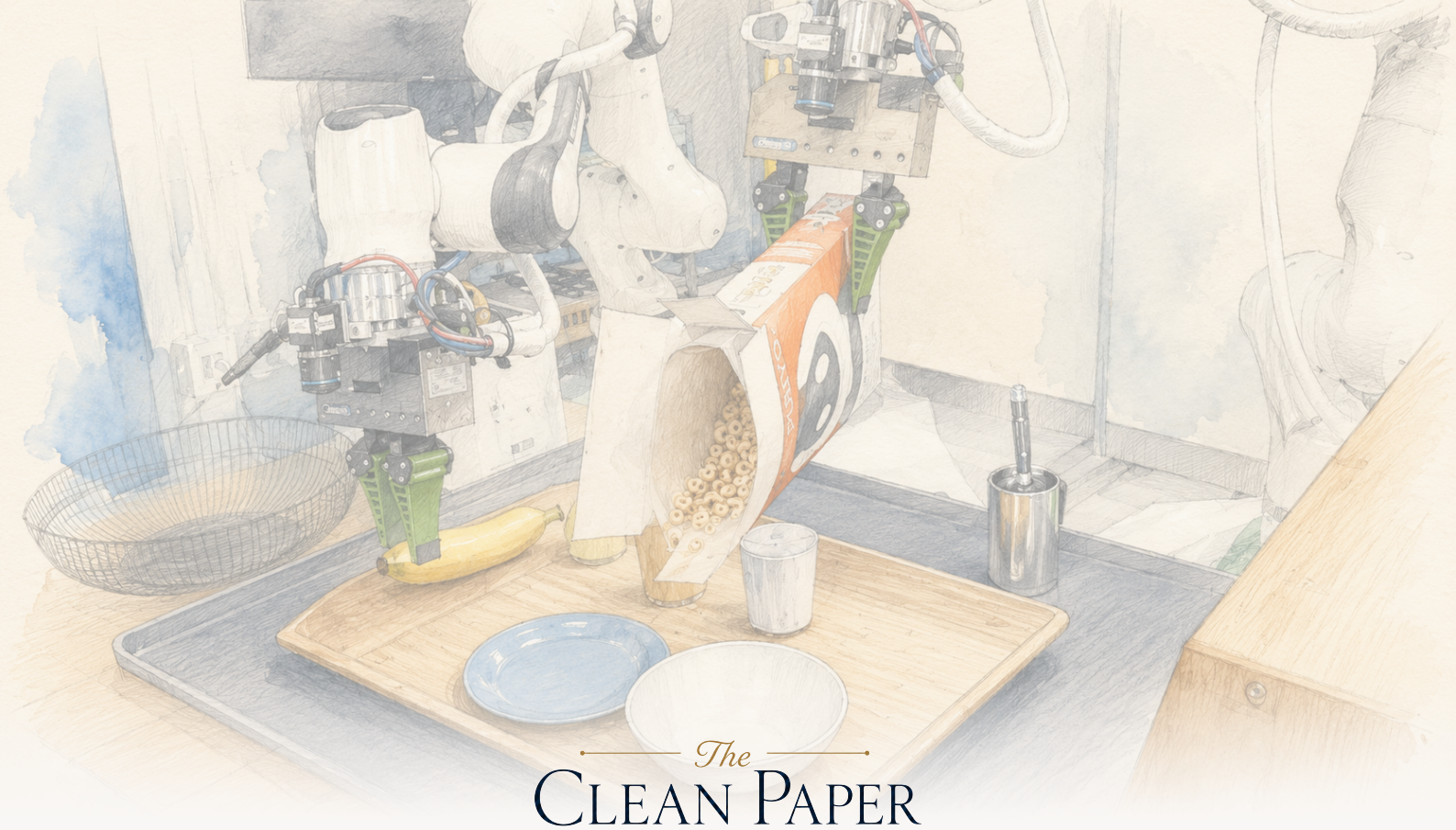


The
CLEAN PAPER

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

Ar dideli robotų „pamatiniai modeliai“ iš tiesų veikia geriau? Atsargus atsakymas — šiek tiek taip, o dauguma tyrimų to patikimai nepamatuotų

Toyota Research Institute apmokė „didelius elgesio modelius“ – robotų politikas, iš anksto mokytas maždaug 1700 valandų įvairių manipuliavimo duomenų – ir neįprastai griežtai palygino jas su nuo nulio mokytomis vienos užduoties politikomis: akli, atsitiktinės tvarkos, didelės imties bandymai (apie 1800 realių ir daugiau kaip 47 000 simuliacinių) su formalios statistikos analize. Po papildomo mokymo dideli modeliai vidutiniškai veikė geriau, reikalavo maždaug 3–5 kartus mažiau konkrečios užduoties duomenų ir buvo atsparesni pasikeitus sąlygoms. Tačiau be papildomo mokymo jie nuosekliai nelaimėjo, dalis efektų buvo labai maži, o paprastas duomenų normalizavimo pasirinkimas turėjo didesnį poveikį nei architektūra. Tai pamatuota parama krypčiai, o ne bendros paskirties zero-shot robotas ar „emergentinis šuolis“.

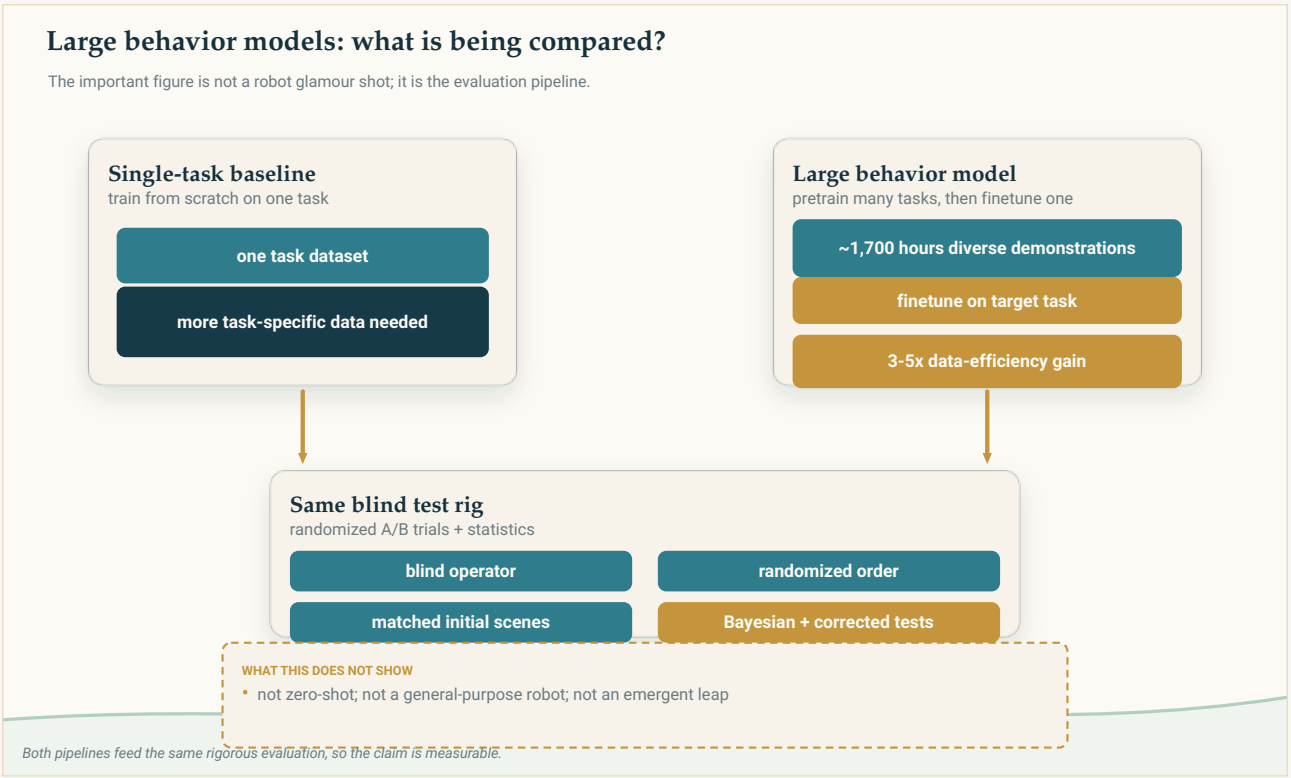
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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

Robotikos straipsnis, kurio tikroji tema – sąžiningas matavimas

Robotiką užplūdo „pamatinių modelių“ banga. Iš kalbos ir vaizdų AI pasiskolinta idėja viliojanti: užuot mokius robotą kiekvienos užduoties atskirai, galima vieną didelį **„elgesio modelį“ (LBM)** iš anksto apmokyti didžiuliu ir įvairiu demonstracijų rinkiniu, kad jis įgytų platesnių gebėjimų ir greitai prisitaikytų prie naujų užduočių. Entuziazmas ir investicijos milžiniški. Antraštė beveik parašo pati save: *bendros paskirties robotų smegenys jau čia.*

Toyota Research Institute darbas įdomus būtent todėl, kad tokios antraštės nerašo. Pagrindinis jo indėlis nėra įspūdingesnis robotas. Tai griežtas iš pažiūros paprasto klausimo patikrinimas: *ar didelio modelio metodas iš tiesų veikia geriau ir kaip apskritai tai patikimai išmatuoti?* Autorių teigimu, tokio statistinio atsargumo robotikos tyrimuose dar trūksta. Rezultatas yra naudingas „taip, bet“ ir perspėjimas, kad nemaža dalis srities gali matuoti triukšmą.



Abu metodai patenka į tą pačią aklą, atsitiktine tvarka atliekamą vertinimo sistemą. Straipsnis neteigia, kad robotų pamatiniai modeliai yra zero-shot bendros paskirties sistemos; jis rodo, kad išankstinis mokymas gali gerinti duomenų efektyvumą ir atsparumą, jei tai matuojama kruopščiai.

Original The Clean Paper diagram CC BY 4.0

Ką padarė autoriai

LBM autoriai apibrėžė konkrečiai: tai **difuzija pagrįstos regos ir judesio politikos** – Diffusion Transformer, kuris gauna kamerų vaizdus, trumpą kalbos instrukciją ir roboto sąnarių padėtis, o 10 Hz dažniu išveda trumpas motorinių komandų sekas. Modeliai buvo **iš anksto apmokyti maždaug 1700 valandų** roboto demonstracijų: daugiau kaip 500 skirtingų viduje surinktų užduočių ir viešų duomenų rinkinių. Tada modelis buvo **papildomai mokomas konkrečiai užduočiai**. Visur lyginta su vienos užduoties politika, nuo nulio apmokyta tik tos užduoties duomenimis.

Tačiau straipsnio šerdis yra *vertinimas*, kurį autoriai laiko beveik svarbiausiu rezultatu. Kad neapgautų savęs, jie naudojo:

- **Aklą atsitiktinės tvarkos A/B testavimą** realiame pasaulyje: robotą valdantis eksperimentatorius nežinojo, kuri politika bandoma, o eiliškumas buvo atsitiktinis.
- **Kontroliuojamas ir pakartojamas pradines sąlygas**: prieš kiekvieną bandymą operatoriai derino sceną su uždėtu etaloniniu vaizdu.
- **Didelį bandymų skaičių**: 50 realių paleidimų kiekvienai užduoties, politikos ir sąlygos kombinacijai; 200 kiekvienai užduočiai simuliacijoje. Iš viso – apie **1800 aklų realaus pasaulio bandymų ir daugiau kaip 47 000 simuliacijų**.
- **Tinkamą statistiką**: Bayeso sėkmės tikimybės įverčius ir porinius hipotezių testus su daugybinių palyginimų pataisomis, o ne vien persidengiančių paklaidų juostų vertinimą akimis. Ketvirtadaliui žmonių vertintų bandymų atlikta papildoma kokybės patikra, kad būtų išmatuota ir vertinimo klaida.

[video pending]

Dviejų modelių palyginimas ruošiant pusryčių stalą: kairėje – nuo nulio vienai užduočiai apmokytas modelis, dešinėje – LBM. Abu vaizdo įrašai rodomi 1× greičiu. Tai viena vertinta užduotis, ne bendros paskirties autonomijos įrodymas.

Būtent ši vertinimo sistema yra straipsnio esmė. Autoriai teigia, kad be tokios kontrolės sunku atskirti tikrą pagerėjimą nuo sėkmės ar triukšmo.

Ką jie nustatė

Po papildomo mokymo dideli modeliai vidutiniškai lenkė nuo nulio mokytus vienos užduoties modelius. Sujungus visas užduotis, iš anksto apmokytas ir konkrečiai užduočiai papildomai pritaikytas LBM patikimai lenkė nuo nulio apmokytą modelį tiek simuliacijoje, tiek realiame pasaulyje; skirtumas buvo statistiškai reikšmingas. Atskirose užduotyse LBM beveik visada buvo statistiškai ne blogesnis arba geresnis: 3/3 realių užduočių ir 15/16 simuliacinių.

Ryškiausias laimėjimas – duomenų efektyvumas. Papildomai pritaikytas LBM pasiekė nuo nulio mokyto modelio lygį naudodamas maždaug **3–5 kartus mažiau konkrečios užduoties duomenų**.

Vienoje realioje užduotyje – dengiant pusryčių stalą – LBM, papildomai mokytas tik su **15%** demonstracijų, aplenkė nuo nulio modelį, gavusį **100%** demonstracijų.

Išankstinis mokymas labiausiai padėjo pasikeitus sąlygoms. Kai bandymo aplinka sąmoningai nutolo nuo mokymo sąlygų (*distribution shift*), papildomai mokyto LBM pranašumas **padidėjo**. Viename simuliacijų rinkinyje normaliomis sąlygomis jis statistiškai aplenkė bazinį modelį 3 iš 16 užduočių, o pasikeitus sąlygoms – 10 iš 16. Kadangi realus pasaulis visada šiek tiek skiriasi nuo mokymo duomenų, šis atsparumas praktiškai gali būti svarbiausias rezultatas.

Daugiau išankstinio mokymo duomenų padėjo tolygiai. Našumas nuosekliai kilo didėjant išankstinio mokymo rinkiniui. Bandytame mastelyje **nebuvo staigaus šuolio ar „emergentinio“ lūžio**. Naudinga, prognozuojama ir ne itin dramatiška.

Tačiau bendro modelio be papildomo mokymo istorija nepasitvirtino. Iš anksto apmokytas LBM, naudojamas *zero-shot* režimu, **ne visada** lenkė vienos užduoties modelius. Vienas tinklas galėjo atlikti daug užduočių, tačiau vizija „tiesiog pasakyk robotui, ką daryti“ čia nebuvo patvirtinta. Autoriai dalį problemos sieja su gana silpnu kalbos koduotuvu.

Pagerėjimai buvo tokie maži, kad juos lengva nepastebėti – arba netyčia „atrasti“. Dalis efektų tapo matomi tik su neįprastai didelėmis imtimis ir kruopščiais testais. Autoriai tiesiai rašo, kad, atsižvelgiant į efektų dydžius ir triukšmą, **yra didelė rizika, jog daugelis robotikos straipsnių matuoja statistinį triukšmą**. Be to, paprastas duomenų **normalizavimo** pasirinkimas rezultatams turėjo didesnį poveikį nei architektūros pakeitimai, o išankstinio mokymo normalizavimo **klaida programoje** buvo rasta tik baigus vertinimus.

Ką tai greičiausiai reiškia

Atsargi išvada: didelio masto išankstinis mokymas įvairiais robotikos duomenimis iš tiesų yra naudingas ingredientas. Jis sumažina naujai užduočiai reikalingų duomenų kiekį ir padidina politikų atsparumą, kai realios sąlygos neatitinka mokymo. Tai realiai palaiko kryptį, į kurią investuoja sritis. Tačiau pranašumai yra *vidutiniai ir sąlyginiai*: daugiausia pasirodo po papildomo mokymo ir aiškiausiai matomi sujungus užduotis arba patiriant aplinkos pokytį. Tai ne iš dėžutės išimtas bendros paskirties robotas.

Tylesnė ir galbūt svarbesnė prasmė metodologinė. Straipsnis tampa savotiška matavimo liniuote: parodo, kiek įrodymų iš tikrųjų reikia patikimam teiginiui apie roboto politiką, ir leidžia įtarti, kad nemažai publikuoto entuziazmo paremta per menku duomenų kiekiu. Tokia korekcija sričiai gali būti naudingesnė už dar vieną naują modelį.

Ko tai *neįrodo*

- Tai **ne bendros paskirties robotas**. Pranašumai parodyti konkrečiai difuzinių politikų architektūrai, papildomai mokomai kiekvienai užduočiai, kontroliuojamomis sąlygomis ir iš teleoperuo-

jamų demonstracijų. Tai ne robotas, kuris autonomiškai vykdo bet kokią naują komandą.
- Tai **nepatvirtina zero-shot naudojimo**. Be papildomo mokymo didelis modelis ne visada lenkė vienos užduoties bazinius modelius.
- Tai **nėra „emergentinio šuolio“ įrodymas**. Didinant duomenų mastelį našumas kilo tolygiai, be staigaus gebėjimų lūžio.
- Skaičiai yra **santykiniai ir laboratoriniai**. Absoliutūs sėkmės rodikliai sąmoningai sureguliuoti maždaug ties 50%, kad palyginimas būtų jautresnis; tai nėra realaus pasaulio patikimumo matas. Be to, tirta viena architektūra vienoje laboratorijoje.
- Tyrimas **nepaaiškina, kodėl** konkreti politika vienoje ar kitoje užduotyje pasiseka ar nepasiseka; keli atvejai, kuriuose LBM pasirodė blogiau, pateikiami, bet nepaaiškinami.
- Jis **nieko nerodo apie saugą, autonomiją ar diegimą** už vertinimo sistemos ribų.

Kiek tvirti yra įrodymai?

Pagrindiniams palyginamiesiems teiginiams – kad papildomai mokyti LBM vidutiniškai lenkia nuo nulio bazinius modelius, reikalauja kelis kartus mažiau duomenų ir geriau atlaiko sąlygų poslinkį – įrodymai stiprūs ir neįprastai gerai kontroliuoti: akli, atsitiktinės tvarkos, didelės imties bandymai, formalūs statistiniai testai ir žmogaus vertinimų kokybės patikra. Tai retas atvejis, kai metodika pakankamai tvirta, kad pagrindines išvadas būtų galima priimti beveik pažodžiui.

Svarbius apribojimus iškelia patys autoriai. Jų paklaidų intervalai apima *vertinimo* atsitiktinumą, bet **ne mokymo atsitiktinumą** – du kartus apmokius tą patį modelį galima gauti reikšmingai skirtingas politikas, o ši variacija statistikoje neįtraukta. Realioms užduotims naudota po 50 bandymų, pakanka vidutiniams efektams, bet maži gali likti nepastebėti. Kalbos sąlygoms naudotas gana kuklus koduotuvas, tad didesnių vizijos-kalbos-veiksmo modelių rezultatai gali skirtis. Ir lieka atvirai paskelbta po vertinimo rasta normalizavimo klaida.

Dar viena šaltinio pastaba: šis aiškinimas paremtas autorių preprintu. Žurnale paskelbtos versijos nepavyko gauti, todėl galimi preprinto ir galutinio teksto pakeitimai nebuvo patikrinti.

Kodėl tai svarbu

„Robotų pamatiniai modeliai“ – frazė, beveik kviečianti perdėti. Šį tyrimą lengva perskaityti ir kaip triumfuojantį „veikia!“, ir kaip skeptišką „visa tai per daug išpūsta“. Tikslesnė išvada naudingesnė: įvairių duomenų išankstinis mokymas suteikia realių, išmatuojamų, bet vidutinio dydžio pranašumų – ypač mažina konkrečiai užduočiai reikalingų duomenų kiekį ir didina atsparumą – o mastelio poveikis auga prognozuojamai.

Gilesnė priežastis, kodėl darbas svarbus, yra tai, kad jo griežtumas nukreiptas į pačią robotikos sritį. Parodydamas, kad tikri efektai pakankamai maži, jog pradingtų prastame vertinime, ir kad nuobodus duomenų normalizavimas gali būti svarbesnis už gudrią naują architektūrą, straipsnis argumentuoja už daug stipresnę robotų mokymosi rezultatų matavimo kultūrą. Darbas, kuris savo patikimumą

naudoja skirtumui tarp rezultato ir noro išmatuoti, daro retesnį ir vertingesnį dalyką nei tiesiog užima pirmą vietą lentelėje.

Trumpa santrauka

Toyota Research Institute tyrėjai apmokė „didelius elgesio modelius“ – difuzija pagrįstas robotų politikas, iš anksto mokytas maždaug 1700 valandų įvairių manipuliavimo demonstracijų – ir palygino jas su nuo nulio mokytomis vienos užduoties politikomis taikydami neįprastai griežtą protokolą: aklus, atsitiktinės tvarkos, didelės imties bandymus (apie 1800 realių ir daugiau kaip 47 000 simuliacinių) su formalios statistikos analize. Po papildomo mokymo konkrečiai užduočiai dideli modeliai vidutiniškai pasirodė geriau, reikalavo maždaug **3–5 kartus mažiau užduoties duomenų** ir buvo **atsparesni pasikeitus sąlygoms**; daugiau išankstinio mokymo duomenų našumą gerino tolygiai. Tačiau **be papildomo mokymo** jie nuosekliai nelengė vienos užduoties modelių, keli efektai buvo tokie maži, kad juos parodė tik didelės imtys, o paprastas duomenų normalizavimo pasirinkimas turėjo didesnį poveikį už architektūrą. Tai pamatuota parama robotų pamatinių modelių krypčiai – **ne** bendros paskirties robotas, ne zero-shot generalistas ir ne „emergentinis šuolis“ – kartu su rimtu perspėjimu, kad dalis robotikos tyrimų gali matuoti triukšmą.

Be pagražinimų

Ką rodo straipsnis: Griežtai aklai ir statistiškai pakankamai galingai vertinant, daug užduočių iš anksto apmokytos ir vėliau papildomai pritaikytos difuzinės politikos vidutiniškai lenkia nuo nulio mokytas vienos užduoties politikas, panašų lygį pasiekia su maždaug 3–5 kartus mažiau užduoties duomenų ir geriau atlaiko sąlygų poslinkį; našumas tolygiai auga su išankstinio mokymo duomenų kiekiu.

Kas tikėtina, bet neįrodyta: Kad tie patys pranašumai persikels į daug didesnius regos-kalbos-veiksmo modelius; kad tolygus mastelio dėsningumas tęsis už čia bandyto duomenų intervalo.

Ko tai nerodo: Bendros paskirties ar zero-shot roboto; jokio „emergentinio“ gebėjimų šuolio; paaiškinimų konkrečioms nesėkmėms; absoliutaus realaus pasaulio patikimumo; nieko apie saugą ar autonominį diegimą.

Pagrindiniai ribotumai: Statistika apima vertinimo, bet ne skirtingų mokymo paleidimų atsitiktinumą; 50 realių bandymų vienai užduočiai gali nepastebėti mažų efektų; viena architektūra ir viena laboratorija; kuklus kalbos koduotuvas; normalizavimo klaida rasta tik po vertinimų; analizė remiasi preprintu, o galutinė žurnalo versija nepatikrinta.

Kiek pasitikėti? Daug – kad daugiafunkcis išankstinis mokymas kartu su papildomu pritaikymu suteikia realių, vidutinių pranašumų, ypač duomenų efektyvumo ir atsparumo srityse, ir kad jie išmatuoti neįprastai kruopščiai. Taip pat daug – kad tai **ne** bendros paskirties ar zero-shot robotas ir **ne** emergentinis lūžis. Vidutiniškai – kiek toli pranašumai tęsis didesniuose modeliuose. Ir verta

rimtai vertinti pačių autorių įspėjimą: kai efektai maži, per silpni tyrimai gali publikuoti triukšmą.

Šaltiniai

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/