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

Kas suured robotite „alusmudelid“ töötavad tegelikult paremini? Hoolikas vastus: mõõdukalt jah, ja enamik uuringuid ei suuda seda eristada

Toyota Research Institute treenis „suuri käitumismudeleid“ — robotipoliitikaid, mis eeltreeniti ~1700 tunni mitmekesiste manipulatsiooniandmete peal — ning võrdles neid nullist treenitud ühe ülesande poliitikatega ebatavaliselt range, pimeda, randomiseeritud ja suure valimiga hindamise abil (~1800 pärismaailma, üle 47 000 simulatsioonikatse) koos korraliku statistikaga. Pärast ülesandespetsiifilist peenhäälestust olid suured mudelid keskmiselt paremad, vajasid umbes 3–5× vähem ülesandespetsiifilisi andmeid ja olid tingimuste nihkudes robustsemad; rohkem eeltreeninguandmeid parandas jõudlust sujuvalt. Ilma peenhäälestuseta ei edestanud nad aga järjepidevalt ühe ülesande mudeleid, osa efekte oli nähtav ainult suurte valimitega ning andmete normaliseerimine mõjutas tulemust arhitektuurist rohkem. See on mõõdukas toetus robotite alusmudelite suunale — mitte üldrobot, zero-shot üldmudel ega „emergentne hüpe“ — ning hoiatus, et suur osa robootikast võib mõõta müra.

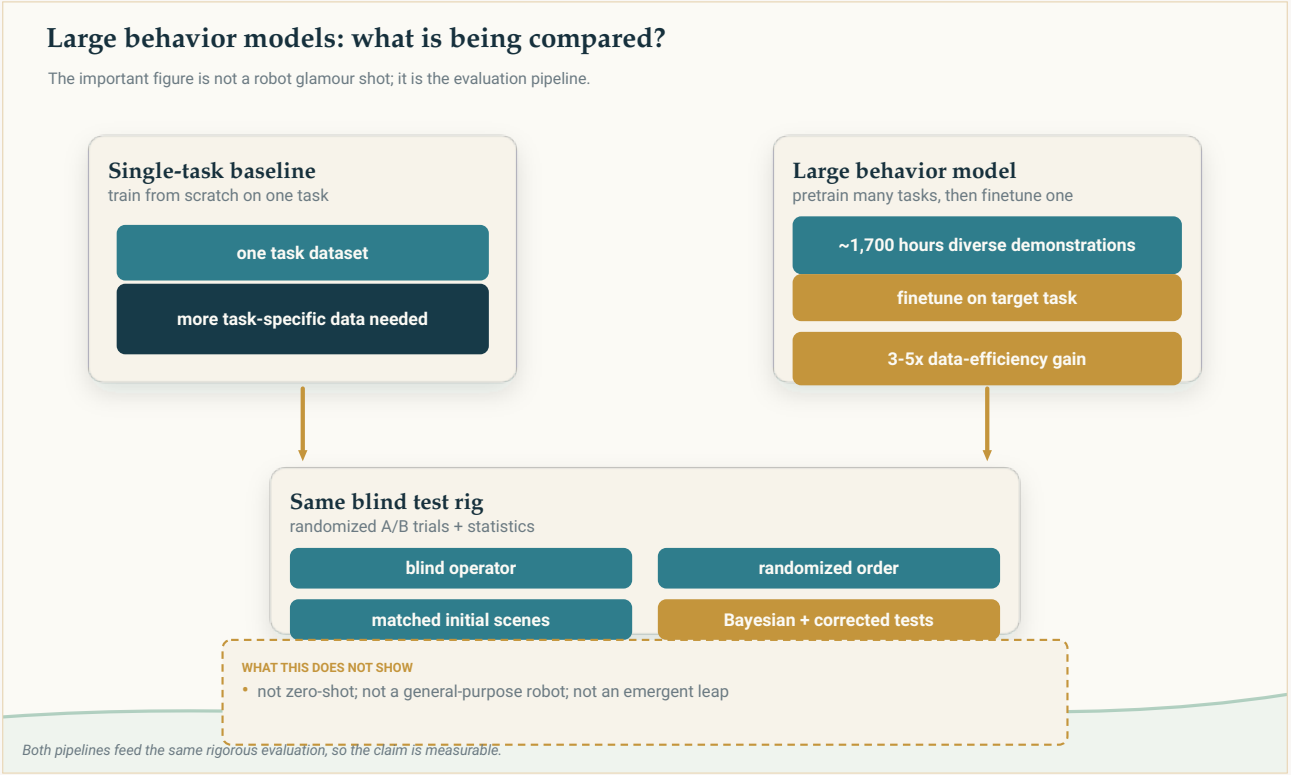
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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

Robootikaartikkel, mille tegelik teema on ausus

Robootikas käib „alusmudelite“ buum. Keele- ja pildi-AI-st laenatud idee on ahvatlev: selle asemel et õpetada robotit ühe ülesande kaupa, treeni üks suur **„suur käitumismudel“ (large behavior model, LBM)** tohutu ja mitmekesise demonstratsioonikogumi peal ning saad süsteemi, mille võimed on laiad ja mis kohaneb kiiresti. Entusiasm — ja investeeringud — on hiiglaslikud. Pealkiri kirjutab end ise: *üldotstarbelised robotiajud on kohal.*

Toyota Research Institute'i artikkel on huvitav just seetõttu, et keeldub seda pealkirja kirjutamast. Selle päris panus ei ole efektsem robot. See vaatab rangelt üht petlikult lihtsat küsimust — *kas suure mudeli lähenemine tõesti töötab paremini ja kuidas me seda üldse teaksime?* — statistilise hoolikusega, mis autorite endi sõnul on valdkonnas ebatavaline. Tulemus on päriselt kasulik „jah, aga“ ja hoiatus, et suur osa robootikast võib mõõta müra.



Mõlemad lähenemised lähevad samasse pimedasse randomiseeritud hindamisse. Artikli väide pole, et robotite alusmudelid on zero-shot üldmudelid, vaid et eeltreening võib hoolika mõõtmise korral parandada andmetõhusust ja robustsust.

Original The Clean Paper diagram CC BY 4.0

Mida autorid tegid

Nad ehitasid LBM-id väga konkreetses tähenduses: **difusioonipõhised visuomotoorsed poliitikad** (Diffusion Transformer, mis loeb kaamerapilte, lühikest keelelist juhist ja roboti enda liigeste

asendeid ning väljastab 10 Hz sagedusel lühikesi motoorsete käskude jadasid). Neid **eeltreeniti umbes 1700 tunni** robotidemonstratsioonide peal — üle 500 majasiseselt kogutud eri ülesande pluss avalikud andmestikud — ja seejärel **peenhäälestati** üksikutele ülesannetele. Kogu töö vältel võrreldakse neid **ühe ülesande poliitikaga, mis treeniti nullist** ainult selle ülesande andmetel.

Artikli süda on siiski *hindamine*, mida autorid käsitlevad peamise tulemusena. Et mitte iseennast ära petta, kasutasid nad:

- **Pimedat randomiseeritud A/B-testimist** pärismaailmas — robotit käitanud inimene ei teadnud, millist poliitikat parajasti testiti, ja järjekord oli juhuslik.
- **Kontrollitud, korratavaid algtingimusi** — enne iga katset sobitas operaator stseeni pildi ülekatte järgi.
- **Suurt katsete arvu** — 50 pärismaailma käivitust iga ülesande, poliitika ja tingimuse kohta; simulatsioonis 200 ülesande kohta. Kokku umbes **1800 pimedat pärismaailma käivitust ja üle 47 000 simulatsioonikäivituse**.
- **Korralikku statistikat** — Bayesi hinnanguid õnnestumise tõenäosusele ja paarikaupa hüpoteesiteste mitme võrdluse parandustega, mitte kattuvate veapiiride silmaga vaatamist. Nad tegid isegi kvaliteedikontrolli veerandile inimese hinnatud katsetest, et mõõta hindamisviga.

[video pending]

Kõrvuti võrdlus mudelitest, mis katavad hommikusöögilauda: vasakul ühe ülesande baasmudel ja paremal LBM. Mõlemad videod mängivad 1× kiirusel. See on üks hinnatud ülesanne, mitte tõend üldotstarbelise autonoomia kohta.

Just see hindamissüsteem ongi asja mõte. Kogu artikkel väidab, et ilma selleta ei saa tegelikku paranemist õnnest eristada.

Mida nad leidsid

Peenhäälestatud suured mudelid edestasid keskmiselt nullist treenitud ühe ülesande mudeleid. Ülesannete peale koondatuna oli esmalt eeltreenitud ja seejärel konkreetsele ülesandele peenhäälestatud LBM nii simulatsioonis kui pärismaailmas usaldusväärselt parem kui sama ülesande andmetel nullist treenitud poliitika ning erinevus oli statistiliselt oluline. Üksikute ülesannete puhul oli peenhäälestatud LBM peaaegu alati statistiliselt sama hea või parem (3/3 pärismaailma ülesannet, 15/16 simulatsiooniülesannet).

Suurim ja selgeim võit on andmetõhusus. Peenhäälestatud LBM saavutas nullist treenitud mudeliga võrdse jõudluse umbes **3–5× väiksema ülesandespetsiifilise andmehulgaga**. Ühes pärismaailma ülesandes — hommikusöögilaua katmises — edestas vaid **15%** demonstratsioonide peal peenhäälestatud LBM nullist treenitud poliitikat, mis oli näinud **100%** demonstratsioonidest.

Eeltreening aitab kõige rohkem siis, kui tingimused nihkuvad. Kui testikeskkonda muudeti sihilikult treeningtingimustest erinevaks („jaotusnihe“), peenhäälestatud LBM-i eelis *kasvas*. Ühes simulatsioonikomplektis oli see tavatingimustes statistiliselt parem 3 ülesandes 16-st, jaotusnihke korral

aga 10 ülesandes 16-st. Kuna päris kasutuskeskkond triivib alati treeningtingimustest eemale, võib just see robustsus olla praktiliselt kõige tähtsam leid.

Rohkem eeltreeninguandmeid aitas sujuvalt. Jõudlus kasvas eeltreeninguandmete lisamisel järkjärgult — testitud skaalal **ei ilmunud järsku hüpet ega „emergentset" võimepurset**. Kasulik, prognoositav, ebadramaatiline.

Kuid lugu üldmudelist ilma peenhäälestuseta ei pidanud vastu. Eeltreenitud LBM, mida kasutati *zero-shot* režiimis — ilma ülesandespetsiifilise peenhäälestuseta — **ei** edestanud järjepidevalt ühe ülesande poliitikaid. Üks võrk suutis teha mitut ülesannet, kuid siin ei täitunud unistus „ütle lihtsalt robotile, mida teha"; autorid seostavad osa probleemist oma väikese keelekodeerija haprusega.

Ja võidud olid piisavalt väikesed, et neid on lihtne mitte märgata — või ekslikult näha. Paljud efektid muutusid nähtavaks alles tavalisest suuremate valimite ja hoolikate testidega. Autorid ütlevad otse, et efektide suurust ja müra arvestades **on märkimisväärne oht, et paljud robootikaartiklid mõõdavad statistilist müra.** Samuti leidsid nad, et igav praktiline valik — kuidas andmeid *normaliseerida* — mõjutas tulemusi rohkem kui arhitektuurimuudatused, ning eeltreeningu normaliseerimisest leiti **viga** alles *pärast* hindamiste lõpetamist.

Mida see tõenäoliselt tähendab

Kaitstav lugemine: suuremahuline eeltreening mitmekesistel robotiandmetel on päris ja väärtuslik komponent — iga uue ülesande jaoks on vaja vähem andmeid ning poliitikad peavad paremini vastu siis, kui maailm ei vasta treeningule. See toetab tegelikult suunda, millele valdkond panustab. Kuid kasu on *mõõdukas ja tingimuslik* — see ilmub peamiselt pärast peenhäälestust ning on kõige selgem koondanalüüsis ja stressitingimustes — mitte universaalse, kohe kasutatava roboti saabumine.

Vaiksem ja olulisem tähendus on metoodiline. Artikkel on sisuliselt mõõdupuu: see näitab, kui palju tõendusmaterjali on tegelikult vaja usaldusväärse väite tegemiseks robotipoliitika kohta, ning vihjab, et suur osa avaldatud elevusest toetub liiga vähesele. Sellist parandust vajab valdkond rohkem kui järjekordset mudelit.

Mida see *ei* tõesta

- See **ei ole üldotstarbeline robot**. Võidud on näidatud kindla arhitektuuri — difusioonipoliitikate — puhul, mida peenhäälestati iga ülesande jaoks eraldi, kontrollitud tingimustes ja teleopereeritud demonstratsioonidest; mitte autonoomse roboti puhul, kes täidab käsu peale suvalisi uusi töid.
- See **ei kinnita zero-shot kasutust**. Ilma peenhäälestuseta ei edestanud suur mudel järjepidevalt ühe ülesande baasmudeleid.
- See **ei ole tõend „emergentse hüppe" kohta**. Skaala kasv parandas tulemusi sujuvalt; siin pole katkendlikkust, mis toetaks lugu „ja siis muutus ta äkki võimekaks".
- Arvud on **suhtelised ja laborikesksed**. Absoluutsed õnnestumismäärad häälestati sihilikult

umbes 50% lähedale, et võrdlused oleksid tundlikumad; need ei mõõda pärismaailma töökindlust ning tegu on ühe labori ühe arhitektuuriga.
- See **ei selgita**, miks konkreetne poliitika õnnestub või ebaõnnestub; mitme ülesande puhul, kus suur mudel oli *halvem*, tulemus esitatakse, kuid põhjust ei seletata.
- See **ei ütle midagi ohutuse, autonoomia ega kasutuselevõtu kohta** väljaspool hindamisraamistikku.

Kui tugev on tõendusmaterjal?

Kesksete võrdlusväidete puhul — peenhäälestatud LBM-id edestavad koondtulemuses nullist treenitud baasmudeleid, vajavad mitu korda vähem andmeid ja on jaotusnihke korral robustsemad — on tõendus tugev *ja* ebatavaliselt hästi kontrollitud: pime, randomiseeritud, suurte valimitega, statistiliselt testitud ja hindamise QA-kontrolliga. See on haruldane juhtum, kus metoodika on piisavalt tugev, et võtta põhitulemusi peaaegu nimiväärtuses.

Ausad piirangud toovad autorid ise välja. Veapiirid kajastavad *hindamise* juhuslikkust, kuid **mitte treeningu** juhuslikkust — kui treenida sama mudelit kaks korda, võib saada märgatavalt erineva poliitika, ja see varieeruvus statistikas ei kajastu. Pärismaailma ülesannetel oli 50 katset, piisavalt keskmiste efektide tabamiseks, kuid väikseid mõjusid võib see mööda lasta. Keeleline tingimine kasutas tagasihoidlikku kodeerijat, seega võivad väited stiilis „ütle robotile lihtsalt, mida teha“ suuremate süsteemide puhul olla teistsugused. Ja autorid avalikustavad otse normaliseerimisvea, mis leiti tagantjärele. Need ei uputa põhitulemusi, kuid on täpselt sellised asjad, mida artikli järgi valdkond sageli kõrvale pühib.

Samast vaimust lähtuv allikamärkus: see seletus põhineb autorite preprindil. Meil ei õnnestunud saada kätte ajakirjas avaldatud versiooni, seega pole kontrollitud, kas preprindi ja avaldatud teksti vahel tehti muudatusi.

Miks see oluline on

„Robotite alusmudelid“ on väljend, mis lausa kutsub liialdama, ja sellist uuringut on lihtne valesti lugeda mõlemas suunas — võiduka „see töötab!“ või tõrjuva „see on üle haibitud“ kujul. Täpsem järeldus on mõlemast kasulikum: mitmekesistel andmetel eeltreenimine annab päris, mõõdetava, kuid mõõduka kasu — peamiselt vähem andmeid ülesande kohta ja suurema robustsuse — ning paranemine skaleerub prognoositavalt.

Sügavam põhjus, miks artikkel loeb, on see, et ta pöörab ranguse omaenda valdkonna vastu. Näidates, et päris efektid on piisavalt väikesed, et lohaka hindamise korral kaduda, ning et igav valik nagu andmete normaliseerimine võib kaaluda üle nutika uue arhitektuuri, väidab artikkel, et suur osa robotõppe edust vajab enne uskumist tugevamat mõõtmist. Artikkel, mis kulutab oma usaldusväärsuse tulemuse ja soovi vahelise piiri valvamisele, teeb midagi haruldasemat ja väärtuslikumat kui edetabeli võitmine.

Lühidalt

Toyota Research Institute'i teadlased treenisid „suuri käitumismudeleid" — difusioonipõhiseid robotipoliitikaid, mis eeltreeniti umbes 1700 tunni mitmekesiste manipulatsiooniandmete peal — ning võrdlesid neid nullist treenitud ühe ülesande poliitikatega ebatavaliselt range protokolli alusel: pime, randomiseeritud ja suure valimiga hindamine (≈1800 pärismaailma ja üle 47 000 simulatsioonikatse) koos päris statistikaga. Pärast ülesandespetsiifilist peenhäälestust olid suured mudelid koondtulemuses usaldusväärselt paremad, vajasid umbes **3–5× vähem ülesandespetsiifilisi andmeid** ja olid **tingimuste muutudes robustsemad**, samal ajal kui jõudlus paranes eeltreeninguandmete kasvades sujuvalt. Kuid **ilma peenhäälestuseta** ei edestanud nad järjepidevalt ühe ülesande mudeleid, mitmed efektid olid nii väikesed, et neid näitasid alles suured valimid, ning tavapärane andmete normaliseerimise valik mõjutas tulemusi rohkem kui arhitektuur. See on tugev, mõõdukas toetus robotite alusmudelite suunale — **mitte** üldotstarbeline robot, zero-shot üldmudel ega „emergentne hüpe" — ning terav hoiatus, et suur osa robootikast võib mõõta müra.

Kaine hinnang

Mida artikkel näitab: Range, pimeda ja statistilise võimsusega hindamisega (≈1800 pärismaailma + üle 47 000 simulatsioonikäivituse) edestavad mitme ülesande peal eeltreenitud ja seejärel peenhäälestatud difusioonipoliitikad (LBM-id) koondtulemuses nullist treenitud ühe ülesande poliitikaid, saavutavad samaväärse jõudluse ~3–5× väiksema ülesandespetsiifilise andmehulgaga ning on jaotusnihke korral robustsemad; jõudlus kasvab eeltreeninguandmetega sujuvalt.

Mis on usutav, kuid tõestamata: Et need eelised kanduvad palju suurematesse nägemis-keel-tegevus mudelitesse (nende keelekodeerija oli väike); et sujuv skaleerumine jätkub väljaspool testitud andmevahemikku.

Mida see ei näita: Üldotstarbelist ega zero-shot robotit (ilma peenhäälestuseta puudus järjepidev eelis); mingit „emergentset" võimehüpet; seletusi konkreetsetele ülesandepõhistele ebaõnnestumistele; absoluutset pärismaailma töökindlust (tundlikkuse suurendamiseks hoiti õnnestumismäärad umbes 50% juures); midagi ohutuse või autonoomse kasutuselevõtu kohta.

Peamised piirangud: Statistika mõõdab hindamise, mitte eri treeningkäikude juhuslikkust; 50 pärismaailma katset ülesande kohta võib väikseid efekte mitte tabada; üks arhitektuur ja üks labor; tagasihoidlik keelekodeerija; andmete normaliseerimise viga leiti pärast hindamist; analüüs põhineb preprindil ja avaldatud versiooni pole kontrollitud.

Kui kindel võiks tavalugeja olla? Väga kindel, et mitme ülesande eeltreening koos peenhäälestusega annab päris, mõõdukaid eeliseid — eriti andmetõhususes ja robustsuses — ning et neid mõõdeti ebatavaliselt hoolikalt. Väga kindel, et see **ei ole** üldotstarbeline ega zero-shot robot ja **ei ole** emergentne hüpe. Keskmine kindlus selles, kui kaugele kasu suuremate mudelitega skaleerub. Ning tõsiselt tasub võtta autorite enda hoiatust: efektid on nii väikesed, et väikese statistilise võimsusega uuringud võivad avaldada müra. Sobiv hoiak: mõõdukas optimism lähenemise suhtes ja terve skepsis robot-AI tulemuste suhtes, millel selline statistiline tugi puudub.

Allikad

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/