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

Rade li veliki robotski „osnovni modeli" zaista bolje? Pažljiv odgovor: skromno da, a većina studija to ne može pouzdano razlikovati

Toyota Research Institute trenirao je „velike modele ponašanja" — robotske politike predtrenirane na ~1,700 sati raznolikih podataka o manipulaciji — i uporedio ih s politikama za jedan zadatak treniranima od nule uz neuobičajenu rigoroznost: slepi, randomizovani eksperimenti velikog uzorka (~1,800 u stvarnom svetu, 47,000+ u simulaciji) i prava statistika. Nakon dodatnog treniranja za pojedini zadatak veliki modeli bili su u proseku bolji, trebali su približno 3–5× manje podataka specifičnih za zadatak i bili robusniji kada su se uslovi menjali; izvedba je postupno rasla s više podataka za predtreniranje. Ali bez dodatnog treniranja nisu dosledno nadmašivali modele za jedan zadatak, neki su efekti bili toliko mali da su zahtevali velike uzorke, a obična odluka o normalizaciji podataka bila je važnija od arhitekture. To je odmerena podrška smeru robotskih osnovnih modela — ne opštenamenski robot, ne zero-shot generalist, ne „emergentni skok" — i oštro upozorenje da velik deo robotike možda meri šum.

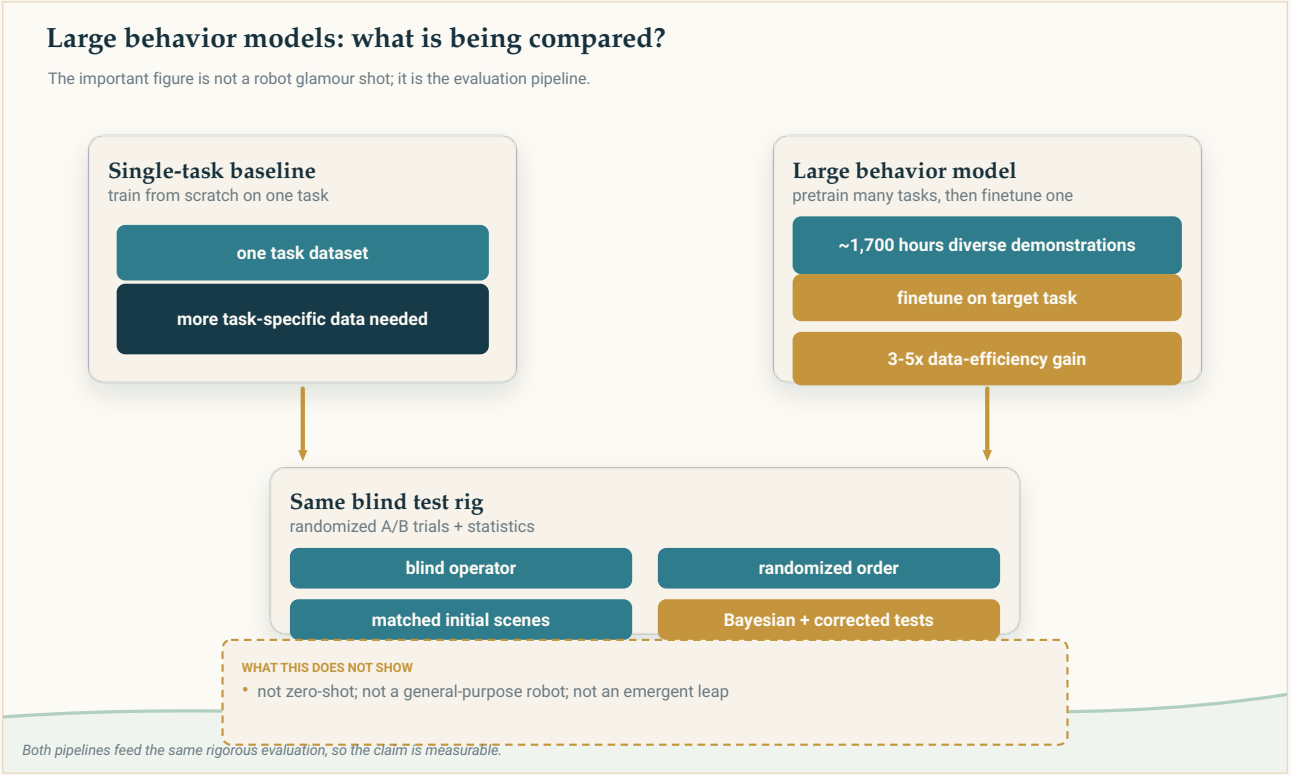
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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

Rad o robotima čija je prava tema poštenje

Robotika je usred navale "osnovnih modela". Ideja, preuzeta iz jezičkih i vizuelnih modela veštačke inteligencije, zavodljiva je: umesto da robota obučavate za jedan zadatak odjednom, istrenirajte jedan veliki **"model ponašanja" (large behavior model, LBM)** na ogromnoj i raznolikoj zbirci demonstracija pa dobijte sistem širokih sposobnosti koji se brzo prilagođava. Oduševljenje — i ulaganja — ogromni su. Naslov se piše sam: *stigao je opštenamenski robotski mozak.*

Ovaj rad Toyota Research Institutea zanimljiv je upravo zato što odbija napisati taj naslov. Njegov stvarni doprinos nije atraktivniji robot. To je tvrd pogled na naizgled jednostavno pitanje — *radi li pristup velikog modela zaista bolje i kako bismo to uopšte znali?* — uz nivo statističke pažnje koji je, prema samim autorima, neuobičajen za tu područje. Rezultat je zaista korisno "da, ali" i upozorenje da velik deo robotike možda meri šum.



Oba pristupa ulaze u isto naslepo, randomizovano vrednovanje. Tvrdnja rada nije da su robotski osnovni modeli zero-shot generalisti, nego da predtreniranje, kada se pažljivo meri, može poboljšati efikasnost podataka i robusnost.

Original The Clean Paper diagram CC BY 4.0

Šta su autori uradili

Izgradili su LBM-ove u konkretnom smislu: **vizuomotorne upravljačke politike zasnovane na difuziji** (Diffusion Transformer koji čita slike kamera, kratku jezičnu uputu i položaje zglobova samog rob-

ota te pri 10 Hz daje kratke nizove motoričkih naredbi). Ti su modeli **predtrenirani na približno 1,700 sati** demonstracija robota — više od 500 različitih zadataka prikupljenih unutar instituta, uz javne skupove podataka — a zatim su **dodatno trenirani (finetuned)** za pojedinačne zadatke. Poređenje je kroz celi rad s **politikom za jedan zadatak treniranom od nule** samo na podacima tog zadatka.

Srce rada ipak je *vrednovanje*, koje autori tretiraju kao glavni rezultat. Kako se ne bi zavarali, koristili su:

- **Naslepo, randomizovano A/B testiranje** u stvarnom svetu — osoba koja je upravljala eksperimentom nije znala koja se politika testira, a redosled je bio nasumičan.
- **kontrolisane, ponovljive početne uslove** — operateri su pre svakog eksperimenata uskladili scenu s preklopnom referentnom slikom.
- **Velik broj eksperimenata** — 50 izvedbi u stvarnom svetu po zadatku, politici i uslovu; 200 po zadatku u simulaciji. Ukupno: oko **1,800 slepih izvedbi u stvarnom svetu i više od 47,000 simulacijskih izvedbi**.
- **Odgovarajuću statistiku** — Bayesove procene verovatnoće uspeha i parne testove hipoteza s korekcijama za višestruke poređenja, umesto procene preklapaju li se trake greške "od oka". Čak su proveli i kontrolu kvalitete na četvrtini eksperimenata koje su ljudi bodovali kako bi izmerili grešku ocenjivanja.

[video pending]

Poređenje modela jedan uz drugi dok postavljaju stol za doručak: levo je osnovni model za jedan zadatak, desno LBM. Oba se videa reproduciraju brzinom 1x. Ovo je jedan vrednovani zadatak, ne dokaz opštenamenske autonomije.

Upravo je taj postupak poenta. Celi je rad argument da bez njega ne možete razlikovati stvarno poboljšanje od sreće.

Šta su pronašli

Veliki modeli nakon dodatnog treniranja u proseku nadmašuju modele za jedan zadatak trenirane od nule. Kada se rezultati objedine preko zadataka, LBM koji je najpre predtreniran, a zatim dodatno treniran za pojedini zadatak pouzdano je nadmašio politiku treniranu od nule na istom zadatku, i u simulaciji i u stvarnom svetu, uz statistički značajnu razliku. Na pojedinačnim zadacima dodatno trenirani LBM bio je statistički jednako dobar ili bolji od modela treniranog od nule u gotovo svakom slučaju (3/3 zadatka u stvarnom svetu, 15/16 simulacijskih zadataka).

Najveća i najjasnija prednost je efikasnost podataka. Dodatno trenirani LBM dostigao je izvedbu uporedivu s modelom treniranim od nule koristeći približno **3–5× manje podataka specifičnih za zadatak**. U jednom zadatku u stvarnom svetu (postavljanje stola za doručak), LBM dodatno treniran na samo **15%** demonstracija nadmašio je politiku treniranu od nule na **100%** demonstracija.

Predtreniranje najviše pomaže kada se uslovi promene. Kada je testno okruženje namerno pomaknuto od uslova treniranja ("distribution shift"), prednost dodatno treniranog LBM-a *povećala se*.

U jednom simulacijskom skupu statistički je nadmašio model treniran od nule na 3 od 16 zadataka u normalnim uslovima, ali na 10 od 16 pri promenjenoj distribuciji. Budući da se stvarna primena uvek donekle razlikuje od uslova treniranja, ta je robusnost verovatno praktično najvažniji nalaz.

Više podataka za predtreniranje pomagalo je postupno. Izvedba je stalno rasla kako su dodavali podatke za predtreniranje — bez **iznenadnog skoka ili "emergentnog" prelaza** na testiranim razmerima. Korisno, predvidljivo i bez drame.

Ali priča o generalistu bez dodatnog treniranja nije izdržala. Predtrenirani LBM korišćen *zero-shot* — bez dodatnog treniranja specifičnog za zadatak — *nije* dosledno nadmašivao politike za jedan zadatak. Jedna mreža mogla je izvoditi mnogo zadataka, ali san "samo mu reci što želiš" ovde nije potvrđen; autori deo problema pripisuju krhkosti svog malog jezičkog enkodera.

A dobici su bili dovoljno mali da ih je lako propustiti — ili prividno proizvesti. Mnogi efekti postali su vidljivi tek uz veće uzorke nego što je uobičajeno i uz pažljive testove. Autori direktno navode da, s obzirom na veličinu efekata i šum, **postoji znatan rizik da mnogi radovi iz robotike mere statistički šum.** Otkrili su i da je jedna sasvim obična odluka — način na koji se podaci *normaliziraju* — uticala na rezultate više nego promene arhitekture, te da je **greška** u normalizaciji pri predtreningu otkrivena tek *nakon* završetka vrednovanja.

Šta to verovatno znači

Odbranjivo tumačenje: predtreniranje velikih razmere na raznolikim podacima robota stvaran je i koristan sastojak — smanjuje količinu podataka potrebnih za novi zadatak i čini politike otpornijima kada se stvarni svet ne poklapa s uslovima treniranja. To zaista podržava smer na koji se područje kladi. Ali dobici su *umereni i uslovni* (uglavnom se pojavljuju nakon dodatnog treniranja, a najjasniji su u agregatu i pod stresom), a ne dolazak gotovog opšteg robota.

Tiše, važnije značenje je metodološko. Rad je zapravo merno ravnalo: pokazuje koliko je dokaza zaista potrebno da bi se iznela pouzdana tvrdnja o robotskoj politici i sugeriše da se mnogo objavljenog uzbuđenja oslanja na premalo podataka. To je korektiv koji je ovoj područja potreban više nego još jedan model.

Šta ovo *ne* dokazuje

- Ovo **nije opštenamenski robot.** Prednosti su pokazane za određenu arhitekturu (difuzijske politike), dodatno treniranu za svaki zadatak, u kontrolisanim uslovima i iz teleoperiranih demonstracija — ne za autonomnog robota koji na zahtev obavlja proizvoljne nove poslove.
- **Ne** potvrđuje **zero-shot** upotrebu. Bez dodatnog treniranja veliki model nije dosledno nadmašivao osnovne modele za jedan zadatak.
- Ovo **nije** dokaz **"emergentnog skoka"**. Skaliranje je donosilo postupna poboljšanja; ovde nema diskontinuiteta koji bi podupirao priču "a onda je odjednom postao sposoban".

- Brojevi su **relativni i laboratorijski**. Apsolutne stope uspeha namerno su podešene prema ~50% kako bi poređenja bile osetljive; one nisu mera pouzdanosti u stvarnom svetu, a rad proučava jednu arhitekturu iz jednog laboratorija.
- **Ne** razrešava **zašto** određena politika uspeva ili ne uspeva, a više konkretnih zadataka na kojima je veliki model bio *lošiji* navedeno je bez objašnjenja.
- Ne govori **ništa o bezbednosti, autonomiji ili primeni** izvan sistema za vrednovanje.

Koliko su dokazi jaki?

Za centralno uporedne tvrdnje — da dodatno trenirani LBM-ovi u agregatu nadmašuju modele trenirane od nule, trebaju nekoliko puta manje podataka i robusniji su pri promeni distribucije — dokazi su snažni *i* neuobičajeno dobro kontrolisani: naslepo i randomizovano testiranje, veliki uzorci, statistički testovi te kontrola kvalitete bodovanja. Ovo je red slučaj u kojem je metodologija dovoljno čvrsta da se glavne zaključke može uzeti gotovo po nominalnoj vrednosti.

Poštene ograde navode sami autori. Njihove trake greške obuhvataju slučajnost *vrednovanja*, ali **ne** slučajnost *treniranja* — istrenirate li isti model dvaput, možete dobiti primetno drukčiju politiku, a ta varijacija nije obuhvaćena statistikom. Zadaci u stvarnom svetu imali su po 50 eksperimenata, dovoljno za srednje efekte, ali moguće premalo za male. Jezično uslovovanje koristilo je skroman enkoder, pa se tvrdnje o "samo reci robotu što da radi" mogu drugačije ponašati u većim sistemima. Tu je i otvoreno priznanje da je greška normalizacije pronađena naknadno. Ništa od toga ne ruši glavne nalaze, ali upravo su to vrste stvari za koje rad tvrdi da ih područje prečesto gura pod tepih.

Još jedna napomena o izvoru, u istom duhu: ovo objašnjenje zasniva se na preprintu autora. Nismo uspeli pribaviti objavljenu verziju iz časopisa, pa nismo proverili da li postoje promene između preprinta i objavljenog teksta.

Zašto je važno

"Robotski osnovni modeli" izraz je kao stvoren za preterivanje, a ovakav rad lako se pogrešno pročita u oba smera — kao trijumfalno *"radi!"* ili podcjenjivačko *"sve je prenapuhano"*. Tačno tumačenje korisnije je od oba: predtreniranje na raznolikim podacima daje stvarne, merljive, ali umerene koristi — pre svega manje podataka po zadatku i veću robusnost — a poboljšanja s razmerom rastu predvidljivo.

Dublji razlog važnosti rada je to što njegovu rigoroznost usmerava na sopstveno područje. Pokazujući da su stvarni efekti dovoljno mali da nestanu u lošem vrednovanju i da dosadna odluka poput normalizacije podataka može nadjačati pametnu novu arhitekturu, iznosi argument da velik deo napretka u učenju robota treba čvršća merenja pre nego što mu poverujemo. Rad koji sopstveni kredibilitet troši na čuvanje granice između rezultata i želje radi nešto ređe i vrednije od osvajanja prvog mesta na skali.

Kratak rezime

Istraživači Toyota Research Institutea trenirali su “velike modele ponašanja” — difuzijske robotske politike predtrenirane na ~1,700 sati raznolikih podataka o manipulaciji — i uporedili ih s politikama za jedan zadatak treniranima od nule koristeći neuobičajeno rigorozan protokol: slep, randomizovan i velikog uzorka (≈1,800 eksperimenata u stvarnom svetu i 47,000+ simulacijskih), uz pravu statistiku. Nakon dodatnog treniranja za svaki zadatak veliki modeli pouzdano su bili bolji u agregatu, trebali su približno **3–5× manje podataka specifičnih za zadatak** i bili **robusniji kada su se uslovi promenili**, a izvedba je postupno rasla s količinom podataka za predtreniranje. Ali **bez dodatnog treniranja** nisu dosledno nadmašivali modele za jedan zadatak, nekoliko efekata bilo je toliko malo da su ih otkrili tek veliki uzorci, a obična odluka o normalizaciji podataka bila je važnija od arhitekture. To je čvrsta, odmerena podrška smeru robotskih osnovnih modela — **ne** opštenamenski robot, ne zero-shot generalist i ne “emergentni skok” — uz oštro upozorenje da velik deo robotike možda meri šum.

Provera bez ulepšavanja

Šta rad pokazuje: Uz rigorozno, naslepo i statistički dovoljno snažno vrednovanje (≈1,800 izvedbi u stvarnom svetu + 47,000+ simulacijskih), difuzijske politike predtrenirane na više zadataka i zatim dodatno trenirane (LBM-ovi) u agregatu nadmašuju politike za jedan zadatak trenirane od nule, postižu uporedivu izvedbu s ~3–5× manje podataka specifičnih za zadatak i robusnije su pri promeni distribucije; izvedba postupno raste s količinom podataka za predtreniranje.

Šta je moguće, ali nije dokazano: Da se te koristi prenose na mnogo veće modele vid–jezik–akcija (njihov je jezički enkoder bio mali); da se glatko skaliranje nastavlja izvan testiranog raspona podataka.

Šta ne pokazuje: Opštenamenskog ili zero-shot robota (bez dodatnog treniranja → nema dosledne prednosti); ikakav “emergentni” skok sposobnosti; objašnjenja neuspeha na konkretnim zadacima; apsolutnu pouzdanost u stvarnom svetu (stope uspeha podešene su blizu 50% radi osetljivosti); išta o bezbednosti ili autonomnoj primeni.

Glavna ograničenja: Statistika obuhvata slučajnost vrednovanja, ali ne varijabilnost između različitih treniranja; 50 stvarnih eksperimenata po zadatku može propustiti male efekte; jedna arhitektura i jedan laboratorij; skroman jezički enkoder; greška normalizacije podataka otkrivena je nakon vrednovanja; analiza se zasniva na preprintu (objavljena verzija nije proverena).

Koliko poverenja treba imati opšti čitalac? Visoko da predtreniranje na više zadataka uz dodatno treniranje daje stvarne, umerene koristi — naročito u efikasnosti podataka i robusnosti — i da su one izmerene neuobičajeno pažljivo. Visoko da ovo **nije** opštenamenski ili zero-shot robot i **nije** emergentni skok. Srednje koliko će se dobici nastaviti pri većim modelima. I vredi ozbiljno shvatiti upozorenje samih autora da su efekti u područja dovoljno mali da studije s premalom statističkom snagom mogu prijavljivati šum. Primeren stav: odmeren optimizam prema pristupu i zdrava skepsa prema rezultatima robotske AI bez ovakve statističke podloge.

Izvori

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/