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

Da li veliki robotski „temeljni modeli" zaista rade bolje? Pažljiv odgovor: skromno da, a većina studija to ne može pouzdano utvrditi

Toyota Research Institute trenirao je „velike modele ponašanja" — robotske politike prethodno trenirane na ~1,700 sati raznovrsnih podataka o manipulaciji — i uporedio ih s politikama za jedan zadatak treniranim od nule uz neuobičajenu strogost: slijepi, randomizirani eksperimenti s velikim uzorcima (~1,800 u stvarnom svijetu, 47,000+ u simulaciji) i prava statistika. Nakon dodatnog treniranja za pojedinačne zadatke veliki modeli bili su u prosjeku bolji, trebali su približno 3–5× manje podataka specifičnih za zadatak i bili su otporniji kada su se uvjeti mijenjali; učinak je postepeno rastao s više podataka za prethodno treniranje. Ali bez dodatnog treniranja nisu dosljedno nadmašivali modele za jedan zadatak, nekoliko efekata bilo je toliko malo da su bili vidljivi tek na velikim uzorcima, a običan izbor načina normalizacije podataka bio je važniji od arhitekture. To je odmjerena podrška pravcu robotskih temeljnih modela — ne robot opće namjene, ne zero-shot generalista, ne „emergentni skok" — i jasno upozorenje da veliki dio robotike možda mjeri šum.

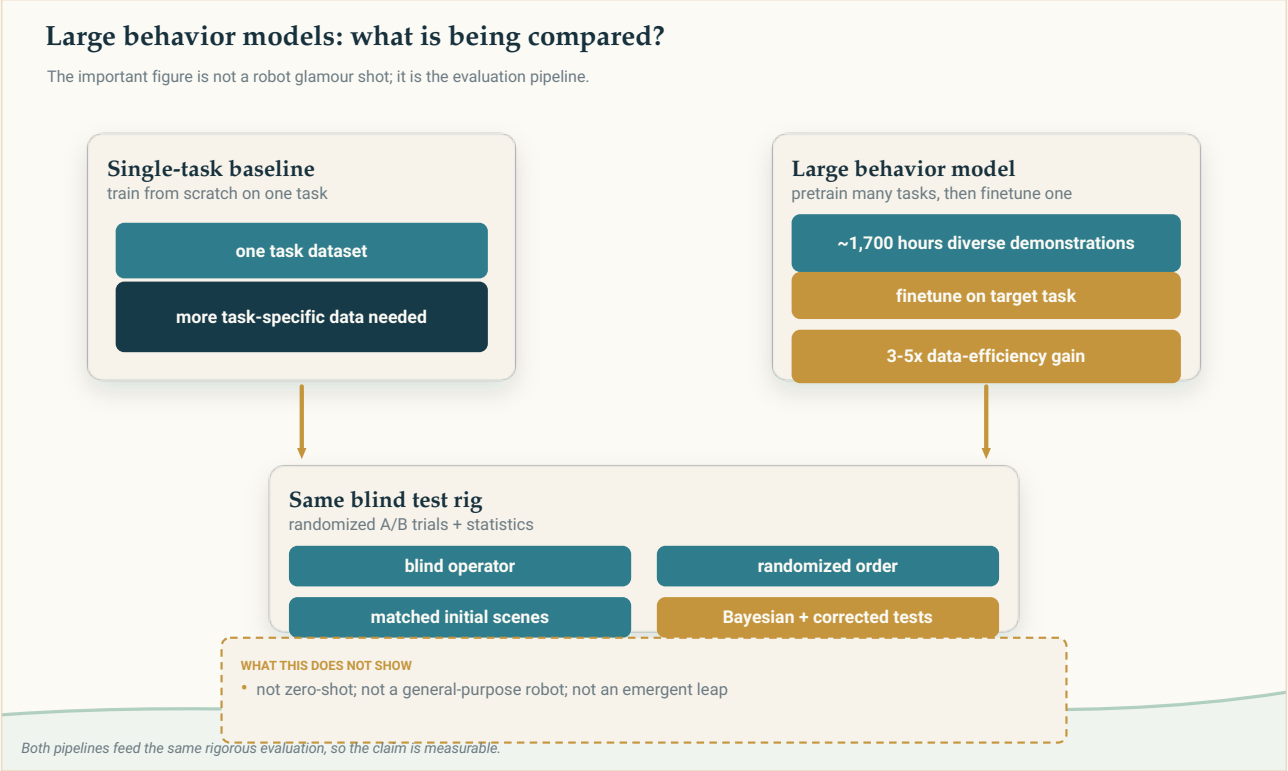
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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 "temeljnih modela". Ideja, posuđena iz jezične i slikovne vještačke inteligencije, zavodljiva je: umjesto da robota uvježbavate 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 općenamjenski 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 uopće znali?* — uz razinu statističke pažnje koja je, prema samim autorima, neuobičajena za to područje. Rezultat je zaista korisno "da, ali" i upozorenje da velik dio robotike možda mjeri šum.



Oba pristupa ulaze u isto slijepo, randomizirano vrednovanje. Tvrdnja rada nije da su robotski temeljni modeli zero-shot generalisti, nego da predtreniranje, kada se pažljivo mjeri, 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 temeljene na difuziji** (Diffusion Transformer koji čita slike kamera, kratku jezičnu uputu i položaje zglobova samog robota

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 cijeli 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:

- **Slijepo, randomizirano A/B testiranje** u stvarnom svijetu — osoba koja je upravljala eksperimentom nije znala koja se politika testira, a redoslijed je bio nasumičan.
- **Kontrolirane, ponovljive početne uslove** — operateri su prije svakog eksperimenta uskladili scenu s preklopnom referentnom slikom.
- **Velik broj eksperimenta** — 50 izvedbi u stvarnom svijetu po zadatku, politici i uslovu; 200 po zadatku u simulaciji. Ukupno: oko **1,800 slijepih izvedbi u stvarnom svijetu i više od 47,000 simulacijskih izvedbi**.
- **Odgovarajuću statistiku** — Bayesove procjene vjerovatnoće uspjeha i parne testove hipoteza s korekcijama za višestruke uporedbe, umjesto procjene preklapaju li se trake greške "od oka". Čak su proveli i kontrolu kvalitete na četvrtini eksperimenta koje su ljudi bodovali kako bi izmjerili grešku ocjenjivanja.

[video pending]

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

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

Šta su pronašli

Veliki modeli nakon dodatnog treniranja u prosjeku nadmašuju modele za jedan zadatak trenirane od nule. Kada se rezultati objedine preko zadataka, LBM koji je najprije predtreniran, a zatim dodatno treniran za pojedini zadatak pouzdano je nadmašio politiku treniranu od nule na istom zadatku, i u simulaciji i u stvarnom svijetu, 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 svijetu, 15/16 simulacijskih zadataka).

Najveća i najjasnija prednost je efikasnost podataka. Dodatno trenirani LBM dosegnuo je izvedbu uporedivu s modelom treniranim od nule koristeći približno **3–5× manje podataka specifičnih za zadatak**. U jednom zadatku u stvarnom svijetu (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 promijene. Kada je testno okruženje namjerno 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 promijenjenoj distribuciji. Budući da se stvarna primjena uvijek donekle razlikuje od uslova treniranja, ta je robusnost vjerovatno 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" prijelaza** na testiranim razmjerima. Korisno, predvidljivo i bez drame.

Ali priča o generalistu bez dodatnog treniranja nije izdržala. Predtrenirani LBM korišten *zero-shot* — bez dodatnog treniranja specifičnog za zadatak — *nije* dosljedno nadmašivao politike za jedan zadatak. Jedna mreža mogla je izvoditi mnogo zadataka, ali san "samo mu reci što želiš" ovdje nije potvrđen; autori dio problema pripisuju krhkosti svojeg malog jezičnog enkodera.

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

Šta to vjerovatno znači

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

Tiše, važnije značenje je metodološko. Rad je zapravo mjerno ravnalo: pokazuje koliko je dokaza uistinu potrebno da bi se iznijela pouzdana tvrdnja o robotskoj politici i sugerira da se mnogo objavljenog uzbuđenja oslanja na premalo podataka. To je korektiv koji području treba više nego još jedan model.

Šta ovo *ne* dokazuje

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

- Brojevi su **relativni i laboratorijski**. Apsolutne stope uspjeha namjerno su podešene prema ~50% kako bi uporedbe bile osjetljive; one nisu mjera pouzdanosti u stvarnom svijetu, a rad proučava jednu arhitekturu iz jednog laboratorija.
- **Ne** razrješava **zašto** određena politika uspijeva ili ne uspijeva, a više konkretnih zadataka na kojima je veliki model bio *lošiji* navedeno je bez objašnjenja.
- Ne govori **ništa o sigurnosti, autonomiji ili primjeni** 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 promjeni distribucije — dokazi su snažni *i* neuobičajeno dobro kontrolirani: slijepo i randomizirano testiranje, veliki uzorci, statistički testovi te kontrola kvalitete bodovanja. Ovo je rijedak slučaj u kojem je metodologija dovoljno čvrsta da se glavne zaključke može uzeti gotovo po nominalnoj vrijednosti.

Ograničenja otvoreno navode sami autori. Njihove trake greške obuhvataju slučajnost *vrednovanja*, ali **ne** slučajnost *treniranja* — istrenirate li isti model dvaput, možete dobiti primjetno drukčiju politiku, a ta varijacija nije obuhvaćena statistikom. Zadaci u stvarnom svijetu imali su po 50 eksperimenta, dovoljno za srednje učinke, ali moguće premalo za male. Jezično uslovljavanje koristilo je skroman enkoder, pa se tvrdnje o "samo reci robotu što da radi" mogu drukčije 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 temelji se na preprintu autora. Nismo uspjeli dohvatiti objavljenu verziju iz časopisa, pa nismo provjerili postoje li promjene između preprinta i objavljenog teksta.

Zašto je važno

"Robotski temeljni modeli" izraz je kao stvoren za pretjerivanje, a ovakav rad lako se pogrešno pročita u oba smjera — kao trijumfalno *"radi!"* ili podcjenjivačko *"sve je prenapuhano"*. Tačno tumačenje korisnije je od oba: predtreniranje na raznolikim podacima daje stvarne, mjerljive, ali umjerene koristi — prvenstveno manje podataka po zadatku i veću robusnost — a poboljšanja s razmjerom rastu predvidljivo.

Dublji razlog važnosti rada jest to što njegovu rigoroznost usmjerava na vlastito područje. Pokazujući da su stvarni učinci dovoljno mali da nestanu u lošem vrednovanju i da dosadna odluka poput normalizacije podataka može nadjačati pametnu novu arhitekturu, iznosi argument da velik dio napretka u učenju robota treba čvršća mjerenja prije nego što mu povjerujemo. Rad koji vlastiti kredibilitet troši na čuvanje granice između rezultata i želje radi nešto rjeđe i vrednije od osvajanja prvog mjesta na skali.

Kratak sažetak

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: slijep, randomiziran i velikog uzorka (≈1,800 eksperimenta u stvarnom svijetu 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 promijenili**, a izvedba je postupno rasla s količinom podataka za predtreniranje. Ali **bez dodatnog treniranja** nisu dosljedno nadmašivali modele za jedan zadatak, nekoliko učinaka 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, odmjerena podrška smjeru robotskih temeljnih modela — **ne** općenamjenski robot, ne zero-shot generalist i ne “emergentni skok” — uz oštro upozorenje da velik dio robotike možda mjeri šum.

Provjera bez uljepšavanja

Šta rad pokazuje: Uz rigorozno, slijepo i statistički dovoljno snažno vrednovanje (≈1,800 izvedbi u stvarnom svijetu + 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 promjeni 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 jezični enkoder bio malen); da se glatko skaliranje nastavlja izvan testiranog raspona podataka.

Šta ne pokazuje: Općenamjenskog ili zero-shot robota (bez dodatnog treniranja → nema dosljedne prednosti); ikakav “emergentni” skok sposobnosti; objašnjenja neuspjeha na konkretnim zadacima; apsolutnu pouzdanost u stvarnom svijetu (stope uspjeha podešene su blizu 50% radi osjetljivosti); išta o sigurnosti ili autonomnoj primjeni.

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

Koliko povjerenja treba imati opći čitalac? Visoko da predtreniranje na više zadataka uz dodatno treniranje daje stvarne, umjerene koristi — posebno u efikasnosti podataka i robusnosti — i da su one izmjerene neuobičajeno pažljivo. Visoko da ovo **nije** općenamjenski ili zero-shot robot i **nije** emergentni skok. Srednje koliko će se dobici nastaviti pri većim modelima. I vrijedi ozbiljno shvatiti upozorenje samih autora da su učinci u području dovoljno mali da studije s premalom statističkom snagom mogu prijavljivati šum. Primjeren stav: odmjeren 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/