This Week in Robotics

THE LINE

A single-author paper found a numerical bug in the default fine-tuning recipe of the library you are using, and the effect size is larger than any architectural result this year. Full-Weight VLA Fine-Tuning Without an fp32 Master Copy Silently Erases Most Updates (preprint, 2610.31859, full-text read) shows that when trainable weights sit in bf16 with no fp32 master copy, AdamW rounds away every update smaller than |w|/512 — and that 9 of 12 open VLA codebases ship a default or documented recipe that trains this way, four of them on full weight matrices, **including LeRobot's SmolVLA default and its own π0/π0.5 guides**. On LIBERO-Object, π0.5 goes from 5.1% to 80.9% with an fp32 copy at a ten-times-lower learning rate. The loss curve never tells you. 99.3% of 300 sampled SmolVLA fine-tunes on the Hub store the action expert in bf16. This is §6.

The second story is that evaluation honesty had its best week on record, and almost all of it is deflationary. SO-101 Bench (preprint, cs.CV, solo independent author, full-text read) shows an Isaac Lab twin of a USD 120–150 arm ranking five sim-trained policies exactly as the physical arm does (Spearman ρ = 1.00, MMRV = 0.000) while overestimating absolute success by 9 to 31 points. A paired evaluation of π0.5 inference settings on BEHAVIOR-1K finds eight of twelve ablations null, including dropping flow-matching denoising from 20 steps to 1 (+0.01 ± 0.02 — free), and fails to replicate a prior high-resolution-rendering gain. CuraEval proves by Parseval reduction that a published demonstration-quality score is episode length wearing a Fourier transform (Spearman 0.998–0.999 against length), and then shows that length-based selection flips from +10.0 to −1.0 points **on the same seeds** when you switch from max-over-checkpoints at 50 rollouts to final-checkpoint at 200. One log line claiming progress flips a VLM judge's verdict on robot video.

The third is that open hardware got a USD 400 bill of materials with retail links. OpenJaws (preprint, NVIDIA Seattle Robotics Lab) is a fully open sensorized UMI gripper at **USD 398.83 itemized**, against UMI's own 5,389, under a commercially permissive licence with hardware, software, data and sim assets released — and it publishes the result this thread has wanted for months: a controlled link from localization error to policy success, 10 mm to >30 mm training-set ATE dropping success from 90% to 30%.

Pin your optimizer dtype before you pin anything else.

Window: 2026-10-26 → 2026-11-01. Queue: 7 part files, ~340 scored lines, 48 at 4+, **six at 5** — against 25 at 4 and none at 5 last week.

Coverage note. The window had exactly one arXiv cs.RO announcement cycle worth the name: a 129-entry catch-up on 10-30 covering two weekend days, enumerated 129/129 via `papers.cool/arxiv/cs.RO?date=`, which is the only surface here that honours `&skip=N`. No cycle fired on 10-31 or 11-01, control-verified on two surfaces. Three gaps you should price into everything below. First, **cs.CV and cs.LG were never enumerated** — fourteen consecutive sweeps — and that is no longer a tail risk: cs.CV ran 233 entries on 10-30 alone, and two of this week's best items (SO-101 Bench and the UED-on-quadrupeds study) carry cs.CV and cs.LG banners respectively and were unreachable by cs.RO enumeration. Both arrived by accident. Second, IDs 2610.37667–2610.37798, a 132-ID hole, were never announced as cs.RO new submissions and are not verified as swept. Third,

the 10-22 cycle count (arxiv.org 120 entries/69 new against papers.cool's 78) is unreconciled for a fifth week and I am retiring it rather than carrying it further; it has produced no new evidence since W43.

One method note that is itself a finding: the harvester logged two papers' categories from the surface that surfaced them rather than from the PDF banner, and both were wrong. Set source type from the banner.

---

1 Policy learning & VLAs

SO-101 Bench — the sim-eval question, answered for the arm you own

SO-101 Bench: Accessible Spatial Manipulation — preprint, arXiv 2610.36420v1, 27 Oct 2026, cs.CV banner, Truman Hickok, independent, no affiliation. Full text read. Code and Hugging Face assets linked from the paper.

What is architecturally new: nothing — and that is the point. The contribution is an Isaac Lab digital twin of a real SO-101 setup, 138 household objects split into seen (48), unseen-instance-seen-class (43) and unseen-class (47, of which **37 are kept private** to stop test-set leakage), four tasks scored as five families from simulator state, 916 feasibility-checked primary-evaluation episodes over 15 task × split cells, plus ROBUST (178 episodes, held-out backgrounds and robot colours), PLAN (48, rotate-to-fit) and MEM (51, remember an occluded referent) subsets. Physics at 240 Hz, policy at 30 Hz, overhead plus wrist camera at 640 × 480.

What the eval actually covers, and this is the part worth your attention: **the direction of the sim2real claim is the unusual one**. Most real2sim work drops real-data-trained policies into a twin. Here five policies that never saw real data from the setup — four VLAs and an LLM agent — run with *identical weights* in the twin and on the physical arm, in an environment absent from training, on the same 45 layouts with object footprints projected into the overhead view so both domains start matched. The ranking is preserved exactly: agent 84%→64%, Galaxea 0.5 67%→53%, GR00T N1.7 56%→24%, π0.5 49%→18%, MolmoAct2-SO101 24%→16%. ρ = 1.00, MMRV = 0.000, Pearson r = 0.90 [0.66, 0.97]. Separately, independently trained real-only and sim-only GR00T N1.6 policies find the same cells hard (r = 0.86).

The leaderboard is the deflationary half. Nine learned policies, every one fine-tuned on the same 3,629 simulated demonstrations under an equal 36 h × 4 × A100 budget. Best learned: InternW0-Δ 61.4 [58.5, 64.1] and Galaxea 0.5 60.8 [57.9, 63.6], statistically tied. Then GR00T N1.7 45.8, ImageWAM 43.0, π0.5 41.0, LaWAM 39.9, both MolmoAct2 variants under 25, FLUX 3 Action 8.0. Binning one familiar object is 90–93%; next-to and between are 34–44% for the same policies; InternW0-Δ falls 72%→18% on next-to from seen objects to unseen classes. **No learned policy exceeds 17% on PLAN.** Starting MolmoAct2 from its SO-100/101-post-trained checkpoint does not beat its general checkpoint (20.9 against 23.8, overlapping intervals) — prior embodiment exposure is not a substitute for task coverage.

Then the uncomfortable result: an LLM agent (GPT-6.1 Sol writing code in the PyRUA-Lean harness, handing grasping to Galaxea 0.5) scores **83.2** and solves 75% of PLAN, and the harness ablation localises the lead. On 60 paired episodes: 35% with direct tool calls, 67% with PyRUA-Lean, 85% with the Galaxea hand-off, against 53% for Galaxea alone. It holds on hardware, 64% against 53%.

Where it is soft, stated by the author: five policies and 45 layouts in one real environment, so MMRV's 95% CI runs [0.000, 0.169] and ρ = 1.00 on five points is five points. The fixed 36 h budget left MolmoAct2 (0.5–0.6 epochs) and FLUX 3 Action (0.2) visibly undertrained with loss still falling, so the bottom of the leaderboard measures the budget, not a ceiling. The agent's time limit counts control steps only — the arm holds still while the model thinks, and median wall time to rollout stop is 197 s for the agent against 38 s for Galaxea 0.5. Much of the real-robot bin-task gap is a lighter plastic bin, and many large objects were scaled to 0.7–0.9× in sim. One arm, rigid objects.

Reproducible at one-arm scale: yes for *evaluation*, which is the point — the twin, the datasets (`so101_bench_sim`, 3,629 episodes; `so101_bench_real`, ~4,400 episodes) and eleven pinned checkpoint repos are released. Fine-tuning a policy to leaderboard standard is 36 h × 4 A100; running the benchmark against an existing checkpoint is not. The paper's own evaluation hardware includes a workstation RTX 3090 and rented A5000s.

The inference-settings question got a third-party paired answer

What Actually Matters? A Paired Evaluation of Inference Settings for π0.5 on BEHAVIOR-1K — preprint, 2610.37917, Kiel + George Mason + Ewha, full-text read, cs.RO banner 27 Oct. Third-party benchmark under BEHAVIOR-1K's official rules on the public test split, which is rarer than it should be. Twelve inference ablations, eight null. The two that matter: **denoising 20 steps to 1 is null at +0.01 ± 0.02**, so the entire flow-matching sampling budget is free money on this benchmark; and chunk 26 → 13 costs ΔQ −0.16 and 11 points of success while adding 752 steps, so chunk length is where the cost actually lives. Best-of-4 sampling: null. Advantage guidance at β = 2: worse, −0.03. A prior high-resolution-rendering gain **did not replicate**. Snapshot-restart before the first progress drop gains +0.076 and cuts 29% of sim steps but does not raise successes per sim step, and is net harmful unless stage belief is restored with sim state. Soft: nothing released beyond videos, compute undisclosed, zero real episodes, and the control is only two baseline runs per cell, which is thin for a paper whose contribution is statistical discipline.

Read it against Coupling Execution Horizon and Sampling Depth in Flow-Matching VLA Policies (preprint, 2610.32102, project page only, abstract-level), which reaches a compatible and more specific claim: coarse integration *hurts* frequent replanning and *helps* long open-loop execution. Both are training-free and act on knobs you already have. **INFERENCE** \2192 the two together are the first genuine convergence in the replan-latency cluster, which has produced eleven-plus papers in six weeks with no cross-comparison — and they agree that the denoise budget is the cheap knob and the chunk horizon is the expensive one.

World modelling bounded a second time, independently

What Does World Modeling Actually Improve in Robot Action Learning? — preprint, 2610.34593, CASIA, abstract-level, no code. Controlled ablations find future-video-prediction gains on seen tasks **do not survive distribution shift**. That is the same shape as W43's AW-0 result, where hybrid code-as-policies execution scored 20.6 under randomization against plain π0.5's 27.7 — but from a different group with a different method, which is what the thread needed. Flag as unreproduced single-lab until someone runs it, but note that SO-101 Bench's leaderboard is independent corroboration at the ranking level: WAMs neither dominate nor collapse there (InternW0-Δ 61.4 top, ImageWAM 43.0 above π0.5's 41.0, LaWAM 39.9).

The constructive counterweight is **GAWM** (preprint, 2610.37091, abstract-level, **code released**): predict the future as compact tokens in a frozen DINOv3 latent and regress chunks, **under 100M parameters**. A sub-100M world-action model with released code is the first thing in this cluster a one-GPU builder can actually train end to end, and it is the only item in the WAM pile this week that clears that bar.

Cross-controller transfer without SysID

Cross-Controller Transfer via In-Context Behavior Adaptation — preprint, 2610.39280, Stanford (Finn, Sadigh), cs.RO banner 29 Oct, full-text read, project page only, no code, no licence. Architecturally: condition a π0.5 backbone on three calibration trajectories of **commanded-versus-achieved pairs** rather than on SysID gains or ground-truth controller metadata. 2,880 real trajectories, 4 tasks × 6 controllers × 4 gain settings, all real, one WidowX at 50 Hz, no sim. Zero-shot on a held-out controller: 50.3% against 17.5%, beating SysID at 27 and ground-truth metadata at 12, with a same-controller upper bound of 75. The ablation is load-bearing in the right direction: **relabeling alone hurts**, 0.24 → 0.15. Soft: 20 trials per cell gives roughly ±11 pp standard error, so only the aggregate holds; no seeds. The reproduction cost here is the data, not the GPU — 2,880 real trajectories across six controller configurations is the expensive part, and that is a useful thing to know about a result that looks cheap.

---

2 Manipulation, dexterity & sensing

OpenJaws — preprint, 2610.36598, 27 Oct, NVIDIA (Murali, Sundaralingam, Birchfield, Coumans), open-jaws.github.io, full-text read. This is the hardware item of the week and it is unusually complete. A modular 3D-printed sensorized UMI gripper with a **low-cost global-shutter fisheye stereo pair** instead of monocular VIO, exposing an **ISO 9409-1-50-4-M6 flange** so it mounts on common collaborative arms without a bespoke adapter, built deliberately with **no custom PCB** so that nothing has to be

fabricated before you can build one. BOM is itemized with retail prices and source links: Dynamixel XC430-W150-T 137.89, stereo camera 99.99, OpenRB-150 28.64, down to the wire strippers — **USD 398.83 total, 0.408 kg**. Against UMI's 5,389 (including a WSG-50), FastUMI Pro's 840, LEGATO's 1,218, DexUMI's 1,316.

The localization result is the technical core. On their new UMI-Loc-Bench-v1 (90 handheld demonstrations, nine tasks, AprilTag-cube ground truth with the cube *masked* so methods cannot track the fiducials), stereo COLMAP reaches 5.2 mm / 0.8° full-trajectory ATE and cuVSLAM stereo 12.4 mm / 1.9°, against UMI's own modified ORB-SLAM3 mono+IMU pipeline at 18.4 mm / 4.0° — and the coverage numbers are the brutal part: **all stereo methods track essentially the whole trajectory; ORB-SLAM3 mono+IMU covers 9% and UMI-SLAM 37%**. ViPE, the learned monocular model, recovers trajectory geometry well (3.5 mm scale-fitted) but its native-scale ATE is 48.2 mm with a median fitted scale of 0.556, so it cannot produce metric action labels. Adding an IMU to stereo ORB-SLAM3 improves mean ATE 5.1 → 2.0 mm but drops coverage 100% → 18%.

Then the ablation that earns the paper its place in this ledger: inject synthetic piecewise-constant drift to raise training-set translational ATE from ~10 mm to >30 mm, and **policy success falls from 90% to 30%** on the same task. That is a direct, controlled measurement of what label noise costs, and it is the first time this ledger has seen the localization-quality-to-policy-success link quantified rather than asserted. Also notable: a baseline-calibration sweep showing a ±5 mm error on the 59.7 mm stereo baseline moves mean ATE from 3 mm to 6–8 mm while rotational ATE moves 0.1°, so baseline error is almost purely a scale error.

Soft: policy evidence is thin — Diffusion Policy on a UR10, 200 demos for block stacking and 50 for pick-and-place, **90% over 10 trials each**, which is one seed and ten rollouts. One failure per task. No VLA or WAM was run on the data, which the authors flag. The agentic real-to-sim pipeline (GPT-6 Astra building Isaac Lab twins from demonstrations with no ground-truth object poses) succeeds on 7 of 10 gear-insertion demonstrations, which is interesting and unvalidated. Reproducible at one-arm scale: yes, more so than anything since OpenTac — USD 399, a printer, TPU for the fingers, and an arm with an ISO flange.

Language-Grounded Force/Torque Pretraining for Extrinsic Contact Perception — preprint, 2610.39314, full-text read, **code and dataset released**. Six-axis force history alone, aligned to frozen sentence embeddings, beats a supervised multi-task baseline on extrinsic contact perception. One-GPU reproducible, sim-to-real, and it publishes its own load-bearing negative: **the frozen readout fails**, so the alignment is doing work the frozen features cannot. This is the week's best force item on artifact grounds and the strongest candidate the *force-as-first-class-input* thread has had with code attached.

Force-Driven Hook Disassembly via Constraint-Aware Object-Centric RL — preprint, 2610.38513, Osaka + AIST, cs.RO banner 28 Oct, full-text read. Force is the **sole exteroceptive input**; a particle-filter belief over extraction direction feeds both policy and reward. Zero-shot sim-to-real onto a real end-of-life appliance, 100/100/90/75% across four tasks, with the belief ablation lifting cover success 45% → 75%. Honest about the mess: path efficiency 29–47% with 49–139 direction reversals (the motion is jittery), the deployment filter was retuned against the training filter without retraining the policy, failure mode 1 is the *grasp choice* rather than the controller — and a plain force-following baseline did better there. Trained on one generic star peg at 0.55 mm tolerance, deployed on unseen fasteners. **Nothing released**: no code, no environment, no project page, no promise, and no compute numbers at all. Self-built task suite, n = 20 per task.

CARD: Contact-Anchored Retargeting for Dexterous Manipulation (preprint, 2610.36539, CMU + Amazon FAR + UMD + Berkeley, abstract-level, no code) retargets human demonstrations while **preserving contact force targets** through a shared hand-contact representation and a tactile tracking controller. TouchBridge (2610.40349, full-text read, no code) does tactile-conditioned human-to-robot transfer with no object mesh, pose estimate or physics simulation required, and reports pretraining as load-bearing. FLoRe (2610.35862, Amazon FAR + NYU + CMU, abstract-level, no code) trains a diffusion policy on force-sensing handheld demonstrations with a hybrid compliance whole-body controller — force commanded, not inferred post hoc. Four papers in one week where the measured wrench is the retargeting or control *reference*; none with code.

SwapBot — preprint, 2610.40646, full-text read, one-arm scale. **USD 10 3D-printed tool adapters with embedded tactile sensing** beat both dexterous hands and universal grippers in a low-data regime. Open-sourcing is promised and not delivered, which on this ledger's record is a coin flip. If it lands it is the cheapest tactile end-effector story in the file.

GelSlim 5.0 exists — flag closed. W43 flagged GelSlim 5.0 for a full read while noting that no 5.0 repo, paper page or cost had surfaced. gelslim5/gelslim5-vistacfusion (release, 2026-10-29, **CC-BY-4.0**, card read in full) resolves it: 55,069 frames, 4,669

real plus 50,400 simulated under one loader, 12 objects, a documented real-to-sim ratio ladder, ground-truth contact depth in metres, 17.5 mm square tactile FOV, UV-responsive contact layer. No evaluation numbers in the card. One-GPU sized and permissively licensed, which makes it the cleanest visuo-tactile set with a real *and* simulated half this ledger has recorded.

Two data rigs released rather than promised: **UMI-Live** (2610.39385, full-text read) ships self-contained bimanual UMI grippers needing no tracking rig, with hardware, code and **40 hours of episodes** released — though the policy evidence is one seed on one task. And amirsh1376/Skill-UMI (release, created 28 Oct, **CC-BY-4.0, not gated**, card read in full): 397 episodes, 322,114 frames, three tasks, a RealSense D455 at the wrist plus **two GelSight Minis, one per finger**, per-frame sub-task labels, paper-backed by arXiv 2609.38494. Stored in a Zarr/OXT-FTP-1 layout rather than LeRobot, so budget a conversion. No evaluation numbers.

---

3 Platforms & embodiment

Demonstrated, with the demonstration honestly bounded. **SONIC-K1** — preprint, 2610.39327, 29 Oct, Vilnius + Kent State, full-text read. NVIDIA GEAR's SONIC whole-body motion tracker was trained on **128 GPUs for roughly 3 days on 700 h of in-house mocap**. This port to the 22-DoF Booster K1 keeps SONIC's architecture, PPO settings and rewards unchanged, retargets 9,991 Bones-SEED clips with UMR, and trains from scratch for 50,000 iterations with 16,384 environments on **one shared H100 in 135.4 hours** — about 103 h had the GPU been theirs alone, and roughly **1–1.5% of SONIC's compute**. It tracks 97.4% of 1,985 held-out clips in MuJoCo under SONIC's own 0.25 m success rule at 15.4 mm MPJPE-L, 93.6% on PHUMA, 99.0% on 100STYLE. Code **Apache-2.0**, weights **CC BY-NC 4.0**, evaluation tools released; the motion data is deliberately not redistributed.

The paper is in this section rather than §6 because it is scrupulous about what it has not shown. **The real-robot evidence is qualitative**: upper-body and slow whole-body motions run smoothly on hardware, stepping and turning are visibly jittery, and there is no motion capture for the robot so no hardware tracking error is reported. One seed, no ablations, one hyperparameter set. The hold-out comes from the same dataset as training. The G1 columns copied from the SONIC paper are explicitly labelled as scale-setting rather than head-to-head. And the robustness sweep locates the likely cause of the leg problem: the policy survives PD gains from 0.6× to 1.4×, 70% motor strength, a 2 kg payload, 3× observation noise and 3× dense pushes at ≥97% no-fall, but **latency falls off a cliff — 98.5% at 20 ms, 56.0% at 40 ms** — and they did not model actuation delay in training.

Which is exactly the gap that **Delay Randomization Widens the Delay Margin of Learned Humanoid Motion Trackers** (preprint, 2610.32277, abstract-level, nothing released) fills: randomizing 0–50 ms command delay during training lifts average survival at 40 ms delay **from 15% to 95%**, with an ablation confirming training without delay randomization does not produce the gain. Two papers in the same window, one with a 40 ms cliff and one with the one-line training-time fix for it, neither citing the other. **INFERENCE** \2192 this is the cheapest sim-to-real intervention in the file and the mechanism is embodiment-agnostic — if you have a comms or inference lag anywhere in the loop, randomize it.

The deflationary curriculum result. Is It the Curriculum or Just the Easy Start? — preprint, 2610.40364, **cs.LG banner** 29 Oct, full-text read, **code MIT**, model Apache-2.0. Adds two controls separating "trained on easy levels" from "actually curated levels" across 14 unsupervised-environment-design methods on four quadrupeds. Domain randomization on uniform levels appears in **zero of 20 top groups**; plain Terrain-CL (Rudin 2021) is in 13 of 20 and ACCEL in 11. But the controls do most of the work: DR → DR-easy on Barkour is 20.9% → 59.5%. Honest limits the authors state themselves: one PPO student at one hyperparameter set; Go1 and Barkour curves are **still rising at 1B steps**, so those are budget rankings not convergence rankings; and on Spot all 14 methods fall within 3.8 points and the correlation with every other robot is negative (ρ −0.61 to −0.28), which reads like a floor effect rather than a finding. At 0.5–0.9 GPU-h per run and 306 GPU-h for all 350 runs on one H100, it is also the only one of the nine studies in this cluster you could rerun.

Staged, or video-only. AGIBOT coordinated 110 production humanoids in a Guinness-certified dance (trade press, 10-24 event): every technical detail — 5G fleet control, centimetre UWB, millisecond sync, formation readjustment on dropout — is company-claimed via press, with no paper, code, dataset or product. The only claim worth anything is that these were standard production units. DeepMind paired Genie 4 with Gemini Argon to drive a simulated Apollo playing checkers (trade press): research demo, no

paper, no code, simulated embodiment. XPENG says IRON's solid-state battery entered small-batch production (company claim via trade press) with no capacity, mass or price. Audi is piloting Agile Robots' mobile humanoids for parts sequencing (trade press): named automaker, named vendor, no specs and no throughput.

And a correction to carry: part 6 of the queue logged a report that Tesla would open Optimus reservations in early December. Musk said no on X within a day (trade press quoting a principal's denial, 10-31T03:58Z). The original claim was secondhand investor commentary; Tesla operates no first-party dated newsroom, so nothing in this area can be primary-sourced from here. Treat the whole Optimus reporting surface accordingly.

Purchasable, allegedly. Two vendor spec cards bracket the listed price floor for whole humanoids and neither has a primary source: Mantis Standart at **USD 1,400** (140 cm wheeled, 50 kg, Blue Worm, Xi'an — an order of magnitude under the usual wheeled humanoid and about ten times an SO-101 arm; no payload, torque, DoF, SDK or ROS 2 detail, so the price is the only datum) and Lumen H1 at **USD 40,000** (169 cm, 69 kg, Wuhan). Both are spec sheets with no manufacturer URL verified. The 1,400 figure is low enough that I would treat it as unconfirmed rather than cheap.

---

4 Open source & tooling

Ranked by what reduces friction this month.

1. **LeRobot PR #5107** — merged 10-26. Dataset running statistics were accumulated in float32, which gave **zero or biased normalization std**. That silently corrupts the normalization of any dataset you build. Together with §6 this makes two numerical-precision bugs in LeRobot's default path found in one week, both of which fail quietly and neither of which shows up in a loss curve.

2. **Isaac Lab's reset and write paths are actively unreliable — three correctness fixes in three days.** #8634 (2026-10-30T22:01:21Z): additive observation-noise bias was **accumulating across episode resets**, so any RL run relying on noise randomization was training against drifting noise statistics. Results from before and after this commit are not comparable. #8616 (2026-10-31T08:16:25Z): `reset_joints_within_limits_range` was clamping joints it was not asked to touch. #8831: scalar joint-friction writes were wrong in **both** PhysX and OvPhysX. Same-day companions to #8634 include FORGE Euler ranges during action clipping (#8804) and `fix_root_link` for articulations rooted at a fixed world joint (#8164) — the second of which matters specifically for a single statically mounted arm. If you tuned contact-rich behaviour on Isaac Lab before 10-31, that tuning is suspect. Note also that `/tags` reads `v3.0.0-beta2` at head, which is **tag-name ordering, not recency**: GA v3.0.0 shipped 2026-10-20 and the ledger had it right while the queue reopened it for three sweeps.

3. **MuJoCo: three same-night numerical-robustness fixes in the contact and solver path**, no release cut. a63f591 (elliptic cone Hessian T^3 underflow), 8d60957 (Cholesky NaN guard), c4137ef (implicit actuator derivatives under tendon saturation). Ten commits in window, nine after the 10-31 cutoff. Separately c835c40 ships agent skills bundled with Menagerie, testdata, notebooks and `mujoco.batch`, and cb53b2b documents a prebuilt standalone Studio desktop app. Tag head still **3.17.0** — after two model-affecting releases in nine days last month, a quiet fortnight is good news.

4. **LeRobot: native Long-WAM training and inference in-tree** (#5008, 2026-10-29T14:35:12Z, plus the inference frame-resize device fix in 9069415/#5144 that unbroke it). A world-action-model policy head trainable from the standard CLI with no fork. This is the clearest confirmation yet of W43's **INFERENCE** `\2192` the release cadence is dead and the code surface is not, so watch `src/lerobot/policies`.

5. **LeRobot's 10-30 batch, four behavioural defaults worth checking before you trust a loss curve.** Thirteen commits to main; head 3f797de at 2026-10-30T20:10:40Z. #5096 makes TF32 configurable — **check the new default before comparing loss curves across this commit**. #5117 adds native EMA with pluggable decays, which means **the exported checkpoint becomes the EMA copy**. #4974 changes the dataset loader path for Lance episode-subset indexing to

O(num_episodes). #5112 keeps GR00T N1.7 camera frames on device: fewer copies, higher VRAM. Also #5103 warmup-stable-decay scheduler and #5156 rollout latency and memory logging.

6. **lerobot-omx/omx-act-models** — release, 2026-10-30, **Apache-2.0**, card read in full. Four ACT checkpoints for one OpenManipulator-X arm with something almost nothing on the Hub carries: **a real success rate and honest compute**. 16 of 30 physical rollouts succeed; 50k steps at batch 8, **6 h 32 m on one GPU**, final loss 0.082, trained on 50 episodes/25,445 frames and 20 episodes/23,327 frames. It documents a 0→2k→4k→20k fine-tune adaptation ladder with per-step loss and the resume gotchas (delete `optimizer_fused` on older LeRobot; raise steps or resume exits instantly). The most friction-reducing single artifact in the window for anyone with one small arm.

7. **DexSuite** — preprint, 2610.37799, Mila/McGill + Hugging Face + Kinova + Amazon FAR, full-text read. Ten arms × twelve end-effectors composing at runtime, **4,700 episodes released in LeRobot and HDF5 format**, FlexiTac tactile simulation at 12×32 on 2 mm pitch. The tooling is the contribution and it is real. Discount the numbers: self-built benchmark, **zero real episodes**, **no licence stated**, the tactile simulation is penetration-to-force and not validated, and the headline hand-versus-gripper gap is a null that flips with policy choice. To their credit they publish their own noise floor (7 pp at 50 episodes) and flag that ACT's in-training max exceeds a fresh seed by 7 pp, which is checkpoint-selection bias named out loud. The transferable idea is the embodiment-independent TCP plus workspace AABB.

8. **CuraEval** — preprint, 2610.38221, Colorado Mines + USC ICT + Northeastern, full-text read, code released under **MIT** — correcting the harvester, which logged no licence; the repo root carries a LICENSE file and the README states MIT. On a third-party benchmark (robomimic multi-human), 267 training runs and 44,100 evaluation episodes. Headline: a published PSD-based demonstration-quality score reduces by Parseval to $T^2 \times$ mean squared position — Spearman **0.998–0.999 against episode length** — and selects identical 150-demo subsets on the square task at 50%. Geometric-entropy leave-one-out is **actively harmful** (41.7% against 62.5%, below all three random draws). **No curated condition beats full data after Holm correction in any setting.** The protocol finding is the one to remember: length-based selection flips from +10.0 to −1.0 points on the same seeds when you move from max-over-checkpoints at 50 rollouts to final-checkpoint at 200 rollouts, which implicates FAKTUAL (max over 20 checkpoints) and CUPID (last ten) directly. CUPID also costs 36 A40-GPU-hours and its three scorers correlate only 0.16–0.31 with each other. The authors state their Holm families were specified retrospectively, which is the right disclosure and also the paper's main softness, alongside zero real episodes.

9. **Boyun7/LEAP-VLA** — release, created 2026-10-25, **MIT**. A 1B VLA (InternVL3-1B base plus a latent action aligner and LoRA) with per-benchmark checkpoint folders, `dataset_statistics.json` for action denormalization, a runnable inference snippet and a full results table in the card (LIBERO Spatial 98.4 / Object 99.4 / Goal 97.6 / Long 93.6; LIBERO-plus 78.3 across 10,030 perturbed tasks, which is the only robustness number here worth anything). **One billion parameters is genuinely one-GPU reachable**, unlike the 6.4B lingbot and 16B Cosmos families, and MIT is permissive where last month's releases were Gemma-licensed or noncommercial. Caveat: zero downloads and zero likes, so nobody has reproduced it, and every number is author-reported.

10. **Luxonis OAK 4 FFC** — trade press, shipping now. Five camera modules on one vision computer, onboard stereo depth, 52 TOPS, 8 GB, 2.5GbE, **USD 849 dev kit**. The only in-window priced, shipping hardware item inside a small-build budget.

Negative confirmations, in window. LeRobot still **v0.6.1** of 2026-08-03 — the **eleventh consecutive negative**, seventeen weeks post-deal — confirmed via `/tags?per_page=10` rather than `/releases`, which has served stale data on three independent repos now. MuJoCo 3.17.0. Genesis v1.4.5. ManiSkill v3.0.1 from April. RoboCasa v1.0 from February. `openrobotics/reps` **zero commits** — and the checklist repair here matters more than the negative: `ros-infrastructure/rep` is archived and frozen since 2025-10-23, so the old target returns empty forever and would read as "no REP activity" indefinitely. REP 0157 (rosidl redesign for native zero-copy transport) has a dedicated PMC meeting on **2026-11-04**, three days out. ROS Discourse also carried Space Station OS v0.9.0 (ROS 2 Jazzy, Pixi+RoboStack locked environment, 144 tests passing, no licence on the post) and a TU Delft mirror outage.

5 Money & moves

Physical Intelligence is reportedly raising USD 1B at roughly a USD 20B valuation — trade press, reported talks, not a closed round; the figure is attributed to Bloomberg and widely syndicated, with Avenir and Lux named as leading. **Sources conflict and I am not resolving it**: a Dealroom note dated 2026-10-14 reports the same USD 1B at **USD 11B**. Either the valuation nearly doubled inside two weeks of the same raise, or one of the two numbers is wrong. What the signal says regardless: the lab that shipped the first frontier-robot-model API on 10-13 is being priced as infrastructure, not as a research group, and no round has closed.

Agility expects about USD 465M gross at SPAC close, down from an earlier USD 620M estimate — **primary source**, company newsroom, 2026-10-30. Roughly 164M from trust, 201M in a Foxconn-led PIPE at ten dollars, 100M already received, against a USD 2.5B deal value. This is the first time in this ledger that a humanoid capital event has priced **down** against its own earlier guidance, and a 25% shortfall driven by trust redemptions is a public-market verdict rather than a private mark. Set against GXO expanding the Agility partnership to two more US sites starting with Digit 4 (trade press) — repeat customer, not a demo — the read is that the commercial traction is real and the financing enthusiasm is cooling, which is a different and healthier combination than the reverse.

Bleu Robotics, EUR 5M pre-seed — **primary source**, company blog, 2026-10-29. Contrarian Ventures and OVNI Capital co-led; founders are Benoit Berkoukchi (ex-Renault) plus **Jean-Baptiste Mouret and Serena Ivaldi, both ex-Inria**. The thesis is sharp enough to be falsifiable: ten demonstrations per workstation via handheld grippers, roughly **40 minutes** to a trained skill, training and inference **onboard**, demonstrated on Unitree G1 and Franka Research 3, rigid objects only with textiles explicitly out of scope. Two ex-Inria learning researchers taking pre-seed money to bet that industrial humanoid skill acquisition is a 40-minute handheld-gripper problem is the most specific small-data claim any funded company has made in this ledger — and the handheld-gripper mechanism is the same one OpenJaws and UMI-Live just opened up.

Machenta, USD 40M to teach robots tasks from one demonstration — trade press, no method disclosed. Read alongside Bleu: USD 45M across two companies in one week for *demonstration-efficiency as a product*, with the smaller of the two the only one stating a mechanism. INFERENCE \2192 capital is moving from "collect a fleet-scale corpus" to "need fewer episodes", which is the thesis a one-arm builder is forced into anyway.

A3: North American robot orders up 24.8% in Q3 2026 — press release, 2026-10-28. 10,989 units and USD 723.5M; **cobots 1,831 units, up about 56%, now 16.7% of all units**. Semiconductor and electronics +97.6%, automotive components +90.4%, automotive OEM +7.5% in-quarter but **−17% over nine months**, food and consumer goods down. The cobot share is the number to hold: the fastest-growing category in the installed base is the one whose arms accept an ISO 9409-1-50-4-M6 flange, which is exactly what OpenJaws standardised on this week. ROKAE reports 10,000 humanoid arms delivered and names a Schaeffler Taicang deployment (trade press, vendor-stated volume, unaudited).

Boston Dynamics employees petitioned internally against a Spot sale to ICE — trade press. Two lines: this is a retention and culture risk at a named team with deep whole-body control expertise, and it is the first item in this ledger where deployment ethics shows up as a **talent-market** variable rather than a policy one. Where that expertise goes if it leaves is the thing to watch.

Out of window but easy to misfile, logged so the window is not silently filtered: Bedrock Robotics USD 600M Series C at USD 4.1B (2026-10-26, construction autonomy), Foundry Robotics USD 100M Series A (primary 2026-10-28T15:00Z, syndicated 10-29), Opentrons USD 40M Series E (lab automation), Rendezvous USD 20M and Plato USD 16M (off-thesis). And the W43 Generalist USD 400M chase is closed: the round dates to **2026-06-04**, five months stale, rendered as an undated current lead by a trade surface that must not be used as a recency signal for any purpose.

6 Deep read of the week

Full-Weight VLA Fine-Tuning Without an fp32 Master Copy Silently Erases Most Updates — preprint, arXiv 2610.31859v1, submitted 22 Oct 2026, announced in the 10-26 cs.RO cycle. Shubham Singh, UT Austin, single author. I read the

paper in full. Note the date: the submission predates this window, the announcement does not, and it sits above W43's band top of 2610.31670, so this is first coverage rather than a repeat.

I picked this over SO-101 Bench and SONIC-K1 because it is the only paper in the queue that changes what you should do before your next training run, and because its evidentiary standard is the highest in the file.

Setup

Task and embodiment: LIBERO-Object, 454 demonstrations, 500 tries per setting (10 tasks × 50 fixed initial states), in simulation only. Two models: SmolVLA (0.45B, 99.9M trainable) fine-tuned from `smolvla_base` with LeRobot at commit b9cb121 for 10,000 steps at batch 16; π0.5 (4.1B, 4.14B trainable) from the general `lerobot/pi05_base` for 2,500 steps at batch 32. AdamW with $\beta = (0.9, 0.95)$, LeRobot's warmup and cosine schedule, bf16 autocast throughout. Three seeds per SmolVLA setting, two per π0.5 setting, 26 runs in total.

What varies is only **how the trainable weights are stored**: all-bf16; LeRobot's own recipe (SmolVLA's default, which keeps 3.3M in fp32 and 96.6M in bf16; or the `--policy.dtype=bfloat16` flag from LeRobot's π0/π0.5 guides, which stores 3.61B of 4.14B in bf16); or an fp32 copy. Both π0.5 settings load through the bf16 flag and the fp32 condition casts back, so both start from identically rounded weights.

What is held out, and this is unusual enough to state plainly: **the rules were fixed before the runs**. A dated project log recorded each threshold, prediction, amendment and deviation before the runs it covered — a 10-point minimum effect size that also had to exceed the largest within-setting seed gap, a probe criterion, and paired McNemar tests. The log was not filed with an external registry, and the author says so. Two predictions are recorded as wrong: a small-or-no gap was predicted at 10^{-4} and the measured gap was 10–14 points; a sub-3-point re-planning gap was predicted at probability 0.45 and came in at 4.2.

Method — what is genuinely new

The mechanism is arithmetic, not machine learning. bf16 carries 8 significant bits, so the representable neighbours of a weight w are $|w|/256$ to $|w|/128$ apart; round-to-nearest returns w unchanged for any update below **$|w|/512$**. An update of 10^{-5} therefore cannot move any weight with $|w| > 2^{-8} \approx 0.004$, which covers **88% of SmolVLA's action-expert weights**. Mixed-precision training avoids this by keeping an fp32 master copy and rounding only the compute copy; PyTorch's AdamW, handed bf16 parameters, keeps its moments in bf16 and adds the update straight into the bf16 weight.

Inherited: the observation that bf16 storage loses small updates is old (Micikevicius 2017; Zamirai 2020) and has been reported for LLM RL post-training, on-policy distillation, GRPO and ternary models. Two prior robotics datapoints exist — GuidedVLA reports full-bf16 π0 at 10 against 90, and an OpenTau issue reports π0.5 on LIBERO at 0% without and 94% with fp32 masters — but both are single runs at one learning rate.

Genuinely new, and the reason this is a paper rather than a bug report, is three things. First, **E(u) is turned into a lower bound rather than an estimate**: since rounding is monotone, if no element of a step moves more than $c\eta$, then at least a share $E(c\eta)$ of weights is unchanged, where Cauchy–Schwarz bounds $|\hat{m}|/\sqrt{\hat{v}}$ by $c = 1.165$ for $\beta = (0.9, 0.95)$ — though for PyTorch's default $\beta_2 = 0.999$ the constant is 7.3 and E degrades to an estimate, which the paper marks in its table. Second, **a per-step probe** that copies the trainable weights before every 100th optimizer step and counts unchanged elements, validating the bound empirically at every probe. Third, **an audit at two scales**: 12 open VLA codebases read at pinned commits with the LeRobot and openpi rows confirmed at runtime, plus a Hub survey of 300 sampled SmolVLA fine-tunes and 300 π0.5 fine-tunes read from file headers without downloading.

Results — the number that matters, and the baseline

The baseline is LeRobot's own default. On LIBERO-Object with 50-action open-loop chunks, against LeRobot's recipe, the fp32 copy adds:

- **SmolVLA at 10^{-4} (the default): 59.3% → 69.5%, +10.2 points**, seed spread 2.2
- **SmolVLA at 10^{-5}: 4.3% → 49.3%, +45.1 points**, seed spread 4.2

- **π0.5 at 2.5×10^{-5} (the default): 89.0% → 97.7%, +8.7 points**, seed spread 1.6 — which removes 79% of the recipe's remaining failures
- **π0.5 at 2.5×10^{-6}: 5.1% → 80.9%, +75.8 points**, seed spread 3.4

Against all-bf16, SmolVLA gains 14.3 points at 10^{-4} (from 55.2%) and 49.2 at 10^{-5} (from 0.1%). Every run with the copy beats every run without it, and the copy scores higher on all 10 tasks in every comparison. Paired over the 500 shared initial states: at 10^{-4} the copy wins 370 tries where the recipe fails against 217 the other way (exact McNemar $p = 3\times10^{-10}$), and 701 against 25 at 10^{-5}. The probe measures **79–99% of updates erased at 10^{-4} and 96–99.95% at 10^{-5}**; for π0.5, 96.7–99.7% of gradient-receiving bf16 weights per probe.

The audit numbers are the part that generalises. Nine of twelve codebases ship an affected default or documented recipe; four train full weight matrices and lose 12–94% of learning-rate-sized updates (LeRobot, openpi's PyTorch trainer, RLinf, Pion). **289 of 291 sampled full-weight SmolVLA fine-tunes (99.3%) store the expert in bf16**, and in 40 of them a median 22% of elements across six expert weight matrices are still **bit-identical to the base model after training**, against 0.1% for a model trained from scratch. Of 300 π0.5 fine-tunes from the last eight weeks with a training config, 252 set the bf16 flag.

Two findings from the runtime checks are worth repeating because code reading missed them: LeRobot's SmolVLA **accepts `--policy.dtype=float32` and ignores it**, because the action expert inherits bf16 from the vision-language model's config; and LeRobot's π0/π0.5 guides add `--policy.dtype=bfloat16` and *call it mixed precision*, while the library's own accelerator page documents the correct fp32-weights-with-bf16-compute pattern that those guides do not use.

Where it is soft

The author's own limitations section is accurate and I will not improve on it, but three things deserve emphasis.

The scorer fails its own check for SmolVLA. It validates on π0.5 (99.2% against a published 99.0%) but scores LeRobot's published `smolvla_libero` at 67.6% with 50-action chunks against the 88% required in advance, and 93.4% when re-planning every 10 actions. A second check on `HuggingFaceVLA/smolvla_libero` scored 81.4% on five and a half tasks before being stopped, and its shortfall is unexplained. The paper therefore restricts SmolVLA's absolute numbers to within-paper comparisons, and the scorer debugging its own plan called for was not done. That is disclosed, and it still means half the headline model's numbers cannot be compared to anything published.

Two of the five gaps are borderline by the paper's own rule. The 95% intervals on the paired-try gaps are 7.1–13.3 points for SmolVLA's 10.2 and 6.6–10.9 for π0.5's 8.7 — **both contain 10**, the pre-registered threshold. And the two largest effects (45.1 and 75.8 points) are measured on **undertrained models**: at those learning rates even the fp32 runs end 17–20 points below their own default-rate scores. So the honest summary is that the effect at default learning rates is 8–14 points with intervals straddling the threshold, and the spectacular numbers come from a regime where nothing has converged.

Confound not ablated. bf16 moments travel with bf16 weights and the two are not separated, though a simulation puts the moment effect at under 0.2% of AdamW's mean step at $\beta_2 = 0.95$. More interestingly, bf16 storage acts like a **lower effective learning rate**, so "bf16 at a higher learning rate versus fp32 at its best rate" is untested for VLAs — and a cited 187M language-model result runs the other way at 2×10^{-4}. Scope: two models, one LIBERO suite, shortened schedules, SmolVLA mostly on open-loop chunks. Also: the chunking interaction is real and awkward — re-planning every 10 actions shrinks the default-rate gap from 10.2 to 4.2 points, below threshold, so **how much this bug costs you depends on your replanning frequency**.

And one irony worth recording: for a paper about reproducibility, the artifacts are "available from the author on request" and **the trained checkpoints were not kept**.

Steal list

1. **Compute E(η) from your weights before training, then run the probe during it.** E needs only the weight tensors and minutes of CPU. The probe is a dozen lines: copy trainable weights before every *n*th optimizer step, count elements whose stored value is unchanged after. Both are cheap, general, and would have caught this in any of the 8,196 Hub repos.

2. **Pre-register the effect size and the family of tests in a dated log, then publish the predictions you got wrong.** Two of this paper's recorded predictions were wrong and saying so is what makes the rest credible. Costs nothing.
3. **Pair your evaluation on shared initial states and report McNemar, not two means.** All settings face the same 500 initial states and environment seeds here, which is why 370-against-217 is informative where 69.5%-against-59.3% alone is not.
4. **Check whether your config flag is actually honoured at runtime, not just accepted.** `{p.dtype for p in model.parameters() if p.requires_grad}` after the optimizer is built. Code reading marked SmolVLA safe and was wrong.
5. **Treat GPU model as a run-to-run variance source.** Retraining one run on an RTX 5060 instead of a 4090 moved it 52.2% → 48.2%, about as much as a seed. Nobody reports this.

Cost to reproduce a scaled-down version

The paper's own total is **about USD 110 of rented compute**: roughly 19 RTX 4090-hours for SmolVLA's runs, 8.4 H100-hours and 13 RTX PRO 6000-hours for π0.5, plus about 6 more 4090-hours for the re-planning retrains and another ~10 hours across scorer checks and failed pods. Local development on one RTX 5060 with 8 GB.

For you, a scaled-down confirmation on one arm and one GPU is **two runs and one afternoon**: fine-tune SmolVLA on your own 50–100 episodes twice at LeRobot's default 10^{-4}, once as-is and once with the one-line change (pass `dtype=torch.float32` where `smolvlm_with_expert.py` builds the expert, plus `--accelerator.mixed_precision=bf16`), with the probe running in both. 10,000 steps at batch 16 is a few hours on a 4090-class card; the fp32 copy costs **under 1 GB** for SmolVLA (0.8 GB more peak at batch 16). You do not even need the policy evaluation to learn something — the probe alone tells you what fraction of your updates have been landing. For π0.5 the fp32 weights and AdamW states fit on one 80 GB GPU at batch 32 with gradient checkpointing **only if AdamW updates one tensor at a time**, which LeRobot's optimizer config does not currently expose. One caution the paper is explicit about: the copy makes the effective step larger, so a learning rate tuned under bf16 storage may need lowering, and two single-seed RL reports saw collapse at the unchanged rate.

---

7 Relevant to what you're building

`CONTEXT.md`'s `## Active build` section is still an unfilled placeholder — **eighth week running**, and the twenty-first consecutive harvester sweep to log it. Per the standing rule I am not manufacturing workstreams. Section skipped.

What it cost this week, stated specifically because the cost is now larger than it has ever been: **six items scored 5 and not one of them could be scored on the "load-bearing for a named subsystem" clause**, because there is no named subsystem. They were scored 5 on the weaker "changes how a practitioner builds" clause instead. `CONTEXT.md`'s own example block names "cheap tactile sensing with released hardware specs" and "anything making force a first-class input to a learned policy" as 5-grade interests; this week delivered a USD 399 fully-open sensorized data-collection gripper with a controlled label-noise-to-success measurement, a released force-only contact-perception model with code and dataset, and a numerical bug in the exact one-GPU full-finetune path that `CONTEXT.md` names. A filled `Active build` would have let §7 say which of those to do first. It cannot, so it does not.

---

8 Market signal

No new published compensation band this week, for the second week running, and the reason is structural rather than seasonal. The brief's standing board check — a new senior manipulation or VLA role with stated comp — has no working data source. Figure's Greenhouse board and OpenAI's careers search publish neither posting dates nor pay ranges; Eka Robotics' a16z-hosted board showed a posting date of Mar 2025 for a company that unstealthed on 10-20. The only published band in this

ledger remains W42's Skild AI Deep Learning Manipulation Engineer at USD 100,000–300,000 base, where a 3× spread on one requisition is a level ladder printed as a single posting.

The one quantified alternative I found is **not usable as a datapoint and is worth naming so it does not get recycled**: a recruiting firm's analysis of the robot-learning talent pool (trade/vendor, published 2026-08-10, so out of window) claims postings up 340% since 2024, a USD 190k median base, a USD 150k–233k interquartile band from 10,348 pay-stating postings, and that only 41 US profiles title themselves "Robot Learning Engineer." Every one of those figures comes from the firm's own index with no stated methodology, the 340% has no source at all, the headline ratio compares a title search against a self-defined skill proxy, and its geographic claim ("Pittsburgh has the densest per-capita concentration") rests on three people. It is a marketing piece. Its only independently checkable content is its reading list, which is itself a signal: π0.5, OpenVLA, MolmoAct, CogACT, SmolVLA, GR00T N1.x, and contributions to `openpi` or `Isaac-GR00T` named as the differentiator.

Where hiring is visibly scaling. Demonstration-efficiency companies: Bleu (EUR 5M, two ex-Inria learning researchers, onboard training) and Machenta (USD 40M, one-shot imitation) — USD 45M in one week for needing fewer episodes. Data-collection hardware as a product category: Luxonis shipping a five-camera vision computer at USD 849, the glove-and-handheld-gripper vendors from W43 continuing. And the cobot installed base, which A3 puts up 56% year-over-year at 16.7% of all North American units — that is where ISO-flange end-effector work lives.

Table stakes as of this week, and this is a real shift. Flow matching, action chunking and world-action vocabulary were already assumed. Add two things. First, **inference-cost arithmetic**: VLA-HWBench profiles 18 open VLAs across two edge boards and finds action-generation strategy dominates model size; TRIAGE cuts Jetson prefill 48%; Gaussian Distillation buys 1.6–1.9× from a 115-second readout fit on a frozen public checkpoint; and the π0.5 paired study says 20 denoise steps to 1 is free. Being able to say where the milliseconds go is now expected. Second, and more sharply: **numerical and protocol literacy**. After this week a candidate who cannot say what `--policy.dtype=bfloat16` does to AdamW, or why max-over-checkpoints selection inflates a curation result, is visibly behind. Both were obscure a fortnight ago.

Still differentiating. (1) Force and contact as a policy *reference* rather than an input — four papers this week put the measured wrench in the control or retargeting target and only one released code. (2) Evaluation design as a skill: pre-registered effect sizes (2610.31859), paired tests on shared initial states, published noise floors (DexSuite's 7 pp at 50 episodes), named checkpoint-selection bias, published failed predictions. (3) Cost arithmetic as a *result*: SONIC-K1's 135 GPU-hours against 128 GPUs × 3 days, OpenJaws' USD 398.83 against UMI's 5,389, the deep read's USD 110 total. Each of those is a sentence that wins an interview.

Structural read. The model layer is being priced as infrastructure — Physical Intelligence reportedly raising USD 1B — while the public-market valuation of embodiment came in 25% under its own guidance at Agility. The runtime layer is NVIDIA's by default and is shipping three correctness fixes a week. LeRobot has not cut a release in seventeen weeks and its `main` is the most active it has been, which means the credential that matters is a merged PR, not a version number.

9 Threads

`THREADS.md` is updated. **One merge, one promotion from the queue, 12 open and still at cap.**

Where the episode starts is **merged into** *Online correction as the last mile*. The two have converged completely: Training in Deployment, SimPT, Robotic Data Inoculation and Mulligan now appear in both, and both reduce to one question — where do corrective episodes come from. The merge is a consolidation, not a kill; everything in the absorbed thread is carried in full.

The freed slot goes to *Human video as the data substrate, and where it saturates*, promoted on the basis W43 pre-committed to: "if a fourth week does the same, consider recording the non-replication itself as the finding and promoting the thread on that basis instead." This is the fourth week. Volume rose again (TouchBridge, CARD, EgoContact, Ego-CEVR, EgoForesight, EXE, ActiveAlign, Scaling Dexterous Manipulation via Action Representation Pretraining at 10.1% → 44.7% from a frozen human-pretrained action encoder) and **TRI's 30,000-hour saturation finding is still unreplicated after five weeks**. The promotion is explicitly on the non-replication, and the thread now carries its first piece of counter-evidence: EgoContact finds that on 22,523

manually verified contact events from 8,626 Ego4D clips, **no method beats the naive object-bounding-box centre** (bbox PCK@1% 7.36 against the best model's 11.05; PCK@5% 73.78 against 71.82). If human video cannot yield contact points better than a bounding-box centroid, the substrate's resolution limit is a measurable thing rather than a worry.

THREAD	STATUS	WHAT WOULD COUNT AS THE NEXT REAL UPDATE
Force as a first-class action	**advancing — first release with code**	2610.39314 reproduced by a third party; it is the first force-pretraining result here with code *and* dataset, so this is now a one-week experiment. Carried unchanged: ForceBench's matrix rerun with training seeds; CW-WAM's admittance loop outside Ant Group; the still-unrun falsifier (wrench-action policy on position-teleop data matching one on force-feedback data).
Contact sensing without dedicated tactile hardware	**advancing, bar still unmet at ten weeks**	2610.39314 (force history alone, code released) and FEEL advance it, and 2610.38513 runs a real appliance on force as its sole exteroceptive input. None meets the bar, unchanged since W34: a predicted- or estimated-force policy matching a real-sensor baseline on a task whose labels were **not** collected with that sensor. 2610.39314 uses the F/T sensor at inference, so it does not count.
Low-cost open hardware as training substrate	**advancing — strongest artifact yet, again**	OpenJaws beats OpenTac on this thread's own terms: itemized USD 398.83 BOM with retail links, no custom PCB, commercially permissive, hardware + software + data + sim assets, and the first controlled measurement of label noise against policy success (10 mm → 30 mm ATE, 90% → 30%). Next: a group that did not build it reporting success rates on it. SwapBot's USD 10 tactile tool adapters landing. VecTile's acceptance-gated release, now five weeks pending.
Claimed open releases vs actual artifacts	**narrowing, best week on record**	Delivered and verified: OpenJaws, 2610.39314 (code + data), GelSlim 5.0 dataset (**W43 flag closed**), SO-101 Bench (twin, two datasets, 11 pinned checkpoints), SONIC-K1 (Apache-2.0 code, CC-BY-NC weights), UMI-Live, GAWM, CuraEval (MIT), UED (MIT), LEAP-VLA (MIT), omx-act-models (Apache-2.0), Skill-UMI (CC-BY-4.0). DexSuite shipped 4,700 episodes with **no licence**. Pledged: SwapBot, Gaussian Distillation ("soon"), VecTile. Nothing: Robotic Data Inoculation, Cross-Controller Transfer, ScheduleGen (NVIDIA all rights reserved), the π0.5 paired study, 2610.38513. Next: whether DexSuite names a licence, and whether the deep read's author-on-request artifacts become a repository.
World-action models as unified policy + simulator	**advancing — second independent bound**	2610.34593 finds future-video gains do not survive distribution shift, matching AW-0's own OOD regression from a different group. GAWM is the first sub-100M WAM with code, so the thread is finally trainable at one-GPU scale. Next: 2610.34593 reproduced, or AW-0's regression replicated on a different backbone. **RoboHarness (2610.14018) is still unread, now eighteen days, and remains the largest single gap in this ledger.**
Planning on a learned model	**advancing — first real-hardware evidence**	SO-101 Bench's agent result is the closest this thread has come to its unmet bar: a code-as-policy harness plus a learned grasping policy beats either alone **on the physical arm** (64% against 53% over 45 episodes), with a harness ablation isolating the cause (35% direct tool calls → 67% harness → 85% with hand-off). It is not latent planning, so the bar is narrowed rather than met. Next: a diffusion-steering or latent-planning result on real hardware. LaLa-World's LLM-actor-versus-CEM comparison rerun on an arm.

THREAD	STATUS	WHAT WOULD COUNT AS THE NEXT REAL UPDATE
When to replan, and when to think	**advancing — first convergence in the cluster**	Two independent results now agree that the flow-matching denoise budget is nearly free (2610.37917: 20 → 1 step, +0.01 ± 0.02; 2610.32102: coarse integration helps long open-loop, hurts frequent replanning) and that chunk horizon is where the cost is (26 → 13 costs 11 points of success). The deep read adds that replanning frequency **interacts with a numerical bug** — the fp32 gap halves from 10.2 to 4.2 points when SmolVLA replans every 10 actions instead of running 50 open loop. Next: the matched comparison on one backbone and one benchmark, still unrun, now partly pre-empted by 2610.37917.
Paying at training time, not run time	**advancing**	Gaussian Distillation fits a Gaussian readout in 115 s on a frozen public VLA for 1.6–1.9× at batch 1 — and publishes that full covariance does not beat diagonal (57.28 against 57.64), that quadrupling fitting data to 24,000 contexts does not consistently help, that behavioural diversity is **not** preserved, and that its intervals do not establish practical equivalence. 2610.34453 lifts bimanual dexterous success 10.1% → 44.7% with a frozen human-pretrained action encoder. GAWM freezes DINOv3. 2610.39314's frozen sentence embeddings work while its frozen *readout* fails, which is the sharpest statement yet of where the frozen/trained boundary belongs. Next: the matched one-step-action-head comparison, now eight weeks unrun.
Benchmark validity in physical AI	**advancing — the week's dominant thread**	Six results: SO-101 Bench (ρ=1.00 but CI [0.000, 0.169] on five policies, sim optimistic by 9–31 pp); 2610.37917 (8 of 12 ablations null, a prior rendering gain fails to replicate, two baseline runs per cell); CuraEval (a quality score is episode length; selection flips +10.0 → −1.0 on the same seeds with the protocol); 2610.34151 (a log line flips a VLM judge); EgoContact (nothing beats a bbox centroid); 2610.40364 (the controls do most of the work). Plus Isaac Lab #8634 making pre/post results incomparable and LeRobot #5107 corrupting normalization std. **Bar still unmet since W38: seed bands on REAL hardware.** SO-101 Bench gives Wilson intervals on hardware but one seed; the deep read gives three seeds in sim.
Online correction as the last mile	**advancing — absorbs *Where the episode starts***	Robotic Data Inoculation (Berkeley, Levine) is the strongest entry: a critic-trained residual perturbs the human's teleop **live**, so counter-corrections become paired recovery data. Its honesty is the contribution — early-session human success is **below both baselines until 25% in**, a random-noise baseline **actively hurt** the keycap task (so DART-style noise is not free), it only pays where recovery matters, and the residual-magnitude decay plot is the convincing evidence. Nothing released. Also 2610.37547, 2610.37533, 2610.36852, ScheduleGen and a learned speed prior for retiming demos (2610.32676). Carried from the merged thread: Mulligan's correction/counterfactual distinction, Dexterity from Supervised Play, Hardware-as-Regularization. Next: a third-party rerun of EXPO-FT (**eight weeks**, nobody); Mulligan's sampler reproduced by someone who did not write it.
Who owns the open robot-learning stack	**advancing — the shape sharpened**	LeRobot v0.6.1 is the **eleventh consecutive negative** at seventeen weeks post-deal, while `main` absorbed Long-WAM in-tree, thirteen commits in a day and four behavioural default changes. W43's inference is confirmed: watch `src/lerobot/policies`, not `/releases`. The model layer is being priced as infrastructure (PI reportedly at USD 1B) while the embodiment layer priced **down** (Agility's SPAC at USD 465M against its own USD 620M estimate). REP 0157 gets a dedicated PMC meeting **2026-11-04**. `openrobotics/reps` has zero commits and the archived `ros-infrastructure/rep` would have read as silence forever. Next: REP 0157's outcome; whether any platform-holder commitment is written into a foundation or steering body (still none); whether LeRobot cuts v0.7.0 or abandons tagged releases.

THREAD	STATUS	WHAT WOULD COUNT AS THE NEXT REAL UPDATE
Human video as the data substrate, and where it saturates	**PROMOTED from the queue — on the non-replication**	Promoted per the W41–W43 pre-commitment after a fourth week of rising volume and flat evidence. The finding **is** the non-replication: TRI's 30,000-hour saturation result (2610.12464) has gone five weeks without a second corpus. EgoContact adds the first measured ceiling — no method recovers contact points better than an object bbox centre. Against that, 2610.34453's 10.1% → 44.7% is the strongest positive. Next: a second group reproducing the saturation finding on a different corpus; or EgoContact's bbox-centre baseline holding up when someone tries to beat it. Kill date W48 if neither lands.

Closed and still closed:
Simulation fidelity for transmission and contact (killed at W43 as pre-committed in W40). Reopen only if a tendon-hand or underactuated-finger transfer result credits a **named** contact model. Nothing this week does.

Nothing is queued behind the cap
For the first time since W41.

10 Open question

How much of the last year's measured VLA progress is architectural, and how much is a numerical artifact of the baselines?

The deep read makes this unavoidable rather than rhetorical. If 9 of 12 open codebases ship an affected default, if 99.3% of sampled SmolVLA fine-tunes and 84% of sampled π0.5 fine-tunes on the Hub store trainable weights in bf16 with no master copy, and if the fp32 fix is worth 8–14 points at default learning rates and far more below them, then a large share of published "our method beats SmolVLA / π0.5" comparisons were run against a baseline that was discarding most of its updates. Nobody knows how large. The paper does not claim to.

What makes it genuinely open rather than merely alarming is that the sign is not obvious. Three readings are all consistent with what is known. **One:** most papers fine-tune baselines at default learning rates, where the measured gap is 8–14 points with intervals straddling 10, so published margins shrink but survive. **Two:** bf16 storage behaves like a reduced effective learning rate, so any paper that tuned its own method's learning rate while leaving the baseline on defaults got a free advantage that scales with how far below default it went — and the deep read's own 45- and 76-point gaps at ten-times-lower rates show how large that can get. **Three:** since the loss curve gives no warning (final losses were 0.30 and 0.29 for the bf16 settings against 0.26, with success differing by 10–14 points), reviewers had no available signal, so the error is unbiased with respect to who published it and the field's *rankings* may be roughly intact even if its *numbers* are not. Reading two is the one that should worry you, and it is the one with the least evidence.

This is not a one-paper worry, because the same shape appeared twice more this week. CuraEval shows a curation result flipping from +10.0 to −1.0 points **on identical seeds** purely by changing checkpoint selection and rollout count — and names two published methods using the inflating protocol. The π0.5 paired study fails to replicate a prior high-resolution-rendering gain. Three independent findings in one window where a published improvement is plausibly an artifact of the measurement apparatus rather than the method.

What would settle it:

Take three to five published VLA comparisons whose baseline was SmolVLA or π0.5 fine-tuned with LeRobot or openpi's PyTorch trainer at a stated learning rate, and rerun **only the baseline arm**, twice — once as the paper ran it, once with the fp32 copy — on the paper's own benchmark, with the probe recording erased-update fractions in both. Nothing else changes: same data, same steps, same seeds, same evaluation protocol. Report the baseline's two numbers against the paper's claimed margin. If the margins survive, the field got lucky and we can stop worrying. If a third of them invert or vanish, the correct conclusion is that a year of VLA leaderboard deltas is partly a precision artifact, and the remedy — report your optimizer's parameter dtypes and E(η) alongside your results — costs a line in a table.

Cost at one-arm, one-GPU scale: SmolVLA on LIBERO at 10,000 steps and batch 16 is a few hours on a 4090-class card, so **six runs across three papers is under 40 GPU-hours and roughly USD 30 of rented compute**, with no robot time at all. The constraint is not compute; it is that the baseline arms have to be rerun exactly as published, and most papers do not state their optimizer parameter dtypes, which is the whole problem restated.

Cheaper version available to anyone today, and worth doing before the above: run the probe on your *own* most recent fine-tune. If a median 22% of your weight elements are bit-identical to the base model, you already know.

ALSO THIS WEEK (4+ IN QUEUE, NOT COVERED ABOVE)

Nothing scoring 4 vanishes silently. All preprints unless noted; abstract-level unless the queue recorded a full read.

- DynaSlots — action-free, label-free slot-centric 3D dynamics pretraining; per-slot rather than per-scene dynamics, with inverse plus diffusion forward models. Gains over scene-level baselines on Meta-World, Adroit and a real RealMan arm. Code and project page claimed at an **anonymous** github.io URL, so the release is not yet verifiable — check whether it de-anonymizes.
- KiRoSe — Berkeley Automation. Discrete Sparse Motion Tokens from **kinematics only, no vision**, read by an open LLM, segmenting demonstrations zero-shot with transfer to new embodiments and to human motion. Runs on logs you already have. Project site; code likely on past form, not confirmed.
- ThinkWorld (NVIDIA + KAIST, 42 HF upvotes) — Chain-of-Dynamics motion tokens reinforced when the predicted future matches the observed transition, turning unlabeled video into supervision. No code.
- PDRVLA (Peking + JD, 80 HF upvotes) — a lightweight physical-dynamics module predicting future RGB, depth and flow *features* rather than generating video online. The right trade for latency; no code.
- TAPESTRY — actions as single-patch tokens inside a pretrained video diffusion transformer, generated in one DDIM step. Project page only.
- Hy-Nomas (Tencent Hunyuan, 108 HF upvotes) — 8B video world model with action-aware causal attention scoping instructions to windows; **weights, code and benchmark released**. No action conditioning, so it is a world model rather than a policy.
- From 12 to 200 Demonstrations — ACT, Diffusion Policy, SmolVLA and π0.5 scaled across 12–200 demonstrations on one low-cost arm with a four-stage partial-credit ladder. Exactly the shape of study a one-arm builder wants; no code.
- VLA-HWBench (NUS) — 18 open VLAs, two edge boards, seven-stage latency breakdown; action-generation strategy dominates model size. TRIAGE (SNU + CMU) — training-free visual-token reuse plus pruning under a fixed recompute budget, 48% Jetson prefill cut. Neither released.
- PhysLadder — matched sim-real benchmark, 16 contact-rich tasks at three interaction levels; frontier-model zero-shot collapses as interaction depth rises. Authors' own benchmark; follows W43's Astranomy cost finding.
- Pushing Dexterity and Safety for In-Hand Object Pose Reaching (Berkeley + UC Davis) — sim-only in-hand pose reaching with a learned residual safety filter at adjustable conservativeness; **full recipe released**.
- BURGER — blank-canvas drawing as a benchmark that removes the visual shortcut and isolates data scaling and chunking. Finds flow matching scales where π0-FAST does not. Code claimed.

- ScheduleGen (NVIDIA + UT Austin) — per-arm waypoint timings as a MILP, so multi-arm coordination comes from inserted waits. The load-bearing negative: **DexMimicGen generates zero valid trajectories on cross-arm variants** where ScheduleGen gets 64–88%, and 160 real YAM trials give 50–80% with no real fine-tuning. Cross-arm variants are author-defined and built to break independent-arm transfer; one seed; the YAM policies need 8 GPUs, though the robosuite variants are single-GPU. No code, NVIDIA all rights reserved.
- X-EraAI/e05-robotwin-data — release. E0.5 data and weights out: 2,500 episodes, 549,787 frames, LeRobot v2.1, geometry cache, depth in metres with intrinsics and extrinsics. Sim-derived, **no held-out split**.
- Robotic Data Inoculation, DexSuite, Gaussian Distillation, CuraEval, SONIC-K1, SO-101 Bench, OpenJaws, SwapBot, TouchBridge, UMI-Live, 2610.39314, 2610.38513, EgoContact, Cross-Controller Transfer, 2610.40364, 2610.37917, FLoRe, CARD, 2610.34593, 2610.34453, 2610.34151, GAWM, 2610.32102, 2610.32277, 2610.32676, LEAP-VLA, omx-act-models, Skill-UMI, GelSlim 5.0 data, LeRobot #5107, the Long-WAM commit and the Isaac Lab and MuJoCo fixes are all covered in the sections above.

Sources are linked inline. Source types: arXiv entries are preprints unless stated — none of this week's items is peer-reviewed, and CuraEval is a CoRL 2026 workshop submission; Hugging Face cards and GitHub release tags are releases; GitHub commits and pull requests are primary artifacts; humanoidsdaily, humanoid.guide, theaiinsider, therobotreport, technode and TechCrunch are trade press; agilityrobotics.com and bleu-robotics.com are company primary sources; Business Wire is a press release; AGIBOT's 110-humanoid record, DeepMind's Genie 4 Apollo run and XPENG's battery claim are demo or announcement only. Single-lab claims with no released artifact are flagged unreproduced where they appear. Where sources conflict — Physical Intelligence's valuation — both figures are given and neither is picked.

Rate this issue 1-5 and tell me what was noise.