# This Week in Robotics

THE LINE

**Shear-capable tactile skin became a parts list this week, and the parts list is cheap.** Three tactile-hardware papers landed inside one 24-hour announcement cycle, two of them with complete fabrication paths and real closed-loop policy deltas, and both reach the same architectural trick from opposite directions: put a passive structured bump over a small group of cheap pressure-sensitive cells and read shear out of the *differential* across the group. OpenTac (preprint, JHU + MIT) compiles robot surface geometry straight to flex-PCB Gerbers and ships code under MIT, hardware under CERN-OHL-P-2.0 and docs under CC BY 4.0, verified on the project page. VecTile (preprint, Imperial) publishes a full bill of materials at **USD 34.67 per unit, of which USD 25 is an Arduino Nano** — USD 3.57 for the sensing layer itself — and fabricates in-house with a 3D printer and a sub-USD-150 vacuum chamber.

And the money arrived in the same fortnight. Eka Robotics took USD 150M Series A (trade press; a16z and F-Prime leading, USD 1.45B valuation attributed to TechCrunch) on the thesis that contact forces are a policy input rather than a safety check, and Enact Intelligence emerged from stealth in Zurich with USD 30M *pre-seed* for whole-body force sensing read through existing actuators with no added hardware. Those two bets are mutually exclusive in an interesting way — one adds sensors, one refuses to — and this week produced papers on both sides.

The second story is less fun. **The simulator layer broke underneath you twice in nine days.** MuJoCo 3.17.0 (2026-10-22) is the second model-affecting release since 3.16.0 on 10-14, and between them they change `mjNEQDATA` from 11 to 12, make 1D-flex `young` actually honoured where it was silently ignored, deprecate `meaninertia` so solver termination now scales by the inertia-matrix trace, and change actuator shortcut gain inheritance so a `motor` inside a `position` default class is **no longer a position servo**. Isaac Lab v3.0.0 GA flipped quaternion order from `wxyz` to `xyzw`. Set against W40's finding that an unpinned MuJoCo version was worth 26 points on one LIBERO task, that is the headline build risk of the week. Pin before you upgrade.

---

Window: 2026-10-19 → 2026-10-25. Queue: 7 part files, ~300 scored lines, 25 at 4, none at 5.

Coverage note, and this week it is good news. The recall problem that has disfigured every issue since W40 is solved: `papers.cool/arxiv/cs.RO?date=YYYY-MM-DD` honours date/show/skip, prints its own N-of-total counters and serves abstracts inline, so for the first time the whole week has denominators — 10-19 = 68, 10-20 = 160, 10-21 = 88, 10-22 = 78, 10-23 = 69, 10-24 = 0, 10-25 = 0. Two consequences worth holding onto. First, **the 88-entry 10-20 gap carried for four sweeps is closed**: 160/160 enumerated, 73 entries recovered. Second, the retrospective recall numbers are humbling — the 10-23 sweep had enumerated 28 of 69 entries in its own band, 41%, and three consecutive sweeps found the previous sweep's band demonstrably incomplete. Every "nothing happened in X this week" in the W40–W42 issues should be read with that in mind.

Still unfixed: cs.LG and cs.CV were never enumerated on any of the seven sweeps, so cross-list work is sampled, not swept. arxiv.org itself is unusable from this environment (stale listings, 429s, robots-disallowed search, blocked export API), so the enumeration is mirror-internal — if papers.cool silently dropped an entry its counter would just read one lower. Two reasons for mild suspicion about the mirror: it served a 2009 paper with no cs.RO tag as entry 69 of the 10-23 cycle, and it prints at least one title truncated mid-phrase. And arxiv.org's own count for the 10-22 cycle (120 entries, 69 new) is **not reconciled** with papers.cool's 78. Carried forward unresolved.

# 1 Policy learning & VLAs

## AW-0 — code as a third denoised modality

Agentic World Action Models (preprint, Berkeley + NVIDIA + UPenn; Goldberg, Darrell, Yuke Zhu, Jim Fan, Danfei Xu) is the week's most architecturally interesting policy paper. The mechanism: a mixture-of-transformers denoises video, action *and code* jointly, with **discrete masked diffusion for the code expert alongside continuous flow matching for video and action**. The generated programs are not scratch work — they are written into a 50-skill library (23 from the YAM teleop corpus, 27 from egocentric human video) and reused as agentic skills, so the same forward pass that produces a chunk also produces a callable.

Eval: 3K hours of pretraining (2,000 YAM + 1,000 egocentric human), RoboTwin2.0 clean 80.8 against WLA 64.8 and π0.5 54.9, real 8-task 71.1 against DreamZero 46.0 and π0.5 43.6. Unusually rigorous for this literature: **three training seeds × 16 rollouts per task**. Joint denoising runs 0.23 s against 0.55 s autoregressive on one RTX 5090.

The load-bearing part is the negative the authors report themselves: under randomization, hybrid code-as-policies execution reaches 20.6, which is **below plain π0.5 at 27.7**. The world-action model loses to a vanilla VLA out of distribution. Both the code-without-comments and code-without-skills ablations hurt on all eight tasks, so the mechanism is real in-distribution. Read it next to How Representational Priors Shape Generalization in World Action Models (preprint, SNU, abstract-level), which argues WAM generalization divergence comes from the generative backbone's representational geometry — that is a candidate explanation for AW-0's own OOD regression. No code, no weights. The annotation pipeline depends on frontier models (Claude Opus 4.8 + Gemini 3.1 Pro), which is a reproduction cost nobody is pricing. Not reproducible at one-arm scale; the *idea* is.

## LaLa-World — verbalise the actions and one predictor covers every domain

LaLa-World (preprint, NAVER AI Lab + Georgia Tech + Brown + Mila) is the one to actually clone. Numerical actions are written out in natural language, which makes language the shared action interface, which means one JEPA predictor trains across domains *and* over raw video with no action labels at all. Qwen3.5-0.8B predictor on a **frozen** DINOv2-S/14 or DINOv3-L/16. **Code released** for both LaLa-World and LaLa-Policy.

The numbers that matter are cost, not score: 6× less training cost and 13.9× faster planning than RAEv2 under matched encoder, data and conditions; LaLa-Policy competitive with Cosmos Policy at 5.5× faster inference; planning 47.5 (JEPA world models) → 59.4; adding image and video data 59.3 → 61.6 while killing a 6.8-point paraphrase penalty. And an LLM actor replaces CEM at **90 rollouts against 9,000** (Claude Opus 5: 100% on Wall, 53.1% on RoboCasa). Soft spots: no proprioception by design, substep summing discards within-step variation, the LLM actor is *worse* than CEM on PointMaze (31.2 vs 56.2), and the DROID subset is only 20% of the data. One GPU, genuinely.

## ALPS — a frozen 85M encoder that carries physical state

ALPS (preprint, UCSC-VLAA) trains one ViT-B/16 token set to carry physical state, language alignment and pixel reconstruction simultaneously. **Weights Apache-2.0** (786 MB), code on GitHub, SHA256SUMS published, and eight LIBERO policies released at 50 demos per task. 85.6M parameters. Lowest LIBERO state-probe error of any ViT-B encoder, matching V-JEPA 2-L at 3.5× the size; Flickr30k I2T R@1 93.2; reconstruction L1 0.120, beating every encoder not trained to reconstruct. The comparison is the valuable artifact: fifteen frozen encoders under one protocol. Caveats: still images and text only, no video, no action supervision, and float16 overflows in deeper blocks — use bf16. This is the cheapest frontend swap on the list.

## Frontier models as robot policies is now a nine-paper cluster, and the answer is still "no"

Opened as a cluster on 10-19, it closed the week with GPT-6-Astra embodied intelligence (twelve benchmarks, eight real configs, code released: competence crossed, practicality not — in-context demo cost and latency bind), HexaShell (a Codex agent writes its own triangulation tools; RoboDojo SOTA 39.7%; the limit is contact, not reasoning), GPT-GEAR (RoboDojo 26.92, LIBERO-GeoMotion 11.1 zero-shot and 50.0 with demos, real ROKAE 40–80% over five trials, package storage 0%), the NVIDIA SRL

Astranomy study (blog; 74.8% on RoboLab-120 at **30.9M tokens and USD 36.40 per episode**), and Astra is a native visuo-tactile policy (19 sim tasks, touch helped on six). Taken together: semantic competence is not the bottleneck, contact and latency are, and the cost per episode is three orders of magnitude off anything deployable. That is a clean, defensible interview answer and it did not exist a month ago.

One more worth a line because it is a negative with a number on it: **To Route or to Adapt?** (preprint, abstract-level) tested eight expert-routing schemes for diffusion planners and **none beat the static default**; an oracle per-scene selector buys only 0.9 m, which caps the whole approach, while reward fine-tuning on 20 target scenes cuts mean lateral offset 12.4 m → 6.9 m. Driving domain, but the design lesson transfers: do not build the router. Its own side finding is the more useful one — open-loop offsets improved while closed-loop performance degraded.

---

## 2 Manipulation, dexterity & sensing

This was the week's centre of gravity. §6 is the full read of OpenTac; here is everything around it.

**VecTile** (preprint, Imperial College London, Manipulation and Touch Lab) is the cheapest credible shear skin I have seen and the only one with a complete, named-part BOM: USD 34.67 per unit (USD 14.67 with an ATmega328P clone in place of the Arduino), USD 3.57 for the sensing layer, 6 mm thick, 4×4 array at 10 mm pitch, each taxel reading a 3-DoF force vector, 343 Hz raw frames and 25 ms measured end-to-end lag against an ATI Mini40. Materials and settings are all specified down to the Mooney-Rivlin constants (Dragon Skin 30, C10 0.198, C01 0.049) and the print profile. Accuracy is **held out properly** — 0.39 N normal and 0.16 N shear RMSE under a four-fold direction-rotation split, not in-sample — and the flat-fitted decoder transfers to R=85 mm convex and concave surfaces for 0.05–0.06 N extra RMSE. Policy result: diffusion policy on UK 3-pin plug insertion, 84 demos, 25 rollouts at unseen socket locations, 48% vision-only → 64% normal-only → **88% full 3-DoF**, and the gain is in *corrections* (35 → 61 → 86%) rather than first-attempt alignment. Plus a scalpel-balloon reflex task where 20 of 20 trials stopped before puncture and normal force alone never tripped the threshold. The catch: **designs and code release only on paper acceptance**, so nothing is downloadable today. Hysteresis, drift and durability explicitly not characterised; no seeds on the policy; sensitivity dies above roughly 5 N per taxel.

**GelSlim 5.0** (preprint, Michigan, abstract-level) is the third tactile-hardware paper of the same cycle: one fingertip camera acquiring both the external view and local contact geometry, attacking occlusion. Flagged for a full read next week — but note the harvester's standing flag that no 5.0 repo, paper page or cost has surfaced beyond 3.0/4.0, so treat the version number cautiously.

**T-WLA** (preprint, MIT + Harvard + UC Davis + Berkeley, full-text read by the harvester) is the architecturally cleanest answer to the rate mismatch between vision and touch: cache the visual memory so a **119 Hz tactile editor rewrites unexecuted actions without rerunning vision** (vision 39.7 Hz, touch 119.1 Hz, against 69 Hz for the best baseline). Future-image and future-pressure heads are training-only and removed at inference — the fourth instance of the "pay at training time" pattern this ledger has tracked. Real mean success 17.0 → 50.0 against SmolVLA at **30 demos per task**, which is a one-arm budget. The load-bearing caveat is in the ablation: **tactile itself buys only 47.25 → 50.0; the 17 → 47 jump is the visual branch.** Cup nesting stays at 17% either way. Four tasks, no seeds on real, and the world-prediction ablation deltas (3.960 vs 3.963) are noise. Code claimed on the project site, unverified.

**CW-WAM** (preprint, Ant Group, full-text read) is the strongest pure force-in-the-loop result of the week and the most statistically honest paper in the queue. Paired motion-wrench tokens with pair-causal attention, a 10 s / 150-interval wrench memory, and — this is the mechanism — **the predicted wrench is the admittance reference**, a slow-fast split with 20 Hz admittance over 1 kHz impedance. 1 mm-clearance peg insertion 25 → 70%, adapter 40 → 95%, force RMSE 4.8 → 2.7 N on the peg. It **reports bootstrap confidence intervals over 10,000 stratified resamples**, ablates the long-term memory and pair-causal attention separately, and publishes an honest negative: adapter insertion force RMSE *worsens* against WAM+Wrench because the policy learns a search behaviour. In-distribution only, hand-tuned gains, predefined compliant directions, 8×H20 for 50k steps, seed 42, project page but no code.

**ContactBench 1.0** (preprint, abstract-verified) deserves more attention than its score suggests: 3D-printable objects with **tunable resistance that look identical**, which is exactly the right way to separate contact reasoning from visual shortcuts. Its headline is a load-bearing negative — visual-tactile policies gave no real gain on it. Pair with **ContactKernels** (preprint), which shapes MPC sampling with a wrench-estimated contact point and normal and takes real insertion 53.3 → 90.0% **with no learning at all**, and with **Tactile In-Hand Object Rotation** (preprint, Bristol, LEAP Hand + TacTip), where a **symmetric finger layout beats the anthropomorphic one by 32% rotation at 19.9% lower fingertip load** — a morphology ablation on real hardware, which is rare, and a direct argument against copying the human hand.

On the no-extra-hardware side: **ContactFlowFilter** (preprint, KAIST) recovers contact count, locations and force vectors from proprioception alone via a flow matching plus particle filter mixture, and **Learning-Enhanced Observer for Sensorless Task-Null-Space Compliant Control** does hybrid pose-force control with a wrench observer and no sensor. Together with LocoPilot (§3) and Enact's funding, sensorless force is now a funded thesis rather than a paper cluster.

Data and rigs: **SharpaDex v1.0** (release, CC-BY-NC-4.0) is 300 hours of real bimanual dexterous teleop — 28,993 episodes, 59 tasks, two 22-DoF hands, 32.4M frames, 60D tactile force/torque plus deformation and raw tactile video, 224,087 language segments — and its card is unusually honest about the seven torso-torque dimensions and twelve TCP force/torque slots that are **always zero**. No paper, no baseline, no success rates, and the source is the hand vendor. The **Origami S1 skill-capture glove** (trade press + company blog) is the right idea executed in hardware: the glove and the robot hand share both joint layout and tactile sensors, so joint angles map directly to joint targets with **no retargeting solver** and the touch domain gap is removed by construction. The company's own blog concedes that matched joint angles do not guarantee matched contact and that the model for that gap is still being developed — which is the FOMO-Dex and FOMI lesson from W42 restated by a vendor. Waitlist only, no price, nothing open.

---

## 3 Platforms & embodiment

**Demonstrated.** SuperMimic (preprint, Berkeley; Abbeel and Sreenath, abstract-verified) is the week's clearest capability result and it is a compute result: analytic policy gradients backpropagated through a learned differentiable world model take humanoid whole-body control to **3 GPU-hours against a 128-GPU-week baseline**, 25 failures against 179 out of 129,514 motions, zero failures at 9.8 hours on two GPUs, with real Unitree G1 transfer. Caveats that matter: the benchmark clips are *training* motions, flat ground only, not data-matched, no code stated. If this reproduces, a thousandfold GPU-hour reduction on whole-body control is the most one-GPU-relevant platform result of the month.

LocoPilot (preprint, IIT, abstract-verified) is a low-cost bilateral teleop rig that estimates force **with no F/T sensor** (1.27 N free, 2.25 N loaded) and renders it through a passivity-constrained haptic channel; 289 demos, 11D unified action, Unitree Go2 + AgileX Piper-L. But it reports **one rollout as a feasibility check with no success rate and no comparison**, so the rig is the contribution, not the result.

PHD-Hand (preprint, NJIT, full-text read) is the open-hardware near-miss of the week: linkage-driven 3-DOA digits with no tendons and no gears, **13 parts per digit against 27+ for every compared peer**, 200 mm hand length (closest to human of any design compared), budget FFF printer with the full print profile given, generic mini RC servos, a 0.6 mm spring-steel drive rod. It attempts all 33 GRASP taxonomy types and demonstrates screwdriver and crescent-wrench operation that tendon hands structurally cannot do, because the drive is bidirectional and backdrivable. **Nothing is released** — open-sourcing is listed as future work — there is no cost figure anywhere, durability is an anecdotal ~200 h with two drive-rod breaks, and there are no force measurements in this paper.

**Staged, or video-only.** Unitree's H2-X kung fu video and XPENG's IRON gesture demo are both demo video only, no method, no success rate; XPENG's single number (0.2 s visual-reasoning refresh) is company-stated. Unitree separately announced **Rehearsal**, a self-improvement pipeline where an AI reads papers, writes code, tests in sim then on hardware and humans score the output — reported secondhand with **no primary post on unitree.com at all**. Roadmap, not result.

Figure's leaked all-hands (trade press, single-source recording) claims Figure 04 design lock — which was **also claimed in May 2026**, so treat it as a restatement — a fourth hand iteration at human parity with no dexterity benchmark, Helix 3.0 as a screen

teaser, and the 1,000th Figure 03. The one number worth keeping is the admitted one: **Helix 2.5 home trials at 237 of 420, 56%**, which the article itself calls a reliability gap. The 200-hour logistics run was robots swapping, not one robot.

**Purchasable, eventually.** The Agile ONE SN (trade press) is a 1.4 m, 31-DoF research humanoid with **two 20-DoF Wuji tactile hands**, Jetson Thor, shipping preconfigured with GR00T 1.7, whole-body control and Isaac Sim/Lab assets. The preconfiguration is the actual story — it is a build-friction product, not a capability product. Availability early 2027, no price, weight disagrees between sources (43 kg per Agile, 40 kg per NVIDIA), and the article itself warns that skill transfer from SN to the industrial variant still needs demonstrating. Note also that **Wuji is now the default dexterous hand in this feed** — it appears in T-WLA, DONIC and Agile ONE in one week.

The most consequential platform item is a commoditisation story: the Unitree G1 has become the default body for American humanoid startups (trade press, read in full). Lumi ships a 29-DoF G1 with Jetson Thor at USD 39,990 founder edition; Voxel is estimated at USD 20,000 for 2028; Hugo's G1 base is an unconfirmed visual inference from a monitor in a video. One investor calls the category a wrapper. The load-bearing detail for a small build: **the USD 13,500 base G1 does not support secondary development.**

---

## 4 Open source & tooling

Ranked by what reduces friction this month.

1. **MuJoCo 3.17.0** (2026-10-22, Apache-2.0) — new `mjbatch` thread-pool batch-rollout API, which is the useful part, inside the second model-affecting release in nine days. Breaking: `mjNEQDATA` 11 → 12 with welds recomputing `relpose` in `mj_setConst`; `mjs_addTexture` signature change; `mjs_fuseStatic` and `mjs_discardVisual` return values; **1D-flex `young` now honoured where it was previously ignored**, so models that set it now bend or fail to compile; `meaninertia` deprecated and solver termination now scaling by the inertia-matrix trace, so PGS, `noslip` and islands-off runs stop at different iteration counts. 3.16.0 adds the silent one: **actuator shortcuts now inherit gains only from a matching shortcut or a `general` default**, so a `motor` in a `position` default class is no longer a position servo. Keyframes are no longer copied by `replicate`, and flex shear stiffness and quadrature were corrected, so existing flex models behave differently. MJX ships in the same version line.
2. **Isaac Lab v3.0.0 GA** (2026-10-20, BSD-3-Clause) — the Newton/MuJoCo-Warp backend and Kit-less runs, finally non-EA. **Quaternion order changed `wxyz` → `xyzw`**, which is a silent breaker; `ProxyArray` now needs `.torch` or `.warp`; `AppLauncher` deprecated. Pins Isaac Sim 6.2, Python 3.12, PyTorch 2.11, CUDA 13.0, Warp 1.17, Newton 1.6.1, RSL-RL 5.5.1. Method note that cost this ledger three sweeps of contradiction: **the `/releases` index page renders stale and still shows v3.0.0-EA as Latest**; the direct `/releases/tag/<tag>` URL is the truth.
3. **FineART-VLA-subtask-midtrained** (release, 2026-10-23, Gemma licence) — a genuinely new LeRobot policy head from Scale AI and the LeRobot team. It predicts the next subtask in language, then acts on it; two modes from one checkpoint, trained with flow matching plus text cross-entropy plus FAST action-token cross-entropy. π0.5 base (Gemma 2B VLM + Gemma 300M action expert), 4.1B parameters, 9.35 GB, three camera views, 14-D state and action, 50-step chunks at 10 flow steps, 300k training steps. Data is the FineART release: **40,543 episodes, 1,718 hours, 533,913 subtask annotations, 151 tasks** on a stationary ALOHA-style Trossen rig, mixed 30% subtask prediction / 70% subtask-conditioned action. No eval numbers on the card — they are in arXiv 2609.36416. 9.35 GB fits one GPU. **The code is in `lerobot` main, not in a tagged release**, which is the method note that matters: watch `src/lerobot/policies`, not `/releases`.
4. **InternW0 checkpoints** (release, 2026-10-23, **MIT on a world-action checkpoint**, which is the permissive outlier) — slow Wan2.2 VideoDiT planner plus fast ActionDiT executor, 7.35B parameters, 14.7 GB per checkpoint, with a runnable `run_libero_eval.sh` and a disclosed post-training recipe (10 epochs, lr 5e-5, batch 48 per process, seed 42, step 600). LIBERO average 98.6 over 2,000 rollouts, which is a saturated benchmark; the honest number is **RoboTwin 2.0 randomized-unseen at 68.0**. Not self-contained — it needs Wan2.2-TI2V-5B, dinov3-vitb16, a LIBERO checkout, a LeRobot conversion and a UMT5 text cache, is not a `transformers.AutoModel`, and 7.35B plus a 5B video backbone is not comfortable on one consumer GPU. Third-party assets are separately licensed, so MIT does not cover the stack.

5. **lerobot/lingbot_vla_v2_base** (release, updated 2026-10-22, Apache-2.0) — 6,375.9M-parameter VLA weights in the LeRobot org with a `lerobot` library tag, citing arXiv 2607.06403. I checked the harvester's open question: this is **not** a mirror of `robbyant/lingbot-vla-v2-6b`, which has the same parameter count but cites a different paper (2508.02317), carries no licence tag and was last updated 7 July. The signal is downstream volume — at least a dozen third-party `lingbot_vla_v2` finetunes were pushed between 10-22 and 10-25, including SO-101 and RoboTwin variants. 6.4B is not a one-GPU full finetune; LoRA is what the community is actually running.
6. **Genesis v1.4.5** (2026-10-22, Apache-2.0) — faster non-convex collision, 8× SDF (author-reported), AMD RDNA4 crash fixes, and a **persistent kernel cache**, which is the real friction win. Breaking: the custom Genesis differentiable-simulation tensor is removed.
7. **OpenTac's own toolchain** (see §6) — OpenTac-CAD runs hosted at `opentac.github.io/cad` or locally via `uv run opentac cad`; OpenTac-Sim is built on MuJoCo and reads sensors from the manifest, with MuJoCo Menagerie robots supported.
8. **mjfit** (MIT) — procedural MuJoCo assembly generator producing exact convex collision MJCF for pegs, holes, threads and gears from any outline, backing Blind Precise Assembly. Flag resolved: the repo exists, 17 commits, no tagged releases.
9. **batchplan-ros** (2026-10-24) — send many candidate grasp poses in one request, batch-solve IK and plans, return the shortest valid path, as a MoveIt plugin on ROS 2 Jazzy. Rust core, CPU-only with optional wgpu, browser Panda demo. **No licence stated**, no benchmarks, single author, self-described as early. Grasp-candidate batching is a real friction point, so worth watching rather than adopting.

Negative confirmations that are now planning assumptions, not watch items: **LeRobot is still v0.6.1 from 2026-08-03 — the tenth consecutive negative, sixteen weeks post-deal**, with the v0.7.0 tag 404ing and v0.7.0 still an open roadmap issue (#3832). ManiSkill is still v3.0.1 from April (and has moved to `github.com/mani-skill/ManiSkill`). No ROS 2 distribution release and no REP — and that last one must be downgraded from a negative to *unverified*, because the REP surface now has no working route from here at all: `ros.org` is behind bot protection and the GitHub REP repo is robots-disallowed.

---

## 5 Money & moves

Interpretation, not reporting.

**Eka Robotics, USD 150M Series A** (trade press, a16z and F-Prime leading; USD 1.45B valuation attributed to TechCrunch). Pulkit Agrawal (MIT) and Tuomas Haarnoja (ex-DeepMind) raising a nine-figure A on "contact force is a policy input, not a safety check" is the clearest price the market has put on this ledger's oldest thread. Note the tight coupling to the research record: the same Agrawal group published Blind Precise Assembly this week (73% zero-shot peg insertion across eight pegs over 237 trials from **joint encoders alone**), and OpenTac's first author thanks Agrawal as his postdoctoral advisor. No paper, no code, no sensor spec, no named customers; the 1,500-picks-per-hour-versus-600-human figure is company-claimed.

**Enact Intelligence, USD 30M pre-seed** (trade press; Redalpine, Notion, Firstminute, with Yann LeCun, Arthur Mensch and Thomas Wolf as angels; team from 1X, DeepMind, Physical Intelligence). USD 30M *pre-seed* is extraordinary, and the thesis is the exact inverse of Eka's: whole-body force awareness read through existing actuators with no extra hardware and nothing fragile. That is the same bet as ContactFlowFilter and the sensorless wrench-observer papers, now capitalised. **Two funded, mutually exclusive answers to the same question in one week is the most informative market fact of the issue.**

**Dream Labs, USD 90M + the Dream API** (trade press). The claim — control robots from new manufacturers with no data, via video pretraining — is a very strong zero-shot assertion with no method, paper or benchmark attached. The structural read matters more than the claim: this is the **second frontier-robot-model API** after Physical Intelligence's 10-13 beta, so the model layer is being productised as a service. Pricing not stated, which is the only number that would make it relevant at one-arm scale.

**Formic, USD 50M Series B** (trade press; Mercato leading, debt from S2G). A robots-as-a-service operator with 750k production hours, 100+ facilities and 99.6% uptime pivoting to general-purpose robots **by adapting Physical Intelligence foundation models for factory work**. That is the first concrete PI commercial deployment datapoint in this ledger, and it says the near-term business model for foundation policies is adaptation by an incumbent operator, not direct sale.

**Fort Robotics, USD 20M and a Nasdaq SPAC** (trade press; Tiger Global leading ~USD 10M with Mark Cuban Companies plus a USD 10M PIPE commitment). The customer list is the signal, not the round: Agility, DoorDash, Cobot, Zoox, Textron **and Google DeepMind**, on 19,500+ deployed units across 600+ customers — a real deployed number rather than a projection. Safety as infrastructure is apparently the layer that monetises first.

**Maxinsights, USD 30M seed** (trade press, announced via LinkedIn; Peak XV leading). The disclosure worth noting is the customer list: **Genesis AI and Dyna Robotics**. The data-vendor layer underneath the named labs is becoming visible, and MaxGlove is a force-sensing glove built with Genesis AI — so the glove-as-data-rig pattern now has three independent commercial instances (Origami S1, Maxinsights, Daimon's Umi device).

**Agility standardises Digit 5 on NVIDIA IGX Thor, exclusively** (trade press, no financial terms). Single-vendor compute lock-in on a humanoid heading to GA, plus joining NVIDIA's Halos safety inspection lab. Read alongside the Agile ONE shipping preconfigured with GR00T and Isaac assets: NVIDIA is becoming the default *runtime*, not just the default trainer, and the W42 framing holds — the model layer is being rented, the embodiment layer owned.

**IDC: FCC curbs cut US humanoid revenue 58% by 2030** (trade press citing IDC). The FCC Covered List now blocks new foreign robot authorizations, Chinese vendors are ~95% of affected supply, and Intel received a one-year waiver on 10-23 with conditional approvals available in exchange for onshoring commitments. Two readings: policy is now a live variable in robot hardware availability, and — given the §3 finding that American humanoid startups build on Unitree G1 bodies — **the supply constraint lands on exactly the companies least able to absorb it**. Caveat carried forward: IDC's own 2025 baseline of 18,500 shipments against IFR's ~7,000 sales for the same year is a 2.6× gap nobody has reconciled, so treat the forecast as directionally interesting and numerically unsettled.

---

## 6 Deep read of the week

**OpenTac: An Open Ecosystem for Customizable, Scalable, and Transferable 3D Tactile Sensing** — preprint, arXiv 2610.31661v1, 22 Oct 2026. Johns Hopkins + MIT (Slepyan, Bethala, Gu, Dangle, …, Hager, Bharadhwaj, Bhattad, Jatavallabhula, Thakor). Project page opentac.github.io, code github.com/opentac/opentac. I read the paper and verified the project page independently.

### Setup

Task set spans four real embodiments and three problem types. Policy learning: peg-in-hole (from LeFlexiTac) and plug-in-socket (from TacOBench) on an **SO-101 arm**, and key-in-lock on **bimanual YAM arms**. Contact-rich control: ball rolling on a UR5 and cup-holding under impulsive load on a **Dexmate Vega-1 Pro**. Data scale is small and deliberately so: **100 demonstrations per SO-101 task, 50 evaluation trials per condition** at held-out start poses; 25 trials on YAM; 20 trials per condition on the Vega-1. Characterisation: 31 compiled and fabricated boards; five objects pressed at 95 locations for cross-sensor transfer (5,290 predictions); a 10-hour single-bump endurance run. **No vision anywhere in the policy experiments** — tactile plus joint angles only, which is what makes the ablations legible.

### Method

Two things are new, and only one of them is the sensor.

The sensor is a four-layer stack totalling 1.6 mm: a 0.8 mm 3D-printed bump lid, 0.1 mm of Velostat-class piezoresistive film, a 0.1 mm two-layer flex PCB where interdigitated copper fingers form a taxel at each row-column crossing, and a 0.6 mm base plate. The shear trick is purely mechanical and it is the inherited part — one elastomeric bump sits over a 2×2 taxel group, and with baseline-subtracted signals z_uv for u,v ∈ {−1,+1}, you read `h_z = Σ z_uv`, `h_x = Σ u·z_uv`, `h_y = Σ v·z_uv`. Normal load is the common mode; tilting the bump creates a pressure differential that shows up in the signed sums. The paper credits Sun et al. for the bump-over-four-cells principle, so the contribution here is not the transduction.

**The contribution is the compiler.** OpenTac-CAD takes a 2D outline, a developable surface selected on a 3D robot model, or a sentence, and emits fabrication-ready flex-PCB Gerbers, a 3D-printable bump case, and a machine-readable **manifest** describing geometry, electrical addressing and bump grouping. The manifest is the actual architectural idea: it is read by OpenTac-Sim (FEM tetrahedral bump deformation in MuJoCo, instantiating whatever sensor the manifest describes), by OpenTac-Transfer (maps calibrated contact fields between skins through physical coordinates and local frames, with no paired cross-sensor data needed), and by OpenTac-Xampling (adaptive readout). One description, four tools — that is what makes the geometry customisable without rewriting the software, and it is what AnySkin's fixed electronic layout and eFlesh's externally supplied magnetometer electronics each leave undone.

Xampling is the other genuinely new piece. Raster scanning costs N samples regardless of where contact is. Xampling assigns a budget $M \ll N$, spends a quarter of it searching the full mask and the rest following regions active in the previous frame, and degrades gracefully to raster in the worst case.

## Results

**The number that matters: key-in-lock on bimanual YAM, 25 trials — proprioception only 18%, OpenTac-1D (normal force) 43%, OpenTac-3D (normal + shear) 62%.** Shear roughly doubles what normal force buys, on a task where no vision is available and the policy must distinguish key-in-slot from key-on-edge from no-contact.

The baseline that matters is a different one. On SO-101 peg-in-hole, tactile takes OpenTac-3D from 66% to 82%, against 50 → 58% for FlexiTac and 52 → 60% for OpenTac-1D. Plug insertion: OpenTac-3D-full 70% against its own pose-only baseline at 54%. Crucially, **neither normal-only condition improves on its own proprioception-only baseline** on these trials, which is a stronger claim for shear than the headline is — 1D touch is worth nothing here.

Supporting numbers. The compiler: 31 boards, **zero DRC violations, every fabricated board functional**, 0.7–2.6 s per design, 153–467 taxels at 2–4 mm pitch needing only 35–46 channels because an R×C matrix costs R+C lines. Shape-aware packing covers 75/90/89/87/93% of the five highlighted regions against a best-rectangular-grid baseline at 57/90/79/87/93 — identical on near-rectangular regions, +18 and +10 points on the tapered YUBI finger and LEAP palm. Nine natural-language agent sessions (Claude Opus 5 driving the compiler, with web search) all produced DRC-clean layouts, median 6.5 minutes, each logging its dimensional sources and assumptions; those were not fabricated. Three-axis accuracy against an OptoForce reference: $R^2$ 0.95/0.93/0.96, RMSE 0.11/0.16/0.15 N. Xampling: 67 fps raster on 1024 taxels rising monotonically to **1286 fps at M=16, 19.2×**; on a tennis ball dropped 30 cm with a ~3 ms contact, **raster detected 2 of 6 impacts in a single frame each, while M=16–256 detected 3 of 3 and recorded several frames of the event**. Cup-holding under load on the Vega-1, 20 trials: proprioception 0 held, **wrist F/T sensor 9 held at 5.5 ± 1.3 cm slip, OpenTac-1D 9 held at 5.8 ± 2.1, OpenTac-3D 16 held at 2.6 ± 1.2** — a USD-2-per-cm² skin beating a commercial wrench sensor on slip, because the relevant signal is local and tangential and a wrist sensor cannot see it.

## Where it's soft

- **Roughly half the policy gain may be rubber, not signal.** The OpenTac-3D *pose-only* baseline is 66% where FlexiTac-pose is 50% and OpenTac-1D-pose is 52%. The paper names the reason — "adding bumps changes contact mechanics as well as sensitivity to shear" — and that is exactly why the per-sensor pose baseline exists. But nothing isolates it. Of the 32-point gap between FlexiTac-pose (50) and OpenTac-3D-full (82), **about 16 points arrive before any tactile channel is read**. This is structurally the same caveat as T-WLA's (§2), and it is this week's recurring pattern.
- **Three-axis calibration is in-sample.** $R^2$ 0.95/0.93/0.96 is a single bumped 4×4 patch over ±2 N, and the figure caption says so. VecTile, by contrast, reports a held-out four-fold direction-rotation split. Do not compare the two numbers.
- **Calibration needs equipment a small build may not have.** The three-axis procedure as described requires a reference F/T sensor (OptoForce) plus a force-controlled arm. VecTile's five-minute manual calibration against an F/T reference is the cheaper protocol, and it still needs the reference.
- **No bill of materials.** The only cost figure is ≈ USD 1.95 per cm², explicitly **sensor only**, with the readout board called out as excluded and never priced. The project page links a materials page but publishes no BOM. VecTile's USD 34.67 remains the only complete cost disclosure in this literature.

- **Statistics are better than the harvester credited, and still thin.** The policy violins use compact letter display with significance testing — two violins not sharing a letter differ at

  > 90% confidence. That is more than most papers here do. But 90% is a weak threshold, the text never states what the violin distribution is over or how many training seeds were run, the word “seed” does not appear, and the paper’s own phrasing is “highest observed success rate” — the same careful hedge ForceBench used in W42.

- **Shear resolution is one quarter of normal resolution**, because shear is a 4-taxel differential. Never stated.
- **Endurance is n=1.** Ten hours on a single bump, with readings rising for the first two hours before stabilising. The authors recommend a warmup phase or relative rather than absolute readings, which is honest and also a real operational constraint. Hysteresis, drift and temperature dependence are acknowledged in one sentence of the conclusion with no measurements. No crosstalk measurement.
- **Simulation and transfer are weaker than the figures imply, and the authors say so**: simulated signals are deformation proxies, and transfer is evaluated on pose-conditioned **normal-contact shape** (median F1 0.750 against 0.714 for array resizing, 0.788 on co-visible support versus 0.800 same-skin repeatability) — not force magnitude, and **no sim-trained policy is ever deployed**.
- **One external baseline.** FlexiTac only; AnySkin, eFlesh and ReSkin appear in the comparison table but are never run. The key-in-lock fixture is author-built. Flex PCBs are not stretchable and mesh-derived layouts need developable surfaces, so doubly curved fingertips are out.

## Steal list

1. **The manifest.** One machine-readable description of sensor geometry, wiring and grouping, read by acquisition, simulation, transfer and firmware alike. This is a software pattern, not a hardware one, and it is the reason changing the skin does not mean rewriting the stack. Adopt it even with a completely different sensor.
2. **Shear as a signed sum over a 2×2 group under one passive bump.** No new transduction, no extra wires — four existing pressure cells and a printed lid. If you already have any pressure array, a bump lid is the cheapest 3-DoF upgrade available.
3. **The per-sensor proprioception-only baseline.** Train and evaluate a pose-only policy *for each sensor variant*, because mounting hardware changes contact mechanics and demonstration distributions. Almost nobody does this, and in this paper it is what exposes the softest part of the result.
4. **Adaptive partial scanning.** A quarter of the budget searching, the rest tracking last frame’s active region. Firmware-level, geometry-agnostic, and the difference between detecting a 3 ms impact and missing two thirds of them.

## Cost to reproduce a scaled-down version

Hardware, the honest version: one 4×4 bumped patch is a JLCPCB two-layer flex order (tens of dollars), a sheet of LINQSTAT, a 3D-printed TPU lid, and a readout board whose Gerbers are published but whose cost is not — budget USD 100–200 all-in including connectors and a microcontroller, and treat that as an estimate, not a quoted figure, because no BOM exists. VecTile’s published USD 34.67 is the better cost anchor for the same class of device.

Compute is negligible: the policy is an ACT trained on 100 demonstrations with tactile plus joint angles and no vision. That is well under 10 GPU-hours per condition on one consumer card.

The real cost is robot time and the calibration rig. The SO-101 peg-in-hole replication is roughly **100 teleop demonstrations (half a day) plus 50 evaluation trials per condition**. To replicate the comparison properly you need four conditions — 1D-pose, 1D-full, 3D-pose, 3D-full — which is 200 evaluation trials and a second day. Three-axis calibration as the paper describes it needs a reference F/T sensor and a force-controlled arm; without one, you can still get normal-force calibration from a load cell and treat shear as uncalibrated differential counts, which is enough for a policy input and not enough for a published force number.

## 7 Relevant to what you're building

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

One line on what it cost this week, because the cost is now specific rather than general: the harvester filed OpenTac and VecTile at 4 rather than 5 *solely* because there is no named subsystem for them to be load-bearing for, and `CONTEXT.md`'s own example block literally names "cheap tactile sensing with released hardware specs" as a 5-grade interest. Two released-or-fully-specified cheap 3-axis tactile skins with real policy deltas, in one week, is as close to a direct hit as this feed is going to produce.

## 8 Market signal

**No new published compensation band this week.** I looked; the standing weekly board check cannot be executed as specified any more, and that is a finding rather than an absence. Figure's Greenhouse board and OpenAI's careers search publish **neither posting dates nor pay ranges**, so the brief's "new senior manipulation/VLA role with comp" check has no data source on those two surfaces. The only published band in this ledger attached to a title containing "manipulation" remains W42's Skild AI Deep Learning Manipulation Engineer at USD 100,000–300,000 base, and a 3× spread on one posting reads as a level ladder printed as a single req.

**Where hiring is visibly scaling.** Eka Robotics' board lists 11 open roles ten days after emerging from stealth, and the composition is the signal: a **Head of People and a Senior Recruiter** alongside the engineering reqs, which is what a company does before it hires in volume. Exactly one role is a policy title — Machine Learning / Reinforcement Learning Engineer, Cambridge MA, on-site, no comp stated (the board shows a posting date of Mar 2025, which cannot be right for a company that unstealthed on 10-20 — treat board dates on this surface as unreliable). Its requirement list is the useful part, and it is almost exactly the table-stakes list below: RL and supervised learning, physics engines named as Isaac Sim / MuJoCo / PyBullet, ROS 2, transformers and foundation models, expert PyTorch or JAX, reproducible experiments, and **deploy, evaluate and debug policies on physical hardware**. Note what is *not* in the role text despite being the company's entire thesis: force, contact and tactile appear only in the company blurb.

**Table stakes, as of this week.** Flow-matching or diffusion action heads and action chunking are assumed, not differentiating — half the queue is efficiency work on top of them (VLA-Rolling, PaceVLA, SPAR, MambaAct, When One Step Is Enough). Being able to run a released VLA on one GPU and evaluate it honestly is assumed. World-action models are now vocabulary, not expertise.

**Still differentiating, and visibly so.** First, **force and contact as a policy input** — two funded companies on opposite sides of the sensor question in one week, plus CW-WAM, T-WLA, OpenTac, VecTile, ContactKernels, ContactFlowFilter and the sensorless-observer papers. Second, **evaluation honesty as a skill**: the papers that stood out this week did so by reporting bootstrap CIs (CW-WAM), three training seeds (AW-0), compact letter display (OpenTac), or a per-sensor control baseline (OpenTac again) — and being able to say out loud which half of a 32-point gap is hardware and which is signal is a differentiator in an interview because most of the literature cannot. Third, **cost arithmetic**: GPU-hours, episodes, dollars per episode. The Astranomy USD 36.40 per episode figure, SuperMimic's 3 GPU-hours against 128 GPU-weeks, LaLa-World's 90 rollouts against 9,000 — a candidate who quotes these fluently is signalling something different from one who quotes success rates.

**Structural reads worth carrying into a conversation.** The model layer is being rented (two frontier-robot APIs now), the embodiment layer is being bought from Unitree, the runtime is becoming NVIDIA's by default, and the open layer is being filled by hardware vendors and academic labs rather than by framework maintainers — LeRobot has not cut a release in sixteen weeks while shipping a new policy head and 6.4B of Apache-2.0 external weights into its org on `main`. If you work on applied manipulation, the teams that are visibly scaling are the force-and-contact startups, the data vendors underneath the named labs (Maxinsights' customers are Genesis AI and Dyna), and the safety infrastructure layer (Fort's customer list includes Google DeepMind).

## 9 Threads

`THREADS.md` updated. **One thread closed, one promoted from the queue, 12 open and still at cap.** *Simulation fidelity for transmission and contact* hit the W43 kill date set in W40 and the kill condition was not met — no tendon-hand or underactuated-finger sim-to-real result credited the IPC flex contact or Stable Neo-Hookean flex. OpenTac-Sim is the closest thing that landed, and its own conclusion disqualifies it: simulated signals are deformation proxies and no sim-trained policy is ever deployed. So the contact-fidelity half is dead; its surviving reproducibility finding is folded into *Benchmark validity*, where version pinning belongs anyway. *When to replan, and when to think* takes the freed slot, as pre-committed.

| THREAD | STATUS | WHAT WOULD COUNT AS THE NEXT REAL UPDATE |
|---|---|---|
| **Force as a first-class action** | **advancing** | CW-WAM is the first paper in this thread where the predicted wrench *is* the controller reference, with bootstrap CIs and a published negative. W42's falsifier is still unrun: a wrench-action policy trained on position-teleop data matching one trained on force-feedback teleop data. Next: ForceBench's matrix rerun with training seeds and more than one assembly task — and now also CW-WAM's admittance loop reproduced outside Ant Group. |
| **Contact sensing without dedicated tactile hardware** | **advancing — and now contested from both sides** | Enact's USD 30M pre-seed capitalises the thesis; ContactFlowFilter and the sensorless wrench observer advance it; LocoPilot estimates 1.27 N free / 2.25 N loaded with no F/T sensor. But OpenTac and VecTile just made *dedicated* hardware cost USD 2–35, which weakens the thread's premise rather than its evidence. Bar unchanged for nine weeks: a predicted-force policy matching a real-sensor baseline on a task whose labels were not collected with that sensor. |
| **Low-cost open hardware as training substrate** | **advancing — strongest artifact yet** | OpenTac ships MIT code, CERN-OHL-P-2.0 hardware, CC BY 4.0 docs, a hosted CAD tool and policy success rates in one release. RoboParty RP1 added 160 N·m joint modules under CERN-OHL-W-2.0; HX Friends turned out to be Apache-2.0 software with **no CAD and no BOM**. The bar is still the independence clause: a group that did not build it reporting success rates on it. VecTile releases on acceptance; PHD-Hand lists open-sourcing as future work. |
| **Claimed open releases vs actual artifacts** | **narrowing** | Delivered and verified this week: OpenTac (licences confirmed on the project page), InternW0 MIT checkpoints, FineART-VLA, lingbot 6.4B Apache-2.0, ALPS weights, mjfit, Genesis 1.4.5. Pledged, not delivered: VecTile (on acceptance), PHD-Hand, T-WLA, CW-WAM, AW-0, SuperMimic, HATCH (no URL at all). Carried: HumanTouch 1,000 h, TACROSS, SoTa, TouchScale, Praxis-1, StreamPI. |
| **World-action models as unified policy + simulator** | **advancing — first real bound** | AW-0 publishes the bound this thread has needed: under randomization its WAM scores 20.6 against plain π0.5's 27.7. Representational Priors offers a mechanism for why. Next: the two read together, and whether AW-0's OOD regression replicates on a WAM with a different backbone. RoboHarness (2610.14018) is **still unread after eleven days** and is still the biggest single gap here. |
| **Planning on a learned model** | **advancing** | LaLa-World replaces CEM with an LLM actor at 90 rollouts against 9,000 and **releases code**, which makes it the first entry in this thread that can actually be run. It also publishes the counter-case: the LLM actor is worse than CEM on PointMaze. Next: that comparison rerun on real hardware, which remains the thread's unmet bar. |

| THREAD | STATUS | WHAT WOULD COUNT AS THE NEXT REAL UPDATE |
|---|---|---|
| **When to replan, and when to think** | **PROMOTED from the queue this week** | Now eleven-plus papers. This week added PACT-VLA (verify the next chunk against a fresh observation before committing), Robot Metareasoning (learn how long to deliberate), When One Step Is Enough (how few Euler steps flow-matching inference actually needs), FastHarness and TrajWAM. None has code. Next real update: any one of them released and rerun, or a single matched comparison of the latency interventions on one backbone. |
| **Where the episode starts** | **ADVANCING** | Quiet week by comparison. Training in Deployment and SimPT move correction collection into deployment and into simulation respectively. Next, unchanged: a third-party replication of Mulligan's Sobol-plus-farthest-point sampler alone. |
| **Paying at training time, not run time** | **ADVANCING** | T-WLA is the cleanest new instance — future-image and future-pressure heads removed at inference. SharpaDex at 300 h ungated with 71k downloads is the first large tactile corpus in this thread that is actually downloadable, against HumanTouch's and TacClaw-Infinity's unverifiable hour counts. Next: the six-week-unrun matched comparison of one-step action heads on one backbone. |
| **Benchmark validity in physical AI** | **advancing — now carries the version-pinning finding** | Absorbs the surviving half of the killed simulation-fidelity thread. Two model-affecting MuJoCo releases in nine days plus Isaac Lab's `wxyz → xyzw` flip, against W40's 26-point unpinned-MuJoCo result, makes dependency pinning an eval-validity problem rather than a hygiene one. New with a point of view: ContactBench (identical-looking objects, tunable resistance), HATCH (physically replicable escape room, sensor-gated links, run-day seed, open BOM — **no URL given, no results yet**), Sensorimotor Alignment (offline proxy for closed-loop rollouts). Still unmet since W38: a major-lab table with seed bands **and** paired evaluation together — though AW-0's three seeds × 16 rollouts is the closest yet. |
| **Online correction as the last mile** | **ADVANCING** | The deployment-feedback cluster reached seven papers across W42–W43: add TEAR (which evidence source holds authority as conditions drift), Progress Probes (cheap runtime monitors off a frozen VLA), Echo (task status denoised jointly with actions). Next, unchanged: a third-party rerun of EXPO-FT, now seven weeks out with no independent run by anyone. |
| **Who owns the open robot-learning stack** | **advancing — the shape is now explicit** | LeRobot has shipped no release in sixteen weeks (tenth negative, v0.7.0 tag 404s) while shipping a new policy head and 6.4B of external Apache-2.0 weights on `main` — the release cadence is dead, the repo is not. Against that: Isaac Lab 3.0 GA, Isaac OS on Holoscan, a Human 1M Dataset restricted to Cosmos Coalition partners, Agility making NVIDIA its **exclusive** onboard compute, and Agile ONE shipping GR00T preinstalled. Next: whether Isaac OS actually appears under an open licence, and whether any platform-holder commitment is written into a foundation or steering body. Still none has been. |
| **~~Simulation fidelity for transmission and contact~~** | **CLOSED at W43, as pre-committed in W40** | Kill condition not met: no sim-to-real transfer result credited the discrete integrator, IPC flex contact, Stable Neo-Hookean flex or routed-cable modelling in three weeks of looking. The reproducibility half moves to *Benchmark validity*. Reopen only if a tendon-hand or underactuated-finger transfer result attributes transfer to a named contact model. |

**Still queued behind the cap:**

*Human video as the data substrate, and where it saturates* — carried from W41 for the fourth week. This week added more volume (SpatialDex learns 3D-grounded bimanual dexterity from RGB-only human video with no depth rig, EgoHand puts NVIDIA into the egocentric-hand cluster, EgoVine compiles branching experience from real video) and **still no replication** of TRI's 30,000-hour saturation finding. Volume does not promote it. Third consecutive week.

## 10 Open question

### On a bumped tactile skin, how much of the policy gain is the signal and how much is the rubber?

OpenTac's own control design is what raises this, which is to its credit. On SO-101 peg-in-hole the three pose-only baselines are not equal: FlexiTac-pose 50%, OpenTac-1D-pose 52%, **OpenTac-3D-pose 66%**. The 3D condition differs from the 1D condition *only* by the printed bump lid — same skin, same readout board. So before any tactile channel enters the observation, the bump geometry is worth roughly 14 points, and adding the tactile channels is worth a further 16. The paper names the cause ("adding bumps changes contact mechanics as well as sensitivity to shear") and controls for it by training a per-sensor pose baseline, but it never isolates it — and the headline comparison most readers will take away, 82% against FlexiTac's 58%, bundles the two effects.

This is not a quibble about one paper, because the same shape appeared twice more this week. T-WLA's own ablation shows tactile input buying 47.25 → 50.0 while the visual branch carries 17 → 47. VecTile reports its gain as living in *corrections* (35 → 61 → 86%) rather than first-attempt alignment, which is consistent with either reading: better contact feedback, or a compliant nubbed finger that simply recovers from misalignment more gracefully. And compliant-finger-geometry-as-regularisation already has its own W43 paper — Hardware-as-Regularization, where a 3D-printed spiral-spring clutch on the leader arm takes block stacking from 36.7% to 73.3% **with no sensing added at all**. Three independent results in one week where passive mechanics plausibly account for a large share of a gain attributed to sensing.

Why it matters at one-arm scale: the two interventions have completely different costs. A printed compliant lid is a day and a spool of TPU. A calibrated 3-axis skin with readout electronics, a reference F/T sensor for calibration and a tactile branch in the policy is weeks. If most of the benefit is mechanical, the correct first move is to print compliant fingertips and spend the saved weeks on demonstrations.

**What would settle it:**

A passive-bump control. Install the bump lid, collect demonstrations normally, train the policy with the tactile channels **zeroed or shuffled** rather than removed, and evaluate on the same 50 held-out start poses. That isolates grip mechanics from the signal, which neither a flat-sensor baseline nor a pose-only baseline does — a flat-sensor baseline changes the mechanics, and a pose-only baseline on the bumped sensor still has the bump but also has a different demonstration distribution collected with it. Three conditions, one arm: 3D-full, 3D-zeroed, 1D-full, at 100 demos and 50 trials each. One extra ACT training run (well under 10 GPU-hours) and roughly 150 evaluation trials, which is a day and a half of robot time. If 3D-zeroed lands near 66, the bump is the hardware contribution and shear is worth a real 16 points. If it lands near 80, the shear channel is mostly decorative and the field has spent a year instrumenting what it should have been moulding.

Secondary version, cheaper and available to anyone with a printer: take VecTile's or OpenTac's bump geometry, print it onto a *plain* gripper pad with no sensor underneath, and run the same insertion task. If a dumb nubbed pad recovers most of the gain, that is the single most cost-effective finding anyone could publish this quarter.

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.

- Blind Precise Assembly — MIT, Agrawal group, read in full by the harvester. **Joint encoders alone** insert eight different pegs in 73% of 237 trials zero-shot. Proprioception-only sim-to-real, one-arm reachable, and the `mjfit` generator is released. This is the research counterpart to Eka's funding and would have been a §2 lead in a quieter week.
- HexaShell — a thin harness where a Codex agent writes its own triangulation tools; RoboDojo SOTA 39.7%; **code released**; the limit it reports is contact, not reasoning.
- Has GPT-6-Astra Achieved Embodied Intelligence? — twelve benchmarks, eight real configurations, code released; competence crossed, practicality bound by in-context demo cost and latency. §1 covers the cluster.
- Physical Canvas — visuo-tactile world model used as an interactive simulator; GRPO post-training *inside* it lifts real π0.5 success 59.1 → 70.8 and raises sim-real correlation 0.602 → 0.885. Four real tasks, robot unnamed, project page only, no code.
- Video-Point Model — KAIST + ETH + SNU + NYU + RLWRLD; multi-view video world model with a 3D point-trajectory head, actions denoised jointly with video. **Code at `github.com/rootyJeon/VPM`**, VPM-Data 10K clips, VP-WorldBench 1,200 clips, FVD 78.52 → 67.35, real Franka FR3. Needs known intrinsics and poses.
- NVIDIA Flora — release; the GR00T-2 data engine itself, as three checkpoints (captioner, rewarder, temporal segmenter) that turn raw robot video into labelled training data. The card discloses that this is what trained GR00T-2. Licence is **NVIDIA OneWay Noncommercial, research and development only**, and the 27B captioner is not one consumer GPU at bf16. The trade-press "200,000 hours processed daily" figure is not in the card.
- X2Real — trade press, read in full; 44 tasks, 10 dimensions, sim-real correlation 0.84 on progress and 0.74 on success for WALL-OSS-0.5, with a public eval runtime over 73 configs. OOD definitions and data generation withheld, and one counterexample at 70% real / 0% sim.
- NVIDIA Cosmos opens world-model tooling — blog, 10-21, verified this week after two failed fetches. Released code and a HuggingFace collection, with **post-training and RL recipes for GR00T, π0.5 and Alpamayo** as the build-relevant part. No licence and no benchmark figures in the post.
- DONIC — Ropedia; one bimanual dexterous tracker across references, unseen trajectory 86.79% and unseen object 57.32% from one checkpoint with no finetune. Held at 3 by the harvester because it is **Isaac Gym only** and physical transfer is future work; the baselines are the authors' own bimanual adaptations, so the +77 pp delta is not like-for-like.
- LocoPilot, SuperMimic, T-WLA, CW-WAM, ContactBench, ContactKernels, Bristol in-hand rotation, SharpaDex, ALPS, LaLa-World, AW-0, VecTile, InternW0, MuJoCo 3.17.0, Isaac Lab 3.0 GA and FineART-VLA are all covered in the sections above.

---

*Sources are linked inline. Source types: arXiv entries are preprints unless stated; HuggingFace cards and GitHub release tags are releases; humanoidsdaily, therobotreport, theaiinsider, robotics247, unite.ai, Fortune and SCMP are trade press; blogs.nvidia.com, agilityrobotics.com, agibot.com, origami-robotics.com and pi.website are company blogs; Unitree H2-X and XPENG IRON are demo videos. Single-lab claims with no released artifact are flagged unreproduced where they appear.*