Journal of Advanced Robotics, Autonomous Systems and Human-Machine Interaction

PRIME OPEN ACCESS

Volume 2, Issue 3

Research Article

Date of Submission: 12 June, 2026
Date of Acceptance: 14 July, 2026
Date of Publication: 24 July, 2026

The State of Robotic Intelligence: An Exhaustive Analysis of Learning, Control, Vision, and Safety Benchmarks (2024-2026)

Narayana Sanghea Panchumarthy*

Independent Researcher, Cambridge, United Kingdom

***Corresponding Author:** Narayana Sanghea Panchumarthy, Independent Researcher, Cambridge, United Kingdom.

Citation: Panchumarthy, N. S. (2026). The State of Robotic Intelligence: An Exhaustive Analysis of Learning, Control, Vision, and Safety Benchmarks (2024-2026). *J Adv Robot Auton Syst Hum Mach Interact*, 2(3), 01-08.

Summary
The period spanning 2024 to 2026 marks a transformative epoch in robotics, characterized by the collapse of the traditional modular pipeline in favor of end-to-end foundation models and the emergence of standardized, large-scale benchmarks for embodied intelligence. This Transition is not merely incremental but represents a fundamental shift in how robotic agents are trained, evaluated, and deployed across industrial and domestic environments. As artificial intelligence capability accelerates, reaching a point where industry produces over 90% of frontier models, the metrics for robotic success have evolved from simple tasks completion to complex measures of cross-embodiment transfer, zero-shot generalization, and high-fidelity spatial reasoning [1].

The Convergence of Vision, Language, and Action: The Foundation Model Paradigm
The most significant architectural development in this era is the rise of Vision-LanguageAction (VLA) models. These systems represent the architectural thesis of the foundationmodel era in robotics, where a single transformer backbone consumes camera frames and natural language instructions to emit robot motor commands as tokens or continuous control signals [2]. By late 2024, the field moved from experimental demonstrations like RT-1 and RT-2 toward production-ready generalists like Octo and OPENVLA, eventually culminating in the 2025-2026 rollout of high-capacity models such as and Google Gemini Robotics series [2-4]. The technical superiority of VLA models lies in their ability to leverage large-scale pre- training on internet video and text to bootstrap physical understanding. For instance, RT-2 leverages PALI-X pre-training to achieve a 91% success rate on tasks involving novel objects, a feat that pure robotics models struggle to match. The scaling laws in robotics are now being rigorously tested, with evidence suggesting that training one model across dozens of robot types—the "Open X-Embodiment" recipe—lifts a target-robot's score by more than 50% over single-embodiment training. This positive transfer is the holy grail of robot learning, enabling a model trained on a mobile manipulator to improve the capabilities of a tabletop arm by sharing high-level semantic and structural knowledge [2,5-7].

Model Class	Creator	Parameter Scale	Action Representation	Open-Source Status	Primary Benchmarks
RT-X / RT-2	Google DeepMind	55B	Discrete Autoregressive Tokens	Weights Available (RT-X)	Open X-Embodiment, RT-1 Tasks [5, 8]

OPENVLA	Stanford/Berkeley	7.5B	Discrete Tokens	Yes (Apache 2.0)	LIBERO, SIMPLER [8]
PALI-X	Physical Intelligence	~3B (est.)	Flow Matching / Diffusion	No	Real-world Dexterity, Contact-rich tasks [8, 9]
SMOLVLA	Hugging Face	450M	Chunked Prediction (ACT-style)	Yes (Apache 2.0)	LIBERO-Long, SIMPLER [8]
OCTO	UC Berkeley	~93M	Continuous Diffusion	Yes (MIT)	Bridge V2, Open X-Embodiment [8, 10]
ROBOFLAMINGO	Shanghai AI Lab	~9B	Autoregressive	Yes	CALVIN, LIBERO [8]

A critical tension in VLA development is the trade-off between semantic reasoning and control frequency. High-capacity models like the 55B-parameter RT-2 offer unparalleled symbolic understanding—such as "picking up an object Taylor Swift would like" by selecting a heart-shaped item—but are limited by inference speeds of approximately 1 Hz. Conversely, lightweight models like SMOLVLA (450M) can run at 20-30 Hz on standard GPUs, making them far better suited for real-time, high-precision control in dynamic environments, albeit with reduced long-horizon reasoning capacity [4,5,8].

Data Infrastructures and the Open X-Embodiment Standard

The acceleration of robot learning is directly predicated on the availability of data on a scale. Prior to 2024, robot learning was fragmented, with researchers training on tiny, lab-specific datasets. The launch of the Open X-Embodiment (OXE) project, a collaborative effort involving 34 labs and 22 robot types, established the "ImageNet moment" for robotics. By aggregating over 1 million trajectories and 527 distinct skills into a standardized format compliant with the Reinforcement Learning Datasets (RLDS) schema, OXE has provided the first rigorous backbone for analyzing scaling phenomena in cross-embodiment generalization [2,5-7,10].

Data standardization is the secondary, yet equally vital, pillar of this revolution. The industry has largely converged into two formats: RLDS for large-scale multi-robot datasets and the LEROBOT format for lightweight, Hugging Face-integrated training workflows [8]. This standardization allows for the creation of unified data loaders and augmentation protocols, facilitating direct comparability across research groups and enabling zero-shot transfer evaluations that were previously impossible [6,8].

Dataset Name	Primary Repository	Trajectories	Robot Embodiments	Key Characteristics
Open X-Embodiment (OXE)	Google DeepMind / Consortium	1M+	22	Largest cross-platform manipulation corpus [5, 7]
DROID	CMU / Toyota Research	76K	1 (Franka Panda)	564 scenes, 1,500 unique objects, multi-view [5, 10]
BRIDGE V2	UC Berkeley / Stanford	60K	1 (Widow X 250)	High-quality kitchen/domestic manipulation [10]

ROBONET	Stanford	15M Frames	7	Focused on video prediction and visual dynamics [10]
FRACTAL	Google	130K	1 (Mobile Manipulator)	Used for RT-1 and RT-2 instruction following [10]
RH20T	Meta AI	21K	Multiple	Focused on multi-finger dexterous manipulation [10]

By 2026, the data aggregation race has intensified, with industry leaders like Tesla and Figure AI targeting datasets of 10 million or more trajectories. The emergence of "Data Factories," such as those developed by Tutor Intelligence, marks a shift toward the industrialization of data collection, where robots are trained in high-fidelity simulations and fine-tuned on physical hardware to close the reality gap [10-12].

Spatial Intelligence: Moving Toward Metric and Geometric Reasoning

A persistent challenge for the current generation of Vision-Language Models (VLMs) is their relative "spatial blindness." While they excel at 2D image tasks, they often struggle with 3D spatial reasoning, such as judging distance, size differences, or maintaining consistency across multiple camera views. Spatial intelligence is defined as the ability to perceive, reason about, and act in 3D space—understanding not just what objects are, but them geometric relationships and physical properties [13-15].

Recent benchmarks like MV-ROBOBENCH and ESPIRE have been introduced to diagnose these shortcomings. MV-ROBOBENCH focuses on multi-view robotic scenes, requiring the integration of complementary perspectives to mitigate depth ambiguity and occlusion. Results show that even state-of-the-art models remain far below human performance in these multi-camera environments, indicating a significant gap in the 3D spatial knowledge of current training data [14]. The ESPIRE benchmark further extends this by evaluating agents on "spatial-reasoning-centric" tasks in high-clutter simulated worlds, revealing that while reflection can help with object localization; it does not necessarily translate into successful action execution [14,16,17].

To address these limitations, the field is pivoting toward 3D foundation models and real-time reconstruction technologies. 3D Gaussian Splatting, popularized in late 2023, has become the standard for spatial representation in 2026, allowing for 1000 FPS reconstruction and real- time planning at 30 Hz. This technology enables robots to build metrically accurate 3D models of their surroundings, facilitating tasks like surgical robotics or bin-picking in cluttered warehouses where sub-millimeter precision is required [9,15].

The architecture of a spatially intelligent robot in 2026 typically includes a perception layer that lifts RGB-D inputs into dense point clouds using Gaussian Splatting, followed by a spatial reasoning layer like Spatial VLM that grounds language instructions in the resulting 3D geometry [15]. This shift is supported by massive industrial investment, with companies like World Labs raising $230M to build "Large World Models" that understand persistence and 3D consistency from video [15].

Benchmark	**Tasks**	**Metric**	**Key Insight**
MV-ROBOBENCH	1.7K QA items across 8 subtasks	Human-Robot Correlation	Spatial intelligence and task success is positively correlated [14, 16]
ESPIRE	Pick-and-Place in cluttered scenes	Accuracy vs. Acceptance vs. Success	Reflection helps localization but execution remains a bottleneck [17]
3DSRBENCH	Qualitative/Quantitative relationships	Distance/Size Recognition	VLMs lack 3D spatial knowledge from 2D training data [13]

SPATIALVLM	Fine-grained 3D QA	Accuracy (87% Spatial, 92% Dist.)	Scaling 3D spatial data significantly enhances performance [15]
MINDCUBE	3D transformation reasoning	12.94% over GPT-4 baseline	PYSPATIAL outperforms general-purpose MLLMs [18]

Manipulation Benchmarks: Complexity, Contact, and Recovery

Robotic manipulation has historically been divided into atomic tasks like pick-and-place, which are now considered largely solved (95%+ success in ROBOSUITE Lift and Meta- World) [2]. However, the research frontier has moved to contact-rich manipulation and longhorizon task chaining. Contact-rich tasks, such as peg-in-hole insertion or furniture assembly, require a sophisticated blend of force sensing, compliance, and adaptive control [2].

Benchmarks like Furniture Bench and Peg Insertion typically show success rates between 60% and 80%, highlighting the difficulty of handling physical uncertainties and tolerances [2].

Long-horizon tasks remain the most difficult frontier. Success rates for tasks involving 5 or more steps still land below 50% on benchmarks like LIBERO and CALVIN. The primary failure mode is error propagation, where a minor deviation in an early-stage cascades into an unrecoverable failure later in the sequence. To combat this, researchers are developing robust failure recovery and autonomous error correction frameworks [2].

The ROBOFAC (Robotic Failure Analysis and Correction) framework represents a paradigm shift in how failures are addressed. By decomposing robotic failures into six fundamental categories, task planning, motion planning, execution control, perception, environmental, and hardware—Robo FAC allows for the training of lightweight “critic” models that can reason not just that something went wrong, but why. This is critical for industrial applications where “blind” execution can lead to equipment damage or safety incidents [19].

Failure Category	Root Cause	Example Phenomenon
Task Planning	Incorrect decomposition	Omission of a required sub-stage (e.g., forgetting to open a lid) [19]
Motion Planning	Limited spatial reasoning	End-effector position or orientation deviation from target [19]
Execution Control	Physical imprecision	Gripper closure insufficient or object slipping [19]
Perception	Sensor occlusion/lighting	Sharp edge not identified or object misclassified [20]
Timing	Temporal offsets	Executing a subtask at an incorrect moment [19]
Human Error	Teleoperation drift	Incorrect manual guidance or unexpected obstruction [20]

Systems like FDEC (Failure Detection and Error Correction) have demonstrated that using a lightweight visual dynamics model for predictive monitoring, coupled with a VLM for verification, can raise task success by 26.7% .In the industrial context, the ROCO (Robotic Collaborative Assembly Assistance) Challenge emphasizes the need for bimanual coordination and robust recovery from “stuck” scenarios, which are common in gearbox construction and other high-precision assemblies [21,22].

Control Theory and Agile Locomotion: The Humanoid Surge

Control paradigms are evolving from classical Template-based controllers toward end-to-end Reinforcement Learning

(RL). While traditional methods like Zero Moment Point (ZMP) were effective in controlled environments, they were brittle to perturbations and required extensive manual tuning [23]. Modern RL-based approaches allow robots to autonomously discover optimal movement strategies, adapting dynamically to slips, pushes, and uneven terrain.

The Benchmarking of locomotion algorithms involves comparing PPO, SAC, and TD3 across diverse morphologies. Quadrupedal robots have achieved near-perfect reliability in trotting and climbing, with systems reaching climbing speeds 232x faster than the prior state-of-theart. Bipedal and humanoid platforms, such as the Unitree G1 and Booster T1, are now being evaluated using frameworks like AGILE, which provides a standardized pipeline for interactive environment verification and unified evaluation of skills like stand-up, locomotion, and loco manipulation [23-25].

Locomotion Skill	Hardware Platform	Action Space	Training Time	Deployment Performance
Quadruped Ladder Climb	Unitree A1 / Hooked Effector	Joint Position	Simulation-heavy	90% zero-shot success (70°-90° angles) [24]
Humanoid Locomotion	Unitree G1	Leg Position (12 DOF)	10h (GPU accel)	Robust on rough terrain [25]
Fall Recovery (Stand-up)	Booster T1	All Joints (Rel)	15h-25h	Consistent sim-to-real transfer [25]
Nonprehensile Transport	Mobile Manipulator	Tray-balanced Object	Optimization-based	Handles large inertial uncertainty [26]
Bipedal Flip-Walking	Flip Walker (0.78 kg)	Underactuated Dynamics	Ground Reaction	0.2 body lengths/sec on snow/rock [24]

The waiter's problem transporting an object on a tray with uncertain inertial parameters— serves as a high-stakes benchmark for control robustness. Recent optimization-based motion planning frameworks can handle substantial parameter uncertainty, ensuring the object is not dropped even during quick, jerky movements. Similarly, Hierarchical Model Predictive Control (MPC) schemes are being deployed on quadrotors to increase planning horizons by a factor of 5, enabling real-time collision-free navigation through complex, static environments [26,27].

Navigation in Human-Populated Environments: The Autonomy Gap

Autonomous navigation benchmarks are shifting focus from open-space traversal to highly constrained and socially dynamic environments. The Benchmark Autonomous Robot Navigation (BARN) Challenge is the primary vehicle for evaluating navigation in narrow, cluttered spaces [28]. Using a standardized Clear path Jackal robot equipped with a 2D LiDAR and a motor controller capped at 2m/s, participants must navigate 300 pre-generated environments ranging from easy open spaces to difficult narrow gaps [28].

Performance in BARN is evaluated based on a standardized metric that clips traversal time within a range (to) to prevent outliers from disproportionately scaling scores [28]. The 2025 competition results showed that state-of-the-art systems are reaching success rates near 90%, prompting the refinement of scoring bounds to reflect higher efficiency [2,28].

Social navigation (SOCNAV) introduces an additional layer of complexity: robots must respect social norms, maintain comfortable distances, and interpret human cues. A Comprehensive review of 85 SOCNAV papers published between 2020 and 2025 reveals that while learning-based methods frequently outperform model-based approaches in scenarios like navigating doorways, the community still lacks a de facto standard benchmark, leading to contradictory conclusions across different studies. Current trends prioritize models that accurately anticipate human movement and training environments that realistically simulate crowded spaces [29-31].

Navigation Paradigm	Key Benchmark	Evaluation Metric	Deployment Status
Highly Constrained	BARN Challenge	Traversal Time / Success Rate	Commercially deployed in warehouses [28, 32]
Photorealistic 3D	Habitat (Meta)	Navigation Error / SPL	High success (90%+) in structured ENVS [2]
Socially Aware	ROBOSENSE (Track 2)	Human Comfort / Trust	Mostly in laboratory/pilot phases [32, 33]
Urban/Unstructured	IRIS-U²	Collision Probability / Coverage	Commercially deployed for food delivery [26, 32]

Real-world deployment remains uneven. While autonomous navigation is commercially deployed for food delivery and warehouse logistics, tasks requiring objects to be handled, assembled, or manipulated largely remain in the lab. For instance, delivery robots can handle thousands of object types in warehouses because the environment is stable and designed for them, but reliability drops precipitously as the environment becomes more variable, such as in a typical household [32].

Safety, Governance, and Trustworthy Evaluation Architectures

As robots move into public spaces, safety transitions from a hardware constraint to a systemwide application requirement. Collaborative robot (cobot) safety is evolving from physical barriers toward probabilistic, cyber-aware protection systems. [34] This is codified in standards like ISO 15066, which defines modes for Speed and Separation Monitoring (SSM) and Power and Force Limiting (PFL) [34,35].

The foundational challenge in cobot safety is the safety-efficiency trade-off: every safety intervention (e.g., speed reduction) reduces productivity. The technical goal is to make safety responses as precise and proportionate as possible using advanced optical sensors, torque sensors, and proximity monitoring. In 2026, the industry is witnessing the emergence of "Cognitive Safety," which concerns the human operator's ability to anticipate and trust robot behavior [36]. A mechanically safe robot may still undermine cognitive safety if its actions are unpredictable, leading to operator stress and the eventual erosion of trust [34-36].

The NIST AI Agent Standards Initiative (launched in early 2026) is the leading regulatory response to autonomous agents. This initiative aims to address [37,38].

- **Security Controls:** Mitigating privilege abuse and unintended autonomous actions.
- **Identity and Authorization:** Mechanisms to authenticate agents that interact with APIs and physical infrastructure.
- **Interoperability:** Protocols enabling reliable interaction among heterogeneous fleets from multiple vendors.
- **Assessing Performance:** Methodologies for evaluating agent resilience and compliance with emerging governance expectations.

Cybersecurity-integrated safety is a rapidly growing cluster, with 2024 patents from Intel and Fanuc addressing post-attack trajectory recovery and automated hazard identification. The risk vectors for embodied AI expand beyond data loss to physical harm and infrastructure disruption, making cybersecurity a "safety-critical" component [34,39].

Safety Sub-Domain	Mechanism	Standard/Reference	Future Direction
Speed/Separation (SSM)	Modulates speed based on human proximity	ISO/TS 15066 [35]	Perception-uncertainty-aware planning [34]
Force Limiting (PFL)	Caps contact forces to safe thresholds	ISO 10218 [36]	High-fidelity tactile feedback integration [40]

Cognitive Safety	Predictable profiles / Explainability	Cognitive Frameworks [36]	Human-in-the-loop co-learning [27]
Cybersecurity Safety	Trajectory recovery / Intrusion detection	NIST CAISI RFI [39]	Cyber-aware protection systems [34]
Formal Verification	Risk assessment tooling	Fanuc 2024 Patent [34]	Digital twins for certification [35, 41]

Evaluation methodologies must keep pace with these capabilities. The Auto Eval architecture represents state-of-the-art in "Trustworthy Evaluation," establishing metrics for execution quality (how well the task is performed) and source authenticity (distinguishing between autonomous behavior and teleoperated counterfeits with 99.6% accuracy). This are essential for restoring trust in robotic demonstrations, which are often curated or teleoperated to mask failures [42].

Real-World Deployment: Scaling to the Factory Floor

The transition from lab to production is being led by the automotive and logistics sectors. Tesla's Optimus program is the most high-profile example, with the company phasing out legacy Model S and X production lines in Fremont to make room for a humanoid factory capable of producing 1 million units annually [11,12]. Long-term targets in Texas Gigafactory aims for 10 million units per year, supported by the development of the AI5 inference processor and the "Digital Optimus" intelligence layer [12].

Logistics automation is similarly massive. Amazon's fleet of 1 million robots now handles 75% of global deliveries, with systems like "Sequoia" increasing warehouse efficiency by 75%. Surgical robotics has also reached a tipping point, with robotic procedures accounting for 60% of surgeries in major hospitals, led by Intuitive Surgical DA Vinci 5 platform. Despite these successes, the "Sim-to-Real Cliff" remains a critical bottleneck for new players [43].

The stark contrast between simulation rankings and onsite performance in the ROCO Challenge revealed that modular systems, while perfect in simulation, proved brittle onsite due to calibration sensitivity and communication delays. VLA-based approaches, However, demonstrated higher tolerance for real-world sensor noise, suggesting that end-toend learning may be the key to robust deployment [22].

Sector	Deployment Scale (2026)	Market Leader	Primary Technological Driver
Industrial Humanoids	50,000 units by year-end	Tesla Optimus / AGIBOT	VLA-driven generalist policies [12, 43]
Warehouse Logistics	1M+ robot fleets	Amazon Sequoia / Robotics	Autonomous path planning and fleet orchestration [43]
Healthcare	60% of major procedures	Intuitive Surgical	High-precision teleoperation and 3D reconstruction [15, 43]
Service/Consumer	11% growth YoY	Diverse (Food/Goods)	Autonomous navigation in structured ENVS [32, 44]
Advanced Manufacturing	Industrial Stock of 2M units	China (various)	AI-powered robots as core national strategy [44]

In this landscape, "Fleet Orchestration" has become a primary purchasing criterion, as organizations increasingly deploy heterogeneous fleets from multiple vendors. Interoperability challenges can lead to congestion, accidents, and downtime, making autonomous management systems essential for the digital enterprise of Industry 4.0. [9,41,45].

Conclusion: The Path Toward General Purpose Autonomy

The synthesis of robotic benchmarks from 2024 to 2026 illustrates a field that has finally found its "scaling playbook." The shift from narrow, task-specific controllers to high-capacity VLA foundation models, supported by the standardization of multi-robot datasets like Open X-Embodiment, has unlocked emergent capabilities in symbolic reasoning and crossmorphology transfer. However, the journey toward true AGI remains distant, with long- Horizon task success rates and nuanced spatial reasoning still present significant research hurdles.

The benchmarks of the next two years will likely focus on "Agentic AI" systems that combine analytical decision-making with generative adaptability to work independently in complex, real-world environments As regulatory frameworks like the EU's AI Act and US OSHA guidance solidify, the industry will move toward a standard were policy evaluation is a distinct product category, combining simulation, standardized benchmark tasks, and automated regression testing. The robotics industry in 2027 will be defined less by hardware breakthroughs than by the software and data infrastructure maturation that occurred during this pivotal 2024-2026 window. Success for both academic and industrial stakeholders will depend on the development of repeatable data collection workflows, rigorous evaluation systems, and the ability to bridge the persistent gap between digital imagination and physical reality [45,46].

References

1. The 2026 AI Index Report Stanford HAI,
2. Robotics AI Benchmarks 2026: State of the Art - Codesota,
3. Robotics Benchmarks - Robots - CodeSOTA,
4. Google Robotics Foundation Models 2026 | RT-X, Gemini Robotics, and Open X- Embodiment | RoboCloud Hub,
5. Open X-Embodiment & Robotics Foundation Models Benchmark,
6. Open X-Embodiment Dataset - Emergent Mind,
7. Open X-Embodiment: Robotic Learning Datasets and RT-X Models - Jiajun Wu,
8. VLA Models Compared: OpenVLA vs π0 vs SmolVLA vs RT-2 (2026 Guide) | SVRC,
9. 12 Predictions for Embodied AI and Robotics in 2026 | Dylan Bourgeois,
10. Robot Learning Datasets 2025: RT-X, DROID, Bridge & Beyond | RoboCloud Hub,
11. Tesla (TSLA) is preparing 10 million Optimus robots per year production line at Gigafactory Texas - Shack News,
12. From EVs to robotics: Tesla targets 10M Optimus units with new Texas plant,
13. 3DSRBENCH: A Comprehensive 3D Spatial Reasoning Benchmark - ResearchGate,
14. Seeing Across Views: Benchmarking Spatial Reasoning of Vision-Language Models in Robotic Scenes - Microsoft Research,
15. Spatial Intelligence for Robotics: The 2025 Frontier of 3D Understanding | RoboCloud Hub,
16. Seeing Across Views: Benchmarking Spatial Reasoning of Vision-Language Models in Robotic Scenes - arXiv,
17. Espire: A Diagnostic Benchmark for Embodied Spatial Reasoning of Vision-Language Models - arXiv,
18. pySpatial: Generating 3D Visual Programs for Zero-Shot Spatial Reasoning - arXiv,
19. RoboFAC: A Comprehensive Framework for Robotic Failure Analysis and Correction - arXiv,
20. Designing for Error Recovery in Human-Robot Interaction - arXiv,
21. Failure Detection with Zero-Shot Error Correction in Robotic Manipulation - ResearchGate,
22. RoCo Challenge at AAAI 2026: Benchmarking Robotic Collaborative Manipulation for Assembly Towards Industrial Automation - arXiv,
23. 14 IV April 2026 IJRASET,
24. IROS 2025 Program Wednesday October 22, 2025,
25. AGILE: A Comprehensive Workflow for Humanoid Loco-Manipulation Learning - arXiv,
26. ICRA 2026 Program Wednesday June 3, 2026,
27. ICRA 2026 Program Thursday June 4, 2026,
28. ICRA 2026 BARN Challenge GMU CS Department,
29. Social robot navigation: a review and benchmarking of learning-based methods,
30. Social robot navigation: a review and benchmarking of learning-based methods -Frontiers,
31. A short methodological review on social robot navigation benchmarking - arXiv,
32. Where Autonomy Works: Evaluating Robot Capabilities in 2026 | Epoch AI,
33. Competitions - IROS2025,
34. Collaborative robot safety technology landscape 2026 - PatSnap,
35. Collaborative Application Standards Are Shaping Manufacturing Robotics - Benchmark Electronics,
36. A Systematic Review of Safety-Driven Approaches in Human–Robot Collaborative Systems,
37. NIST launches AI Agent Standards Initiative as autonomous AI moves into production,
38. NIST Launches AI Agent Standards Initiative and Seeks Industry Input,
39. 9 March 2026 Peter Cihon, Senior Advisor Center for AI Standards and Innovation (CAISI) National Institute of Standards and T - IEEE-USA,
40. IROS 2025 - Genenrative AI for Robotics and Smart Manufacturing,
41. Physical AI and humanoid robots | Deloitte Insights,
42. Trustworthy Evaluation of Robotic Manipulation: A New Benchmark and AutoEval Methods,
43. Who's Leading the Way in Robotics, Humaniods & Physical AI in 2026 - RAISE Summit,
44. International Federation of Robotics,
45. Top 5 Global Robotics Trends 2026,
46. State of Robotics 2026 Report: $38B Market, 12 Humanoids, VLA Adoption,