

Rise of the Machines: Deep Research on the Most Important Work and Breakthroughs in AI Robotics from the Past 7 Days (August 6-12, 2025)

Introduction: The Humanoid Ascendancy

The theme of a "Rise of the Machines," long confined to the realm of speculative fiction, has found tangible, material expression in the global technology landscape over the past seven days. This is not the rise of a malevolent intelligence, but rather the ascendancy of a specific technological form: the humanoid robot. The period of August 6-12, 2025, will be remembered as a watershed moment when the humanoid form factor transitioned from a subject of laboratory research into a focal point of intense, large-scale commercial and geopolitical strategy. This acceleration is not the result of a single breakthrough, but a powerful convergence of three distinct yet deeply interconnected forces: the maturation of sophisticated, cost-effective hardware; profound advances in the development of embodied artificial intelligence; and the rollout of deliberate, ambitious commercialization initiatives. The week's events, overwhelmingly centered around the 2025 World Robot Conference (WRC) in Beijing, signal a global pivot towards the humanoid as the chosen vessel for general-purpose robotics, designed to operate seamlessly within the complex, unstructured environments built by and for humans.¹

The industry zeitgeist, captured by the International Federation of Robotics and other observers, has decisively shifted toward humanoids, with many experts declaring 2025 as "Year One" of a new industrial era.⁵ The WRC, with its pointed theme of "Making Robots Smarter, Making Embodied Agents More Intelligent," served as the primary nexus for this paradigm shift.² The conference was far more than a simple trade exposition; it was a curated demonstration of national ambition, showcasing not just robotic hardware but a clear strategic reorientation from pure manufacturing prowess to a focus on the AI "brain" that animates the machine. This report will dissect the key developments that defined this pivotal week: the unveiling of new,

highly capable humanoid platforms from companies such as EngineAI and Fourier Intelligence ⁷; spectacular public demonstrations of advanced locomotion and task completion from firms like Unitree Robotics and Galaxea AI ⁴; foundational AI software releases from technology giants like NVIDIA that promise to serve as the nervous system for this new generation of machines ¹⁰; and fundamental breakthroughs in sensory technology from academic institutions that aim to close the gap between mechanical and biological perception.¹¹

Underpinning this flurry of activity is a profound strategic wager being placed by corporations and nations alike. The intense global focus on the humanoid form is not merely a matter of technical curiosity; it represents a calculated bet that a single, general-purpose robotic platform, capable of using human tools and navigating human-centric spaces, constitutes a vastly larger and more transformative market than all specialized robotic applications combined.¹² The logic proceeds in a clear sequence. Specialized robots, such as those found in automated warehouses or on manufacturing assembly lines, are mature, efficient, and highly valuable within their structured, controlled domains.¹⁴ Their market, while substantial, is ultimately finite. The human world—our homes, offices, hospitals, and public squares—represents the largest, most valuable, and most unstructured environment of all. To re-engineer this entire world to accommodate a myriad of specialized robots is economically and practically infeasible. Therefore, the most scalable and logical path forward is to develop a robot that adapts to the

existing human world. The humanoid form, despite its immense technical difficulty, is the natural endpoint of this strategic calculus. The events of this week, from Beijing to Buffalo, demonstrate a collective, global commitment to aggressively pursue this more difficult path, signaling a widespread belief that the formidable technical obstacles are, for the first time, genuinely surmountable.

This technological race is unfolding within a distinct geopolitical context. The 2025 World Robot Conference was not just a commercial event but a carefully orchestrated exercise in technological statecraft. China, through high-level government participation and the co-hosting of the event by state-affiliated bodies, used the conference as a platform to project its rapidly advancing capabilities to a global audience, directly challenging the long-held US dominance in the vanguard of AI and robotics.² The event's narrative was meticulously crafted, with international experts invited to praise China's "miracle" progress and its "unmatched" manufacturing ecosystem.² This narrative deliberately positions China's unique capacity for rapid, large-scale deployment, iteration, and real-world data collection as a decisive competitive advantage over Western development models. The opening of a

dedicated "Robot Mall" and the announcement of the inaugural World Humanoid Robot Games further amplify this message, serving as both a commercial strategy and an international demonstration of technological ambition.¹ This concerted effort makes it clear that the race to develop and deploy humanoid robots has become a flagship battleground in the broader technological rivalry between the United States and China, with the "Rise of the Machines" being as much a matter of national prestige and economic power as it is of scientific progress.

Major Breakthroughs: The Foundational Pillars of Embodied Intelligence

The visible progress in humanoid robotics showcased this week rests upon a foundation of less visible, but arguably more significant, breakthroughs in core enabling technologies. These are not merely incremental improvements but fundamental advances in how robots sense, reason about, and interact with the world. Three such pillars stand out from the past seven days: a revolutionary leap in haptic sensing that brings robotic touch closer to human capability, the release of a powerful AI reasoning engine designed to serve as the "brain" for physical agents, and the debut of a new class of humanoid designed from the ground up for empathetic human-robot interaction. Together, these developments in sensing, reasoning, and interaction constitute the foundational advances that will enable the next generation of humanoid capabilities.

The Sentient Touch: A Leap in Haptic Technology

For decades, one of the most significant barriers to creating truly dexterous robots has been the absence of a sophisticated, human-like sense of touch. While vision systems have advanced dramatically, the ability for a robot to feel an object's texture, perceive pressure, and, most critically, detect the subtle onset of a slip has remained elusive. This deficiency has limited robotic manipulation to clumsy, pre-programmed movements, ill-suited for the delicate and adaptive handling of objects in the real world. A breakthrough announced by researchers at the University at Buffalo, published in *Nature Communications*, directly addresses this grand challenge with the

development of a novel electronic textile, or E-textile, that mimics the function of human skin with remarkable fidelity.¹¹

The technology represents a paradigm shift away from complex and expensive sensor arrays. Instead, it leverages a physical phenomenon known as the tribovoltaic effect, where friction between two materials generates a direct-current (DC) electrical signal.¹¹ This simple yet elegant mechanism allows the flexible E-textile to function as a highly sensitive skin, uniquely capable of detecting not only the static pressure of a grip but also the dynamic shear forces associated with an object beginning to slip. This dual-mode sensing is what gives humans their effortless ability to handle objects of varying weight, fragility, and texture without dropping or crushing them.

The performance of the UB sensor system is, crucially, on par with its biological counterpart. The researchers measured the system's response time to be between 0.76 milliseconds and 38 milliseconds, a range that comfortably falls within the 1 to 50 millisecond reaction time of human touch receptors.¹¹ This speed is not merely an academic benchmark; it is the critical threshold required for real-time, closed-loop feedback control. To prove this, the research team integrated the E-textile onto a pair of 3D-printed robotic fingers mounted on a compliant gripper. In a compelling demonstration, the gripper was able to sense the moment a copper weight began to be pulled from its grasp and immediately tightened its grip to prevent the object from falling.¹¹ This experiment provides concrete evidence that the technology can enable the kind of dynamic grip adjustment that is a prerequisite for any advanced manipulation task, from assembling delicate electronics to performing robotic surgery. The potential applications are vast, spanning manufacturing, logistics, prosthetics, and collaborative robotics, promising to endow machines with the nuanced sense of touch they have so long been missing. The team plans to further enhance the system by integrating it with reinforcement learning, a form of AI that could allow the robot to learn even more dexterous and adaptive behaviors through experience.¹¹

The Reasoning Engine: NVIDIA's Cosmos and the Future of Physical AI

If the E-textile from the University at Buffalo represents a breakthrough in peripheral sensing, the announcements from NVIDIA at the SIGGRAPH 2025 conference represent a breakthrough at the central core of the robotic "brain." NVIDIA unveiled a suite of new software libraries and AI models under the umbrella of "Physical AI," a term the company uses to describe the underlying technology for modern robotics

and autonomous systems. The centerpiece of this announcement is **Cosmos Reason**, a new model that can be understood as a foundational reasoning engine—an operating system for physical reality—designed to give embodied agents a human-like understanding of the world.¹⁰

Cosmos Reason is an open, customizable, 7-billion-parameter Vision Language Model (VLM) that has been purpose-built to address the unique challenges of physical AI.¹⁸ Unlike standard VLMs that are trained primarily on text and images from the internet, Cosmos Reason is designed to excel at understanding how the real world works. It incorporates structured reasoning to comprehend concepts like physics, object permanence, and spatiotemporal relationships.¹⁸ This capability allows the model to move beyond simple object recognition and perform true reasoning, enabling it to plan the sequence of steps an embodied agent, such as a humanoid robot, must take to accomplish a complex goal.¹⁰ For example, it can understand that to retrieve a milk carton from a refrigerator, a robot must first navigate to the kitchen, identify the refrigerator, open its door, locate the carton, grasp it, and then close the door.

This release is strategically significant because it is not a standalone product but the core of a comprehensive, end-to-end development ecosystem. Cosmos Reason is designed to integrate seamlessly with NVIDIA's other platforms: **NVIDIA Omniverse** for creating physically accurate simulations for robot training, **NVIDIA Metropolis** for advanced vision AI, and a new generation of hardware including **NVIDIA RTX PRO Blackwell Servers** and **NVIDIA DGX Cloud** services to run the demanding computational workloads.¹⁰ This integrated stack provides developers with a complete toolkit for building, training, and deploying intelligent robots. The model's applications are multifaceted: it can serve as the high-level reasoning backbone for a Vision-Language-Action (VLA) model, it can be used to automatically critique and caption vast datasets of synthetic training data generated in Omniverse, and it can power runtime visual AI agents that require spatial-temporal understanding of physical operations, such as in a factory or warehouse.¹⁸ The immediate real-world relevance of this technology is underscored by the fact that industry leaders are already exploring its use. Uber, for instance, is working with Cosmos Reason to analyze complex autonomous vehicle behaviors, such as scenarios involving pedestrians, to perform quality analysis and inform the development of its driving policies.¹⁸

The Empathetic Counterpart: Fourier's GR-3 and the HRI Frontier

While much of the industry focuses on performance metrics like speed and payload, Chinese robotics firm Fourier Intelligence unveiled a humanoid robot that represents a different, and potentially equally important, evolutionary path. The **Fourier GR-3 "Care-bot"** is not just a new piece of hardware; it is the physical manifestation of a new design philosophy centered on Human-Robot Interaction (HRI) and the deliberate engineering of empathy and social connection.⁸ The company's stated goal is to create a robot that is "designed to care, not just to work," a mission that has profound implications for the future of service robotics in human-centric environments like hospitals, nursing homes, and eventually, private residences.⁸

This philosophy is immediately apparent in the robot's physical design. The GR-3 intentionally departs from the cold, industrial aesthetic of many of its contemporaries. It features warm, neutral tones, soft-touch cushioned surfaces, and premium automotive upholstery to create a friendly, non-threatening, and approachable presence.⁸ Standing at 1.65 meters tall and weighing 71 kilograms, the robot is engineered for balanced, fluid, and naturalistic movement, with up to 55 degrees of freedom and an animated facial interface designed to foster user comfort and trust.²¹ It is also designed for practicality, equipped with a hot-swappable battery system to support continuous, around-the-clock operation in real-world settings.²¹

The core innovation of the GR-3, however, lies in its advanced cognitive and interactive architecture. The robot is equipped with Fourier's proprietary **Full-Perception Multimodal Interaction System**, which integrates vision (from RGB and structured-light cameras), audio (from a four-microphone array enabling sound source localization), and tactile feedback (from 31 distributed pressure sensors) into a single, unified real-time emotional processing engine.⁸ This allows the robot to perceive and respond to human social cues in a remarkably sophisticated manner.

The most significant technological advance is what Fourier calls its **"dual-path response architecture,"** a hybrid brain that mimics human cognitive processes.²¹ This architecture combines:

1. **"Fast Thinking"**: A rapid, rule-based control system that allows for immediate, reflexive actions. For example, it can process tactile input from its pressure sensors to generate a quick, lifelike social response to being touched.
2. **"Slow Thinking"**: A deep contextual understanding system powered by a large language model (LLM). This path enables complex, situational dialogue and allows the robot to interpret a user's emotion and intent, generating thoughtful

and contextually appropriate responses.

This dual-path system represents a crucial step toward bridging the gap between simple perception and genuine connection, enabling the GR-3 to function as a social companion, a service assistant, or a comforting presence for children and older adults.⁸

The convergence of these distinct breakthroughs in sensing, reasoning, and interaction creates a powerful, self-reinforcing cycle of innovation. Advanced sensors, like the University at Buffalo's E-textile, generate richer, more nuanced data streams about the physical world. This high-fidelity data, in turn, provides the ideal fuel for training more sophisticated AI models, particularly high-level reasoning engines like NVIDIA's Cosmos Reason and the end-to-end VLA models that are becoming the industry standard. A more capable reasoning engine can then make more intelligent use of this sensory input, moving from simply detecting a slip to understanding its implications—for instance, recognizing that a slipping object is fragile and therefore requires a delicate, rather than a forceful, grip adjustment. This improved control and understanding enables more complex and subtle physical interactions, like those envisioned for the Fourier GR-3. These sophisticated interactions, in turn, generate more complex and valuable training data, feeding the next cycle of learning. This virtuous loop—better sensors leading to better data, which leads to better AI, which enables better actions, which generates more complex data—is poised to become a primary engine of exponential progress in the field.

Simultaneously, the contrast between the performance-driven industrial designs seen at the WRC and the interaction-focused philosophy of the Fourier GR-3 reveals a fundamental divergence in market strategy. This is not simply a matter of different target applications; it reflects two competing hypotheses about how to achieve the mass adoption of humanoid robots. The first hypothesis, embodied by powerful machines like the EngineAI T800, is "performance-first." It posits that the most direct path to market adoption is by demonstrating clear, quantifiable utility in high-value tasks, such as heavy lifting or high-speed manufacturing. In this model, social acceptance is expected to *follow* proven utility. The primary challenge is one of engineering and AI control. The second hypothesis, embodied by the GR-3 "Care-bot," is "interaction-first." It argues that for robots to be accepted in sensitive, human-centric environments like hospitals and homes, they must first be socially accepted, trusted, and perceived as safe. By prioritizing HRI and psychological safety, these robots can achieve social integration first, with their utility being expanded and upgraded over time. The primary challenge here is one of social science, design, and HRI. This schism indicates that the industry is hedging its bets, simultaneously

exploring two very different paths to the same ultimate goal. The relative success of these early, pioneering models will provide critical data on which approach—brute-force utility or social integration—is the more viable strategy for breaking humanoid robots out of the lab and into the mainstream.

Demonstrations and Prototypes: The World Robot Conference Nexus

While foundational breakthroughs provide the building blocks, it is their integration into functional prototypes and their demonstration in practical scenarios that signals true progress. The 2025 World Robot Conference (WRC) in Beijing served as a living laboratory for this integration, transforming from a mere trade show into a global stage for tangible capability showcases. The demonstrations on display were not simply technological novelties; they were carefully curated presentations of concrete progress in solving the core challenges of robotics, including dynamic locomotion, long-horizon task planning, and real-world commercialization strategies. From acrobatic combat displays to complex domestic chores and an entirely new retail concept, the WRC provided a comprehensive snapshot of the state of the art in humanoid robotics.

Case Studies in Advanced Humanoid Capabilities

The conference floor was dominated by a new generation of humanoid robots exhibiting capabilities that were, until recently, the exclusive domain of research labs and CGI. Several companies stood out with demonstrations that highlighted significant advancements in hardware design, dynamic control, and AI-driven autonomy.

EngineAI's Industrial and Agile Platforms: Chinese firm EngineAI made a significant impact with the debut of two distinct humanoid platforms, targeting different segments of the emerging market.⁷ The

T800 represents a major step towards industrial-grade humanoids. Making its global

debut, this heavy-duty, full-sized robot stands at an imposing 1.85 meters and weighs 85 kilograms. Its construction, featuring an aluminum alloy exoskeleton, 41 high-degree-of-freedom (DoF) joints, and a solid-state battery, is explicitly designed for robust performance in demanding industrial environments. Equipped with a multi-sensor fusion system for real-time environmental processing, the T800 is positioned to take on heavy-load tasks, setting a new benchmark for strength and endurance in humanoid robotics.⁷ In a clear signal of its combat and performance capabilities, the T800 is scheduled to compete in a "Mecha King" combat tournament later in the year.⁷

In parallel, EngineAI showcased its **PM01** humanoid, a platform focused on agility and dynamic performance. Live demonstrations featured the PM01 executing high-speed running and, critically, performing autonomous fall recovery—a complex maneuver that requires sophisticated balance algorithms and rapid motor control.⁷ With a more fluid and natural gait enhanced by advanced algorithms, the PM01 demonstrated its potential for navigating complex spaces. Strategically, EngineAI is positioning the PM01 as an open-source platform, a move designed to cultivate a global community of developers and accelerate innovation in a collaborative ecosystem.⁷ The company also teased a forthcoming lightweight model, the SA02, with a target price of US\$5,300, indicating a clear strategy to make humanoid technology more accessible to researchers and enthusiasts.⁷

Unitree's Dynamic Athletes: Unitree Robotics, already known for its agile quadrupeds, captured significant attention with demonstrations that pushed the boundaries of dynamic humanoid performance.⁹ The most spectacular of these was a series of robot boxing matches between two of its upgraded

G1 humanoid robots, a spectacle that drew direct comparisons to the science fiction film *Real Steel*.⁹ These were not slow, pre-programmed movements; the robots engaged in fluid combat, delivering punches and kicks, defending against attacks, and demonstrating remarkable resilience by getting back up after being knocked down.⁴

This "Real Steel" moment was a powerful showcase of several key technical achievements. The ability to execute these maneuvers requires precise coordination across multiple joints, extremely fast reaction times, and advanced dynamic balance algorithms to maintain stability while shifting weight and absorbing impacts.⁴ A Unitree marketing manager explicitly stated that the boxing demonstration was more than just entertainment; it provided invaluable experience for developers to refine control policies under the stress of real-world external forces.⁴ The lightning-fast kicks and punches were powered by reinforcement learning (RL) and motion-capture

training systems, where the robot learns complex skills through trial and error in a simulated environment before deploying them on the physical hardware.²⁵ This linkage between advanced AI training methods and spectacular physical performance was a recurring theme at the conference.

Galaxea AI's Domestic Dexterity: While Unitree focused on dynamic athleticism, the startup Galaxea AI provided a compelling demonstration of a different, but equally challenging, capability: autonomous execution of a complex domestic task. At their exhibition, a wheeled humanoid robot, the **R1 Lite**, successfully made a bed in a randomly disordered room in response to a simple verbal command: "Please make the bed".⁴ This demonstration was a prime example of a Vision-Language-Action (VLA) model being successfully deployed to solve a long-horizon task in an unstructured environment.

The task's complexity should not be underestimated. It required the robot to perform a multi-step sequence of actions: visually perceive the state of the messy bed, understand the abstract goal conveyed by the verbal command, and then generate and execute a long sequence of precise physical movements—grasping the edge of the duvet, pulling it straight, smoothing it out, and arranging the pillows. According to Galaxea AI's chief scientist, this was achieved using a single, end-to-end VLA model that directly maps visual perception and language input to sophisticated joint control, showcasing the power of this emerging AI architecture for practical, real-world applications.⁴

Robot Model	Company	Height	Weight	Degrees of Freedom (DoF)	Key Feature/Design Philosophy	Intended Application
EngineAI T800	EngineAI	1.85 m	85 kg	41	Heavy-duty, aluminum alloy exoskeleton, high endurance	Industrial, Heavy-Load Tasks
EngineAI PM01	EngineAI	1.39 m	40 kg	23-24	High-speed running, autonomous	General Purpose, R&D

					us fall recovery, open-source	
Fourier GR-3	Fourier Intelligence	1.65 m	71 kg	up to 55	Empathetic design, "Care-bot" concept, dual-path AI	Healthcare, Eldercare, Social Companionship
Galaxea R1 Lite	Galaxea AI	1.28 m	55 kg	23	Wheeled base, dual-arm manipulation, VLA-driven	Domestic Chores, Service, R&D
Unitree G1	Unitree Robotics	~1.4 m	~40-50 kg	23-43	Extreme agility, reinforcement learning, combat-ready	Entertainment, R&D, General Purpose

The Robot Mall: A Paradigm Shift in Commercialization

Perhaps the most ambitious and strategically significant development of the week was the opening of the world's first dedicated "Robot Mall" in Beijing's high-tech E-Town district.¹ This initiative represents a radical departure from the traditional business-to-business (B2B) sales model that has long dominated the robotics industry, marking a deliberate and large-scale experiment in creating a direct-to-consumer (B2C) market. The mall's very existence is a statement of intent, aiming to normalize human-robot interaction and accelerate cultural acceptance by positioning robots as everyday products and companions rather than futuristic

novelties or industrial tools.¹

The mall's operational strategy is explicitly modeled on the mature automotive industry, adopting a "4S" model that integrates **Sales, Service, Spare Parts, and customer Surveys** all under one roof.¹ This ecosystem approach is designed to build consumer confidence and provide the long-term support necessary for mainstream adoption. The product range on display is vast, intended to capture the entire market spectrum from casual consumers to high-end enterprise clients. Offerings range from simple robotic gadgets priced at approximately US

278to advanced, life-size humanoid units, such as an Albert Einstein replica, valued at nearly US\$97,000.¹

Crucially, the Robot Mall is designed as an immersive, "living showcase" where robots are not just static products on a shelf but are also the active staff.³¹ Visitors can be served by robotic waiters and chefs in the restaurant, interact with robotic baristas in cafes, and receive product information from AI-powered sales assistants.¹ Entertainment zones feature robot competitions, including football and track events, further blurring the line between product demonstration and interactive experience. The timing of the mall's launch was strategically synchronized with the World Robot Conference to maximize international media attention and firmly establish the initiative as a key pillar of China's national robotics strategy.¹

The highly dynamic and interactive demonstrations seen at the WRC, particularly the robot boxing matches, serve a purpose that extends far beyond public entertainment. They function as invaluable exercises in real-world data collection, specifically for bridging the notoriously difficult "sim-to-real" gap. The majority of advanced robotic control policies are developed and trained within simulated environments.²⁹ However, these simulations are always imperfect approximations of real-world physics, dynamics, and sensor noise. An unpredictable, high-impact activity like a boxing match generates vast quantities of crucial data on how the robot's hardware, sensors, and control algorithms actually perform under extreme stress. As confirmed by a Unitree manager, this type of experience is essential for developers to understand and improve multi-joint coordination, reaction time, and balance recovery when subjected to powerful external forces.⁴ Therefore, these public spectacles are, in effect, critical, real-world validation and data-gathering experiments. They provide the high-quality, ground-truth data needed to refine the simulation models, which in turn leads to the training of more robust and capable AI policies.

In a similar vein, the Robot Mall's primary function may not be immediate retail

revenue, but rather to operate as a massive, continuous social experiment. It can be viewed as a large-scale "Social Turing Test" designed to gather unprecedented data on public acceptance and human-robot interaction. The long-term success of service robots hinges on their ability to operate safely, predictably, and acceptably around untrained members of the public—a challenge that is as much about social science and psychology as it is about engineering. A controlled laboratory setting cannot replicate the chaotic and unpredictable nature of a public space. The Robot Mall creates a semi-structured but real-world environment where companies can observe thousands of daily, unscripted interactions between ordinary people and a wide variety of robotic platforms.¹ The "Survey" component of the 4S model represents explicit data collection on user experience and sentiment. The implicit data, however, may be even more valuable. By observing which robots people are drawn to, which they avoid, what types of commands they issue, and how they react to robotic errors or unexpected behaviors, companies can gather a wealth of information. This data is invaluable for refining every aspect of robot design, from physical aesthetics and vocal tonality to approach speed, error-handling routines, and overall interaction logic, effectively running a continuous, at-scale experiment to de-risk the future mass deployment of robots in society.

AI Integration: The Architecture of Embodied Action

The impressive hardware and dynamic demonstrations of the past week are animated by an increasingly sophisticated and standardized architecture of artificial intelligence. The progress on display is a direct result of the convergence of foundational AI models, specialized learning architectures like Vision-Language-Action (VLA) models, and advanced training methodologies. This synergy is creating the "brains" and "nervous systems" capable of translating high-level human intent into precise, low-level physical action. An analysis of the week's developments reveals an emerging software stack for embodied intelligence, mapping a clear path from abstract reasoning to real-world control.

The VLA Revolution: Unifying Perception, Language, and Control

Across both industrial showcases and academic papers, the Vision-Language-Action (VLA) model emerged as the dominant architectural paradigm for endowing robots with general-purpose capabilities. A VLA is a type of multimodal foundation model designed to unify three critical functions—visual perception, natural language understanding, and physical action—within a single framework.³⁴ In practice, this is typically achieved by taking a powerful, pre-trained Vision-Language Model (VLM) and fine-tuning it on large datasets of robotic data that pair visual observations and language instructions with corresponding robot trajectories.³⁴ The result is a model that can watch what is happening, listen to a command, and directly output the low-level control signals needed to execute the requested task.

The most compelling practical application of this architecture was Galaxea AI's R1 Lite robot autonomously making a bed.⁴ This demonstration was a textbook example of a VLA in action. The robot's single end-to-end model processed the visual input of the messy room and the language input of the command "Please make the bed," and from that, generated the entire long-horizon sequence of sophisticated joint movements required to complete the chore.⁴ This ability to translate abstract human goals into concrete physical action is the core promise of the VLA revolution.

The prominence of this approach is not limited to a single company. Leading US-based humanoid developer Figure AI also identifies its core software, **Helix**, as a VLA model, indicating broad industry adoption of this paradigm.³⁶ Concurrently, the academic community is pushing the architecture forward with new research published in the last week. One paper introduces

UniVLA, a novel approach that represents vision, language, and action signals as unified discrete tokens within a single autoregressive framework, allowing it to learn from large-scale video data more effectively.³⁷ Another paper proposes

TriVLA, which uses a triple-system architecture to structure the control problem.³⁸ This vibrant activity across both industry and academia solidifies the VLA as the go-to architecture for building the next generation of intelligent, generalist robots.

The Learning Curve: Advances in Locomotion and Skill Acquisition

Underpinning the VLA architecture are the fundamental learning methodologies used to train the control policies, especially for complex skills like dynamic locomotion. The

demonstrations of high-speed running and combat at the WRC were a direct result of advances in how robots learn to move. Reinforcement Learning (RL) has become the key technique for acquiring these dynamic skills. Companies like Unitree explicitly use RL, often combined with motion-capture data, to train their G1 combat robots.²⁵ In this process, the robot learns through millions of trials in a simulated environment, receiving "rewards" for successful actions, gradually discovering control policies that allow for fluid, agile, and robust movement.

Beyond established RL techniques, new academic research from the past week points toward even more powerful approaches for tackling the grand challenge of locomotion on complex, real-world terrain. A notable paper posted to arXiv introduces **World Model Reconstruction (WMR)**, an end-to-end RL framework designed specifically for "blind" humanoid locomotion—that is, walking without relying on exteroceptive sensors like cameras or LiDAR to see the upcoming terrain.³⁹ The key innovation in WMR is the introduction of an "estimator" network that is trained to explicitly reconstruct the state of the world (e.g., the terrain's slope, roughness, or slipperiness) purely from the robot's own internal sensor histories (e.g., joint positions, velocities, and IMU data). This reconstructed world state is then fed to the locomotion policy, which learns to map this estimated state to actions. The authors introduce a clever "gradient cutoff" mechanism between the estimator and the policy during training. This ensures the estimator focuses solely on achieving the most accurate world reconstruction possible, independent of the policy's immediate needs, which significantly improves the quality of the estimation and the robustness of the final controller. The WMR framework was evaluated in simulations of rough, deformable, and slippery surfaces, demonstrating strong adaptability and a promising path toward navigating challenging real-world environments.³⁹

Another groundbreaking academic paper from the past week proposes a radical reformulation of the robot control problem, casting it as a **"next token prediction"** task, directly analogous to how Large Language Models (LLMs) like GPT work.⁴⁰ In this approach, all of the robot's multimodal sensorimotor data—joint encoder readings, IMU data, motor commands, etc.—are "tokenized" into a unified sequence. A large causal transformer model is then trained to simply predict the next token in the sequence. By modeling the entire joint distribution of sensory inputs and motor outputs, rather than just conditionally predicting actions, the model learns a much richer and more holistic representation of the robot's interaction with the world. The results are remarkable: after being trained on a relatively small dataset of only 27 hours of walking data (from prior policies, motion capture, and even YouTube videos of humans), this generative model enabled a real, full-sized humanoid robot to walk,

zero-shot, in the streets of San Francisco and generalize to commands it had never seen during training, such as walking backward.⁴⁰ This suggests a powerful new path toward learning complex real-world control tasks by leveraging the proven power of generative modeling.

Model/Architecture	Developer/Origin	Type	Key Feature	Role in Robotics Stack
Cosmos Reason	NVIDIA	Reasoning Vision Language Model (VLM)	Understands physics, common sense; plans high-level tasks.	Layer 1: High-Level Reasoning & Planning
Vision-Language-Action (VLA)	Industry-Wide (Galaxea, Figure AI, etc.)	End-to-End Control Model	Translates vision/language commands into low-level actions.	Layer 2: Perception-to-Action Translation
World Model Reconstruction (WMR)	Academia (arXiv:2502.16230)	Reinforcement Learning Framework	Explicitly reconstructs world state from proprioceptive sensors for robust blind locomotion.	Layer 3: Low-Level Control & State Estimation
Generative Sensorimotor Models	Academia (NeurIPS 2024)	Causal Transformer	Treats control as a "next token prediction" problem, enabling zero-shot generalization.	Layer 3: Low-Level Control & Policy Learning

A clear, hierarchical AI architecture is rapidly becoming the de facto standard across the robotics industry. This emerging "stack" for embodied intelligence is modular, allowing for specialized development at each layer while ensuring cohesive overall performance. At the top of the stack, in **Layer 1**, sits a large, foundational reasoning engine like NVIDIA's Cosmos Reason.¹⁸ This layer is responsible for high-level, "slow-path" thinking. It parses complex or ambiguous human commands, applies

common-sense knowledge about the physical world, and breaks down abstract goals into a logical sequence of sub-tasks. For example, it would reason that the command "get me a drink from the fridge" requires the sub-tasks of navigating to the kitchen, identifying the refrigerator, opening the door, and so on.

The output of the reasoning engine is then passed down to **Layer 2**, the Perception-to-Action Translation layer. This layer is dominated by Vision-Language-Action (VLA) models.⁴ The VLA's job is to take the concrete sub-tasks from the reasoning engine (e.g., "open the refrigerator door") and translate them into a sequence of low-level, executable actions (e.g., "move joint X to angle Y, orient gripper Z, apply force F"). It is the bridge between abstract intent and physical motion.

Finally, **Layer 3** consists of the Low-Level Control and Training methodologies. This is where techniques like Reinforcement Learning (RL) and Imitation Learning (IL) are used to train the VLA and the underlying motor control policies.³³ These techniques are used to optimize specific skills, such as stable walking, precise grasping, or dynamic balance recovery. This training is most often conducted in simulation, where millions of trials can be run safely and efficiently. This modular, three-layer stack—Reasoning, Translation, and Control—allows for focused innovation at each level and is a key enabler of the rapid, system-wide progress currently being observed.

This reliance on AI models brings another critical factor to the forefront: data. The success of every layer in this stack is entirely dependent on the availability of massive, diverse, and high-quality training datasets. The primary bottleneck in robotics is shifting from being purely algorithmic to one of data acquisition, curation, and processing. VLA models, for instance, require vast datasets that pair synchronized visual, linguistic, and action data.³⁴ Collecting this data in the physical world is prohibitively slow, expensive, and often dangerous. Consequently, simulation has become the essential "refinery" for creating this data at scale, with platforms like NVIDIA's Omniverse serving as the virtual training grounds.¹⁰ However, raw simulation data is plagued by the sim-to-real gap. This is where a new ecosystem of data-centric tools becomes vital. Models like NVIDIA's Cosmos Transfer-2 are designed specifically to augment simulation data, adding photorealistic textures and varied lighting conditions to improve its diversity and fidelity.¹⁹ Reasoning models like Cosmos Reason are then used to "critique" this synthetic data, automatically filtering and labeling it to ensure its quality and relevance for training.¹⁸ Finally, the real-world demonstrations, such as the robot boxing matches, provide the crucial ground-truth data needed to validate the simulations and close the loop. This creates a powerful

data flywheel: simulated data is used to train a policy, the policy is tested in the real world, the real-world performance data is used to improve the simulation, and the improved simulation generates even better data for the next generation of policies. Companies that master this Sim-to-Real-to-Sim data pipeline will possess a formidable and compounding competitive advantage.

Comparative Advances: The Broader Robotics Ecosystem

To fully grasp the strategic significance of the recent surge in humanoid robotics, it is essential to contextualize these developments within the broader landscape of robotic innovation. While humanoids captured the spotlight, the past week also saw significant advances in non-humanoid systems, particularly in advanced legged platforms and large-scale industrial automation. Analyzing these parallel developments serves as a crucial benchmark, highlighting both the unique, formidable challenges faced by the humanoid form factor and the immense, game-changing potential that drives its pursuit.

Progress in non-humanoid legged systems, primarily quadrupeds, continued at a brisk pace, with a focus on enhancing perception and manipulation capabilities. The Swiss company **ANYbotics**, a leader in industrial inspection robots, announced the launch of a new manipulation payload for its ANYmal research quadruped at the ICRA 2025 conference.⁴³ Developed in collaboration with Duatic, an ETH Zurich spin-off, this payload features a powerful, torque-controlled robotic arm and an onboard NVIDIA Jetson GPU. This integration is designed to enable advanced mobile manipulation tasks, such as agile tool use and object handling, in the complex and unstructured environments found in industrial facilities.⁴³ Similarly,

Unitree Robotics unveiled its new A2 quadruped, which features upgraded perception through the integration of both front and rear LiDAR sensors, enhancing its environmental awareness and navigation capabilities.⁴⁴

EngineAI also entered this space with the debut of its JS01, a high-mobility quadruped built for rugged terrain, featuring bionic mechanics and adaptive shock-absorbing joints to ensure stable performance on uneven surfaces.⁷

In the realm of industrial and logistics automation, the sheer scale of deployment remains a key differentiator. E-commerce and logistics giant **Amazon** announced a

major milestone: the deployment of its one-millionth robot across its global fulfillment network.¹⁵ More strategically important than the number, however, was the announcement of

DeepFleet, a new generative AI foundation model designed specifically for fleet management. DeepFleet acts as an "intelligent traffic management system" for Amazon's entire robotic fleet, using AI to coordinate the movement of thousands of robots within a facility, optimizing their paths to reduce congestion and improve overall efficiency. This system is a powerful example of how AI is being used to optimize logistics at an unprecedented scale.¹⁵

The field of autonomous drones is also experiencing a period of rapid innovation, driven by advances in AI that are enabling new levels of autonomy and collaborative behavior. Key trends include the development of **autonomous swarm intelligence**, where multiple drones can operate as a single, coordinated unit to perform complex tasks like search and rescue without direct human intervention.⁴⁵

Drone-in-a-box systems, which are self-contained, weatherproof docking stations that can automatically launch, land, charge, and offload data from a drone, are enabling fully automated, repeatable missions for applications like industrial inspection and security.⁴⁶ These advanced capabilities have attracted significant military interest, with the U.S. Army actively issuing solicitations for autonomous drones capable of advanced teaming and operating in contested environments.⁴⁷

These advances in the broader robotics ecosystem throw the unique challenge and promise of the humanoid form into sharp relief. The progress in non-humanoid systems underscores a fundamental design trade-off between the specialist and the generalist. Robots like Amazon's Proteus warehouse mover or a specialized sewer-cleaning robot are hyper-optimized for a narrow set of tasks within a specific, often highly structured, environment.¹⁵ Within that domain, they achieve a level of efficiency and performance that a general-purpose robot is unlikely to match. Humanoid robots, by contrast, are the ultimate generalists. They are being designed to operate across a vast and unpredictable array of tasks within the highly unstructured human world. This makes the engineering and AI problem for humanoids exponentially more difficult. A warehouse robot does not need to climb stairs, open a door, or use a tool it has never seen before; a truly useful humanoid must be able to do all of these things and more. The progress in specialized robotics thus provides a crucial performance benchmark. The strategic goal for humanoids is not necessarily to outperform a specialist in its niche—for example, to be a better box-mover than Amazon's Hercules robot—but to be "good enough" at a thousand different tasks,

making its unparalleled versatility its ultimate killer application.

Despite the different form factors, the underlying technological currents are beginning to converge, leading to a significant cross-pollination of AI and hardware concepts across the entire robotics ecosystem. The AI principles used in Amazon's DeepFleet to manage a swarm of warehouse robots share a common lineage with the swarm intelligence algorithms being developed for autonomous drones.¹⁵ The challenges of perception, planning, and control faced by the manipulation payload on ANYbotics' quadruped are directly relevant to the development of a dexterous humanoid hand and arm.⁴³ At the hardware level, there is a clear convergence on common computing platforms. The powerful and energy-efficient NVIDIA Jetson Orin family of modules, for example, is being used as the onboard "brain" in a wide range of platforms, from the ANYmal quadruped to the humanoid robots developed by EngineAI and Galaxea AI.⁴³ This indicates the emergence of standardized hardware and software components that can be leveraged across different robotic forms. This ecosystem effect is mutually beneficial, as a breakthrough in one domain—such as a more efficient perception algorithm developed for a drone—can often be adapted and transferred to others. The humanoid robot, as the most complex and demanding platform, stands to be the greatest beneficiary of this technological cross-pollination, inheriting and integrating advances from across the entire field of robotics.

Applications and Implications: Charting the Near-Term Future

The convergence of advanced hardware, sophisticated AI, and ambitious commercial strategies detailed in this report is not merely an academic exercise. It is actively charting the course for the real-world deployment of humanoid robots. The breakthroughs and demonstrations of the past week provide a clear, albeit early, roadmap of the most promising near-term applications, the significant challenges that still lie ahead, and the strategic trajectory of the industry over the next 12 to 24 months. Synthesizing these findings offers a forward-looking analysis of the imminent impact of the humanoid ascendancy.

Emerging Deployments and Market Trajectory

The announcements and showcases from the World Robot Conference and other global sources point toward a multi-pronged market entry strategy, with companies targeting several key sectors simultaneously.

Industrial & Logistics: This sector is poised to be the first major beachhead for humanoid robot deployment. The controlled, structured, yet complex nature of modern factories and warehouses makes them an ideal proving ground. The debut of EngineAI's heavy-duty T800, with its focus on payload and endurance, is explicitly aimed at industrial applications.⁷ The general emphasis at the WRC on manufacturing scenarios, such as sorting parts on an assembly line, further reinforces this focus.³ The value proposition is clear and compelling: deploying humanoids to perform the "dull, dirty, and dangerous" jobs that are difficult to fill and often lead to worker injury.² Initial tasks will likely involve material handling, machine tending, and quality inspection, leveraging the humanoid form to work alongside existing equipment without requiring extensive infrastructure changes.

Healthcare & Eldercare: Driven by powerful demographic tailwinds, including aging populations in many developed nations, healthcare and eldercare represent a massive long-term market.¹ Fourier Intelligence's GR-3 "Care-bot" is the most direct and sophisticated play in this space to date.⁸ Its design philosophy, centered on empathetic interaction and safety, is tailored for environments like hospitals and assisted living facilities. Initial applications will likely focus on social companionship, patient monitoring, and basic assistance tasks like delivering medication or fetching items. As the technology matures, more complex applications in mobility support and physical rehabilitation will become feasible.⁸

Home & Consumer Services: The ultimate, albeit most distant, ambition for many companies is to place a humanoid robot in every home. The demonstrations of domestic chores, such as Galaxea AI's robot making a bed, are early but powerful signals of this intent.⁴ The launch of the "Robot Mall" in Beijing is a direct attempt to cultivate a B2C market and gauge consumer interest.¹ While widespread adoption is likely many years away due to cost and safety concerns, the industry is already laying the groundwork for a future where robots assist with cooking, cleaning, and other household tasks.

Entertainment & Public Engagement: A new and rapidly emerging application category is "robosports" and public entertainment. The thrilling robot boxing matches at the WRC and the upcoming inaugural World Humanoid Robot Games in Beijing are

creating a new form of spectator sport.⁹ These events serve a dual purpose: they generate revenue and public interest while also functioning as a high-profile R&D platform, pushing the limits of dynamic control and real-world performance in a competitive setting.

Overcoming the Hurdles: A Sober Look at the Challenges Ahead

Despite the palpable excitement and rapid progress, a sober assessment reveals that formidable technical, economic, and social barriers stand between the current prototypes and widespread, reliable deployment.

Cost and Economic Viability: The price of advanced humanoid robots remains a primary obstacle. While some companies are teasing more accessible price points, such as EngineAI's planned US\$5,300 SA02 model, the most capable platforms still cost tens or hundreds of thousands of dollars.¹ The International Federation of Robotics has rightly questioned whether humanoids can represent an economically viable business case in the near term, especially when compared to existing, highly efficient specialized automation solutions.⁵ Achieving a price point that delivers a clear return on investment for industrial customers and is affordable for consumers is a critical challenge.

AI Complexity and Reliability: As noted by Unitree's CEO, the AI "brain" is still a significant bottleneck, with current capabilities falling short of what the hardware can support.⁴ The challenge of developing AI that is not just capable but also robust, reliable, and safe in the infinite variety of real-world scenarios is immense. Ensuring that a robot can handle unexpected events, recover from errors gracefully, and operate safely around untrained humans is a far greater challenge than demonstrating a specific skill in a controlled environment.³

Hardware and Energy Limitations: Fundamental engineering challenges persist. Battery life is a critical concern for practical, all-day operation, with many current models offering only a few hours of continuous use.⁵² Issues of mechanical wear and tear on complex joints, actuators, and drivetrains will become more pronounced with long-term use. Furthermore, energy efficiency is a key sustainability concern that the industry is actively working to address through innovations like lightweight construction and more efficient gripper designs.⁵

Safety, Ethics, and Social Acceptance: The integration of autonomous, human-sized robots into the fabric of society raises profound safety and ethical questions that are only beginning to be addressed.⁵¹ Public anxiety regarding privacy, especially with sensor-laden robots operating in private homes, is a significant hurdle.⁶ Concerns about job displacement, while often nuanced, are widespread and must be managed through proactive public dialogue and investment in workforce retraining.² Building public trust is as critical to the long-term success of this technology as solving any engineering problem.

Strategic Outlook and Recommendations

The developments of the past week suggest a clear trajectory for the next 12 to 24 months, characterized by intense competition, targeted initial deployments, and a relentless focus on data and AI.

Key Predictions:

1. **The Rise of Platform Wars:** The competition will extend beyond hardware manufacturers. A "platform war" will emerge among the providers of the foundational AI models and simulation environments. Companies like NVIDIA with their comprehensive Cosmos and Omniverse ecosystem will compete with the proprietary models developed by tech giants like Google and OpenAI, and open-source alternatives, to become the dominant "operating system" for embodied AI.
2. **Industrial Beachheads:** The first commercially significant, at-scale deployments of humanoid robots will occur in structured industrial environments. Automotive manufacturing and logistics facilities will be the primary targets, with robots performing repetitive, ergonomically challenging tasks on production lines and in warehouses.
3. **Slower "Care-bot" Adoption:** The market for "care-bots" and social companions will develop more slowly. The higher bars for safety, regulatory approval (especially in medical settings), and social acceptance will necessitate a more cautious and lengthy development and deployment cycle. However, this market ultimately represents a larger long-term opportunity.
4. **Data as the Decisive Asset:** Data will be codified as the most valuable strategic asset in the robotics industry. The ability to effectively collect, curate, and leverage massive datasets—from both high-fidelity simulations and real-world

interactions—will become the single most important competitive differentiator. Companies that master the Sim-to-Real-to-Sim data flywheel will out-innovate their rivals.

Strategic Recommendations for Stakeholders:

- **For Investors:** The investment landscape requires a nuanced approach. Focus should be placed on companies that possess not only impressive hardware but also a clear and credible data strategy and a well-defined beachhead market. General-purpose promises without a concrete plan for initial, high-value deployment should be viewed with skepticism. A less risky but equally potent strategy is to invest in the enabling technology providers—the companies building the critical sensors, actuators, and, most importantly, the foundational AI and simulation platforms that the entire industry will depend on.
- **For R&D Leaders:** The strategic priorities are clear. Development efforts must be centered on creating robust, efficient sim-to-real pipelines to accelerate AI training and validation. Significant investment in Human-Robot Interaction (HRI) research is non-negotiable to address the critical challenges of safety, trust, and social acceptance. To avoid costly reinvention and to leverage the state of the art, R&D teams should actively engage with and build upon open platforms and ecosystems, such as NVIDIA's Cosmos and the growing body of open-source VLA models and academic research.
- **For Policymakers and Regulators:** The time for proactive engagement is now. Governments and regulatory bodies must begin the process of developing clear, consistent, and innovation-friendly frameworks for the testing and deployment of autonomous humanoid robots in both public and private spaces. This includes establishing standards for safety, data privacy, and operational liability. Simultaneously, policymakers should foster a broad and informed public dialogue to address societal concerns about job displacement and ethics, ensuring that the transition to a more automated future is managed responsibly and equitably.

Works cited

1. Robot Mall: China Opens First Store Selling Humanoid Robots | Technology Magazine, accessed August 12, 2025, <https://technologymagazine.com/news/robot-mall-china-opens-first-store-selling-humanoid-robots>
2. Foreign experts optimistic about China's robotics industry, accessed August 12, 2025, https://english.www.gov.cn/news/202508/12/content_WS689a789ec6d0868f4e8f4c29.html
3. Beijing 2025 Robot Conference highlights China's humanoid robotics growth,

- accessed August 12, 2025,
<https://dig.watch/updates/beijing-2025-robot-conference-highlights-chinas-humanoid-robotics-growth>
4. China poised for further humanoid robot innovation, commercialization - Guangming Online, accessed August 12, 2025,
https://en.gmw.cn/2025-08/12/content_38212994.htm
 5. TOP 5 Global Robotics Trends 2025 - International Federation of Robotics, accessed August 12, 2025,
<https://ifr.org/ifr-press-releases/news/top-5-global-robotics-trends-2025>
 6. Next 2-3 years a rare window for Chinese humanoid robots, says American professor long focused on robotics development, accessed August 12, 2025,
<https://www.globaltimes.cn/page/202508/1340537.shtml>
 7. EngineAI Unveils Groundbreaking Innovations at World Robot ..., accessed August 12, 2025,
<https://laotiantimes.com/2025/08/11/engineai-unveils-groundbreaking-innovations-at-world-robot-conference-2025/>
 8. Meet GR-3: Beyond Function, Designed to Care. Fourier to Unveil Its First Care-Centric Humanoid, accessed August 12, 2025,
<https://en.antaraneews.com/news/372677/meet-gr-3-beyond-function-designed-to-care-fourier-to-unveil-its-first-care-centric-humanoid>
 9. 'Real Steel' gets real: Robots fight in boxing match at world robot expo in China, accessed August 12, 2025,
<https://economictimes.indiatimes.com/news/new-updates/real-steel-gets-real-robots-fight-in-boxing-match-at-world-robot-expo-in-china/articleshow/123255270.cms>
 10. NVIDIA Research Shapes Physical AI | NVIDIA Blog, accessed August 12, 2025,
<https://blogs.nvidia.com/blog/physical-ai-research-siggraph-2025/>
 11. Scientists give robots a sense of touch with fabric that mimics human ..., accessed August 12, 2025,
<https://www.buffalo.edu/ubnow/stories/2025/08/robotic-touch.html>
 12. Are the robots finally coming? - Visual and data journalism, accessed August 12, 2025, <https://ig.ft.com/ai-robots/>
 13. Humanoid robots offer disruption and promise. Here's why - The World Economic Forum, accessed August 12, 2025,
<https://www.weforum.org/stories/2025/06/humanoid-robots-offer-disruption-and-promise/>
 14. Robotic Trends in 2025: Innovations Transforming Industries | Robotnik®, accessed August 12, 2025,
<https://robotnik.eu/robotic-trends-in-2025-innovations-transforming-industries/>
 15. Amazon deploys over 1 million robots and launches new AI foundation model, accessed August 12, 2025,
<https://www.aboutamazon.com/news/operations/amazon-million-robots-ai-foundation-model>
 16. World Humanoid Robot Games Are Coming! Set for August 2025 at Beijing's Iconic 'Dual Olympic' Venues - 北京市人民政府, accessed August 12, 2025,

https://english.beijing.gov.cn/beijinginfo/sci/latesttrends/202505/t20250509_4085806.html

17. AI Breakthroughs, Billion-Dollar Bets & Backlash – The Global AI News Roundup (Aug 11–12, 2025) - TS2 Space, accessed August 12, 2025, <https://ts2.tech/en/ai-breakthroughs-billion-dollar-bets-backlash-the-global-ai-news-roundup-aug-11-12-2025/>
18. CrowdStrike, Uber, Zoom Among Industry Pioneers Building Smarter Agents With NVIDIA Nemotron and Cosmos Reasoning Models for Enterprise and Physical AI Applications, accessed August 12, 2025, <https://blogs.nvidia.com/blog/nemotron-cosmos-reasoning-enterprise-physical-ai/>
19. NVIDIA Opens Portals to World of Robotics With New Omniverse Libraries, Cosmos Physical AI Models and AI Computing Infrastructure, accessed August 12, 2025, <https://nvidianews.nvidia.com/news/nvidia-opens-portals-to-world-of-robotics-with-new-omniverse-libraries-cosmos-physical-ai-models-and-ai-computing-infrastructure>
20. Maximize Robotics Performance by Post-Training NVIDIA Cosmos Reason, accessed August 12, 2025, <https://developer.nvidia.com/blog/maximize-robotics-performance-by-post-training-nvidia-cosmos-reason/>
21. Meet GR-3: Beyond Function, Designed to Care. Fourier to Unveil Its ..., accessed August 12, 2025, <https://www.prnewswire.com/news-releases/meet-gr-3-beyond-function-designed-to-care-fourier-to-unveil-its-first-care-centric-humanoid-302523202.html>
22. “We’re Teaching Robots to Feel”: Fourier’s GR-3 Care-Bot Launches Global Uproar as Hospitals, Homes, and Tech Critics Clash Over Empathy in Machines - Rude Baguette, accessed August 12, 2025, <https://www.rudebaguette.com/en/2025/08/were-teaching-robots-to-feel-fouriers-gr-3-care-bot-launches-global-uproar-as-hospitals-homes-and-tech-critics-clash-over-empathy-in-machines/>
23. Fourier Intelligence unveils the “Care-bot” GR-3, its first full-sized companion humanoid robot - TechNode, accessed August 12, 2025, <https://technode.com/2025/08/07/fourier-intelligence-unveils-the-care-bot-gr-3-its-first-full-sized-companion-humanoid-robot/>
24. PM01 - ENGINEAI-HOME, accessed August 12, 2025, https://www.engineai.com.cn/product_fore/
25. Humanoid robots shine at Beijing's World Robot Conference - CGTN, accessed August 12, 2025, <https://news.cgtn.com/news/2025-08-11/Humanoid-robots-shine-at-Beijing-s-World-Robot-Conference-1FKH8BjaXyo/p.html>
26. Robots fight in boxing match at 2025 World Robot Conference - YouTube, accessed August 12, 2025, <https://www.youtube.com/watch?v=sHZClcatoVE>
27. China poised for further humanoid robot innovation, commercialization - Xinhua, accessed August 12, 2025,

- <https://english.news.cn/20250811/41ea4de256c34676a1654f1b1addb710/c.html>
28. China poised for further humanoid robot innovation, commercialization - China.org.cn, accessed August 12, 2025, http://www.china.org.cn/2025-08/12/content_118021338.shtml
 29. Harnessing the Power of Reinforcement Learning: A Guide to Unitree G1's RL Control Routine - RoboStore, accessed August 12, 2025, <https://robostore.com/blogs/news/harnessing-the-power-of-reinforcement-learning-a-guide-to-unitree-g1s-rl-control-routine>
 30. Unitree G1-Boxing Iron Fist King: Awakening ! , accessed August 12, 2025, <https://www.unitree.com/mobile/boxing>
 31. China opens world's first Robot Mall: lifelike robots on sale, AI ..., accessed August 12, 2025, <https://timesofindia.indiatimes.com/world/rest-of-world/china-opens-worlds-first-robot-mall-lifelike-robots-on-sale-ai-powered-machines-run-shops-and-cafes/articleshow/123216740.cms>
 32. TVB News | 8 August 2025 World Robot Conference opened in Beijing today - YouTube, accessed August 12, 2025, <https://www.youtube.com/watch?v=kb33g6jIPnw>
 33. unitreerobotics/unitree_rl_gym - GitHub, accessed August 12, 2025, https://github.com/unitreerobotics/unitree_rl_gym
 34. Vision-language-action model - Wikipedia, accessed August 12, 2025, https://en.wikipedia.org/wiki/Vision-language-action_model
 35. [2507.10672] Vision Language Action Models in Robotic Manipulation: A Systematic Review, accessed August 12, 2025, <https://arxiv.org/abs/2507.10672>
 36. Helix: A Vision-Language-Action Model for Generalist Humanoid Control - Figure AI, accessed August 12, 2025, <https://www.figure.ai/news/helix>
 37. Unified Vision-Language-Action Model - arXiv, accessed August 12, 2025, <https://arxiv.org/html/2506.19850v1>
 38. TriVLA: A Unified Triple-System-Based Unified Vision-Language-Action Model for General Robot Control - arXiv, accessed August 12, 2025, <https://arxiv.org/html/2507.01424v1>
 39. Learning Humanoid Locomotion with World Model Reconstruction - arXiv, accessed August 12, 2025, <https://arxiv.org/html/2502.16230v1>
 40. Humanoid Locomotion as Next Token Prediction - NIPS, accessed August 12, 2025, https://proceedings.neurips.cc/paper_files/paper/2024/file/90afd20dc776bc8849c31d61a0763a0b-Paper-Conference.pdf
 41. Humanoid Locomotion as Next Token Prediction - arXiv, accessed August 12, 2025, <https://arxiv.org/html/2402.19469v1>
 42. OpenVLA: An Open-Source Vision-Language-Action Model, accessed August 12, 2025, <https://openvla.github.io/>
 43. ANYbotics Launch of Manipulation Payload at ICRA 2025, accessed August 12, 2025, <https://www.anybotics.com/news/anybotics-anymal-research-mobile-manipulation/>

44. Unitree launches A2 quadruped equipped with front and rear lidar - The Robot Report, accessed August 12, 2025, <https://www.therobotreport.com/unitree-launches-a2-quadruped-equipped-with-front-and-rear-lidar/>
45. How AI and Automation Are Changing Drone Operations in 2025 - The Dronedesk Blog, accessed August 12, 2025, <https://blog.dronedesk.io/how-ai-and-automation-are-changing-the-game-for-drone-operations-in-2025/>
46. Top UAV & Drone-in-a-Box - The 2025 Tech Trends Every Drone Team Should Know, accessed August 12, 2025, <https://blog.dronedesk.io/and-drone-in-a-box-the-2025-tech-trends-every-drone-team-should-know/>
47. Army issues solicitation for 'launched effects' autonomous drones - DefenseScoop, accessed August 12, 2025, <https://defensescoop.com/2025/08/06/army-launched-effects-solicitation-autonomous-drones/>
48. 9 ways robots are helping humans: Robodogs to magnetic slime | World Economic Forum, accessed August 12, 2025, <https://www.weforum.org/stories/2025/07/robots-ai-help-humans-at-work/>
49. ENGINEAI-HOME, accessed August 12, 2025, https://www.engineai.com.cn/product_one/
50. Jetson AGX Orin for Next-Gen Robotics - NVIDIA, accessed August 12, 2025, <https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/>
51. Call for Papers — IEEE RO-MAN 2025, accessed August 12, 2025, <https://www.ro-man2025.org/contributing/callforpapers>
52. Top 12 Humanoid Robots of 2025, accessed August 12, 2025, <https://humanoidroboticstechnology.com/articles/top-12-humanoid-robots-of-2025/>
53. World Robot Conference unveils future trends of embodied intelligent robots, accessed August 12, 2025, https://english.www.gov.cn/news/202508/09/content_WS68968b86c6d0868f4e8f4b90.html