Humanoid Robots 2026: A Comprehensive Review of Technical Foundations, Market Dynamics, and Research Challenges

Nan Wang, Ali Mohammad Saghiri, and Bohan Fan

***Abstract*—The year 2026 marks a pivotal transition in humanoid robotics, characterized by the evolution from research prototypes to commercially viable products. This paper presents a comprehensive review of leading platforms including Tesla Optimus (Gen 2), Figure 03, Unitree G1/H1, Boston Dynamics Electric Atlas, and Fourier GR-1/GR-2. We examine these systems across hardware architecture, software intelligence, operational capabilities, and economic viability. Our contributions include: (1) a systematic taxonomy of design challenges categorized into mechanical, perceptual, cognitive, and integration domains; (2) a market analysis mapping challenges to platform capabilities; (3) identification of theoretical and practical gaps; and (4) bibliometric analysis of publication trends. Our analysis reveals market bifurcation between high-cost industrial units and low-cost research platforms, with significant opportunities for convergence.**

***Index Terms*—Humanoid robotics, embodied AI, design challenges, market analysis, bibliometrics, comparative analysis.**

I. Introduction

FOR decades, humanoid robots remained confined to research laboratories, characterized by hydraulic actuation, prohibitive costs exceeding $500,000, and limited accessibility outside academic institutions [1]. The contemporary era (2024-present) represents a fundamental paradigm shift defined by three converging trajectories: electrification of actuation systems, commoditization through mass production, and integration of embodied artificial intelligence replacing classical control with neural network-based approaches [2], [3].

The significance of humanoid robotics extends beyond technological curiosity. These systems represent the convergence of mechanical design, control theory, computer vision, and machine learning [4]. Their humanoid form factor reflects the anthropocentric design of human environments, suggesting potential for seamless integration into spaces constructed for human occupancy [5].

This paper addresses critical questions facing researchers and practitioners: What technical foundations underpin current platforms? How do commercial systems compare across dimensions? What design challenges remain, and which platforms address them? What gaps exist between theory and practice? How has research interest evolved, and what trends indicate future directions?

A. Related Work

Prior surveys have examined specific aspects of humanoid robotics. Siciliano and Khatib [2] provide comprehensive coverage of robotic fundamentals. Vukobratovic and Stepanenko [6] established foundational stability theory. Recent work by Ding et al. [7] introduced benchmarking frameworks to evaluate multiple capabilities simultaneously. However, no prior work provides a unified treatment of technical foundations, market analysis, challenge taxonomy, and bibliometric trends. This paper fills that gap.

B. Contributions and Paper Organization

This paper makes four primary contributions:

1) A systematic taxonomy of design challenges in humanoid robotics, organized into mechanical, perceptual, cognitive, and integration categories.
2) A comprehensive market analysis mapping specific challenges to the capabilities of existing commercial platforms.
3) Identification of theoretical and practical gaps requiring research attention, with prioritized recommendations.
4) Bibliometric analysis revealing publication trends, research hotspots, and emerging directions.

The remainder of this paper is organized as follows: Section II presents technical foundations and definitions. Section III analyzes market leaders. Section IV presents the challenge taxonomy. Sections V-VII analyze gaps, bibliometrics, and future directions. Section VIII concludes.

C. Scope and Limitations

This review focuses on commercially available or announced humanoid platforms as of 2025. We exclude research prototypes without clear commercialization paths and platforms with fewer than 20 degrees of freedom. This threshold was selected to focus the review on full-body humanoid systems capable of coordinated locomotion, upper-body manipulation, and multi-joint whole-body control. Platforms below this threshold were often task-specific, partial-body, or minimally articulated systems whose design constraints and

Manuscript received April 24, 2026; revised June 10 and June 17, 2026; accepted June 30, 2026. This article was recommended for publication by Associate Editor Shujin Qin upon evaluation of the reviewers' comments.

N. Wang, A. Saghiri, and B. Fan are with the Department of Computer Science, William Paterson University, Wayne, NJ, USA (e-mail: wangn1@wpunj.edu, saghiria@wpunj.edu, fanb@wpunj.edu).

Corresponding Author: Nan Wang.

control challenges differ substantially from those of full humanoids. While such systems can provide valuable technical contributions, they were excluded to preserve comparability across platforms and to avoid diluting the analysis of full-body humanoid robotics. The analysis is based on publicly available specifications, published research, and industry reports [8]. A key limitation of this review is the reliance on manufacturer-reported specifications and performance claims, many of which could not be independently verified due to the limited availability of standardized third-party benchmarking data. Reported metrics such as payload capacity, runtime, locomotion speed, and manipulation performance may therefore reflect differing testing methodologies, operating conditions, and reporting standards across manufacturers. Consequently, comparisons between platforms should be interpreted cautiously [9]. The review emphasizes reported capabilities, architectural trends, and design trade-offs rather than definitive performance rankings.

II. Technical Foundations

A. Definitions and Classification

Following IEEE RAS definitions [10], a humanoid robot is an autonomous or semi-autonomous system possessing a body plan approximating the human form: torso, two arms, two legs, and a head, with degrees of freedom (DoF) enabling bipedal locomotion and bimanual manipulation. This distinguishes humanoids from quadrupeds, wheeled platforms, and fixed-base manipulators.

B. Kinematic Structure

A minimal humanoid configuration includes 12 DoF for locomotion (6 per leg: hip yaw, roll, pitch; knee pitch; ankle roll, pitch) and 14 DoF for manipulation (7 per arm), totaling 26 DoF excluding head and torso. Advanced platforms such as Figure 02 incorporate additional DoF in the torso and highly articulated hands (16 DoF per hand) [11], making more dexterous manipulation possible.

The Denavit–Hartenberg (DH) convention provides the standard framework for forward kinematics [12, 13]. For an n-DoF robot:

$$T_0^n = {}^0A_1 \, {}^1A_2 \ldots {}^{n-1}A_n = \prod_{i=1}^{n} {}^{i-1}A_i, \tag{1}$$

where ${}^{i-1}A_i$ is the 4×4 homogeneous transformation matrix from frame $\{i-1\}$ to $\{i\}$ [14]. Inverse kinematics typically uses analytical geometric solutions for non-redundant manipulators or the Jacobian pseudoinverse and optimization-based methods for redundant systems [13].

C. Actuation Technologies

For many years, humanoid robots were primarily powered by hydraulic actuation systems, which generate motion through the pressure of an incompressible liquid acting on pistons and rotary actuators. This approach enabled very high power density and dynamic performance, but at the cost of efficiency, noise, leakage, and maintenance complexity [15]. In recent years, advances in high-torque electric motors and low-ratio gearboxes have driven a shift toward fully electric actuation, enabling quieter, more efficient, and more easily mass-produced humanoid platforms [16].

Electric actuation modalities, categorized by their defining characteristics, primarily encompass the following types:

1) Harmonic Drive Actuators: Featuring high gear ratios (50:1–160:1) and zero backlash, these are prevalent in high-precision joints such as shoulders, elbows, and torso.
2) Quasi-Direct Drive (QDD): Employing low gear ratios (<10:1), QDD enables high-bandwidth force control (up to 2 kHz) and backdrivability, making it widely adopted for dynamic load-bearing joints including hips and knees.
3) Planetary Gearboxes: With gear ratios of 3:1–50:1, these provide high efficiency (~95%) for light-load applications such as wrists and fingers.
4) Series Elastic Actuators (SEA): Incorporating intentional compliance between motor and load, SEA facilitates precise force measurement and shock absorption, particularly suited for dexterous manipulation tasks.

D. Locomotion Principles

The Zero Moment Point (ZMP) criterion [6] provides foundational stability: the ZMP must remain within the support polygon (convex hull of contact points) for dynamic stability. The ZMP position is computed as:

$$x_{ZMP} = \frac{\sum m_i(\ddot{z}_i + g)x_i - \sum m_i \ddot{x}_i z_i}{\sum m_i(\ddot{z}_i + g)} \tag{2}$$

$$y_{ZMP} = \frac{\sum m_i(\ddot{z}_i + g)y_i - \sum m_i \ddot{y}_i z_i}{\sum m_i(\ddot{z}_i + g)} \tag{3}$$

where x_{ZMP} and y_{ZMP} denote the coordinates of the Zero Moment Point on the ground plane; m_i is the mass of the i-th link; x_i, y_i, z_i represent the position of the i-th link's center of mass; $\ddot{x}_i, \ddot{y}_i, \ddot{z}_i$ are the corresponding accelerations; and g is the gravitational acceleration. Modern approaches increasingly adopt model-free reinforcement learning [17], training policies in simulation before Sim-to-Real transfer. The proximal policy optimization (PPO) algorithm [18] has shown particular success for locomotion tasks.

E. Perception Systems

Humanoid robots rely on integrated perception systems to estimate their internal states and interpret the surrounding environment. Modern humanoid platforms typically employ a combination of proprioceptive and exteroceptive sensors. Proprioceptive sensors include Inertial Measurement Units (IMUs), joint encoders, and force/torque sensors located at the feet and wrists, which provide information about body motion, joint configuration, and contact forces [19]. Exteroceptive sensing is commonly achieved through vision systems such as RGB cameras, depth sensors, and LiDAR, enabling the robot to perceive objects, terrain, and obstacles in the environment [2]. Recent developments also incorporate distributed tactile sensor arrays that provide pressure and contact feedback across robot surfaces, supporting manipulation and safe physical interaction [20].

To combine heterogeneous sensory information, humanoid robots employ probabilistic state estimation algorithms such as the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF), which fuse proprioceptive and exteroceptive measurements to estimate the robot's dynamic state in real time [21][22]. The estimated state vector typically includes the robot base position and orientation, linear and angular velocities, joint positions and velocities, and contact states.

Beyond internal state estimation, perception systems also support environment understanding. Vision-based modules are widely used for terrain estimation, obstacle detection, and object recognition, while navigation tasks often rely on methods such as Simultaneous Localization and Mapping (SLAM) to construct maps and estimate the robot's pose in unknown environments [23]. Recent advances increasingly integrate deep neural networks for visual perception, improving robustness in complex and unstructured real-world settings [24].

F. Control Architectures

Humanoid control architectures span a spectrum from model-based to learning-based approaches. Model-based methods have traditionally dominated humanoid locomotion and manipulation control. Representative techniques include: (1) Zero Moment Point (ZMP) control for stable walking, originally introduced by Zero Moment Point based locomotion planning [25]; (2) Capture Point control for push recovery and balance stabilization [26]; (3) Model Predictive Control (MPC) for whole-body motion planning under dynamic constraints [27]; and (4) Operational Space Control (OSC) for task-space tracking and force regulation [28].

Learning-based approaches have gained increasing attention with the success of deep reinforcement learning in robotics. Policies are often trained in simulation using physics engines such as MuJoCo, Isaac Gym, or RaiSim, and subsequently transferred to real robots through techniques such as domain randomization and sim-to-real adaptation [29]. Among reinforcement learning algorithms, Proximal Policy Optimization (PPO) has been widely adopted for locomotion policy training due to its stability and relatively good sample efficiency [18].

More recently, hybrid control architectures have emerged that combine model-based and learning-based components. For instance, model-based controllers may handle low-frequency balance stabilization and dynamic feasibility constraints, while learned policies are responsible for high-level decision making or adaptive foot placement. Such hybrid designs leverage the robustness and interpretability of model-based control while exploiting the adaptability of learning-based methods [30].

III. MARKET LEADER ANALYSIS

A. Tesla Optimus

Tesla Optimus represents a major effort toward a general-purpose, mass-deployable humanoid robot platform, originally announced as Tesla Bot at the 2021 AI Day and later realized through the Optimus Gen 2 and Gen 3 prototypes [31, 32]. The system is designed to leverage Tesla's self-driving-style AI stack, operating on a Tesla-designed SoC and using vision and proprioceptive data to enable end-to-end learning for navigation and manipulation tasks [31, 32]. Optimus stands $\sim$173 cm tall, weighs around 57 kg, and supports a payload capacity of up to 20 kg. Optimus Gen 2 is equipped with 28 degrees of freedom, including highly dexterous hands with 11 degrees of freedom, enabling it to perform a range of industrial and domestic tasks such as object handling, machine tending, and basic logistics [31, 33]. Its actuation architecture provides agile mobility, with Optimus Gen 2 capable of walking at several kilometers per hour, while the upcoming Optimus Gen 3 is reported to further refine dynamics, balance, and dexterity for sustained operation in factory and household environments [33, 34]. Tesla emphasizes high-volume manufacturability, with plans to start Optimus Gen 3 production in Summer 2026 and scale toward millions of units per year through dedicated production lines at Fremont and Gigafactory Texas, positioning Optimus as a foundational component of Tesla's broader AI-driven industrial automation vision [34, 35]. These developments collectively illustrate the transition of Optimus from a research-oriented prototype into a commercially scalable humanoid platform for real-world deployment [31, 34].

B. Figure 03

Figure 03 represents a significant step toward scalable, general-purpose humanoid robotics, designed to operate in everyday environments while enabling end-to-end visuomotor learning [36, 37]. The system is co-designed with Figure's proprietary vision–language–action model, Helix, which runs on an on-board dual-GPU platform and supports pixel-to-action learning through a redeployable pipeline of human demonstrations [37]. Figure 03 features a redesigned sensory suite with a multi-camera visual system that offers approximately twice the frame rate, one-quarter the latency, and a 60% wider field of view per camera compared to earlier platforms, along with palm-mounted cameras for close-range visual feedback in cluttered environments [36, 38]. Its compliant, tactilely enhanced hands—equipped with fingertip sensors capable of detecting forces as small as 3g—enable precise, dexterous manipulation across diverse object types [38, 39]. Beyond intelligence and manipulation, the platform emphasizes practical deployment through home-oriented safety features, including soft-tissue-like materials, reduced mass, and certified swappable battery systems, as well as usability enhancements such as wireless inductive charging through the feet and improved audio systems for bidirectional human–robot interaction [36, 38]. Engineered explicitly for high-volume production, Figure 03 leverages a vertically integrated supply chain, redesigned mechanical components, and a dedicated manufacturing facility (BotQ), which together reduce unit cost and support large-scale deployment [37, 40]. These combined advances position Figure 03 as a leading example of the transition from experimental humanoid platforms to commercially viable, scalable robotic systems for both domestic and industrial applications [36, 40].

C. Unitree G1/H1

Unitree Robotics offers two complementary humanoid platforms: the compact G1 [41] and the full-size H1 [42], targeting distinct segments of humanoid research and deployment. The G1 is a compact, manipulation-oriented humanoid (∼132 cm tall, ∼35 kg), configurable with 23–43 degrees of freedom and arm payloads up to roughly 2–3 kg, an open SDK supporting ROS2 and high-level programming, making it well suited for education, AI-driven manipulation research, and imitation/reinforcement-learning-based development environments. In contrast, the H1 series is a full-size, mobility-oriented humanoid (∼180 cm tall, ranging 47–70 kg, 19–27 DoF depending on variant), engineered for high-speed bipedal locomotion and industrial deployment, with peak knee joint torque up to 360 N·m and a peak torque density of ∼189 N·m/kg, enabling running speeds up to 3.3 m/s, one of the highest recorded for electric humanoid robots. H1 is equipped with 3D LiDAR, depth cameras, and a quickly replaceable 864 Wh battery that supports roughly 1.5–2 hours of runtime, and is designed for complex navigation, industrial logistics, inspection, and field operations, while G1 emphasizes agile, low-inertia actuation for responsive motion control and real-time decision-making. Together, G1 and H1 exemplify Unitree's strategy of offering a low-cost, research-friendly platform (G1) side-by-side with a high-power, industry-ready platform (H1), covering a broad spectrum of humanoid experimentation from algorithm and perception research to real-world automation scenarios[43, 44, 45].

D. Boston Dynamics Electric Atlas

Boston Dynamics' Electric Atlas represents a high-performance, fully-electric humanoid robot designed for demanding industrial and logistics tasks, marking a major shift from earlier hydraulic Atlas platforms toward practical, energy-efficient deployment [46, 47]. Standing approximately 1.9 m tall and weighing around 90 kg, Electric Atlas integrates 56 degrees of freedom, articulated head-mounted sensors, and fully electric actuation to enable high-power mobility, dexterous manipulation, and autonomous operation in shared human workspaces. The robot can sprint at speeds up to ∼2.5 m/s (∼9 km/h), with a reach of up to 2.3 m, and is capable of lifting loads of ∼50 kg instantaneously and ∼30 kg sustained, making it suitable for palletizing, material handling, and machine tending in industrial settings [48]. Electric Atlas is rated for roughly ∼4 hours of runtime per charge under typical workloads, with support for self-swappable batteries to enable continuous operation, and it is engineered with IP67-grade protection, 360° vision-like perception, and tight integration with Boston Dynamics' fleet management software (Orbit) for enterprise deployment. Unlike earlier choreography-oriented prototypes, the Electric Atlas generation is explicitly positioned as a production-oriented platform that combines dynamic locomotion, robust manipulation, and AI-aware perception pipelines, highlighting Boston Dynamics' strategic move from research-showcase robots toward industrially viable, commercially deployable humanoids.

E. Fourier GR-1 and GR-2

The Fourier GR-1 and GR-2, developed by Fourier Intelligence, form a generational pair of China-designed humanoid robots that emphasize high-power actuation, pure-vision environmental perception, and general-purpose task execution. The GR-1 is a 165 cm-tall, 55 kg humanoid featuring approximately 44 degrees of freedom and Fourier-designed Smart Actuators (FSA) that deliver dynamic mobility at up to about 5 km/h and enable a payload capacity of roughly 50 kg, targeting industrial automation, logistics, healthcare assistance, and AI-driven interaction scenarios [49]. GR-1 is notable for its pure-vision perception stack, which uses an array of RGB cameras and transformer-based BEV / occupancy networks to provide 360° environment understanding and precise, human-like spatial positioning without relying on LiDAR, thereby reducing hardware cost and complexity while supporting navigation and object interaction in real-world settings [50]. The GR-2 represents the next-generation platform, standing approximately 175 cm tall and weighing around 63 kg, with 53 degrees of freedom and 12-DoF tactile-equipped hands that support real-time force- and shape-adaptive grasping for complex manipulation [51]. GR-2 builds on FSA 2.0 actuators with peak torques up to ∼380 N·m, dual-encoder high-precision control, and a detachable battery that extends runtime to about two hours [52], positioning it as a modular, developer-oriented platform for AI research, industrial deployment, and long-term human-robot collaboration in logistics, manufacturing, and service tasks. Together, GR-1 and GR-2 illustrate Fourier's strategy of combining electric-actuation-driven dynamics, vision-centric environmental reasoning, and affordable, large-scale-ready hardware to support the transition from experimental humanoids toward industrially viable general-purpose systems.

F. Comparative Specifications

Together, these platforms illustrate the diversity of contemporary humanoid robotics, spanning AI-driven general-purpose systems targeting large-scale deployment (Tesla Optimus and Figure 03), compact open research platforms (Unitree G1), full-size high-performance locomotion and manipulation leaders (Unitree H1 and Boston Dynamics Electric Atlas), and early commercial deployment systems from emerging manufacturers (Fourier GR-1/GR-2) [31, 36, 41, 42, 46, 49, 51]. This multiplicity of designs, performance envelopes, and market strategies reflects an early yet rapidly expanding stage of the humanoid robotics ecosystem, driven by advances in AI control, high-torque electric actuation, and scalable manufacturing that are increasingly converging across global markets [34, 37, 52, 53].

Table I presents a comprehensive comparison of technical specifications across leading humanoid platforms. The specifications reveal distinct design philosophies. Unitree G1's compact stature (1.32m, 35kg) prioritizes safety in research environments, while Electric Atlas (1.9m, 90kg) emphasizes industrial strength. Payload capacity correlates with mass, ranging from 2-3 kg for G1 to ∼50 kg for Atlas.

TABLE I COMPARATIVE SPECIFICATIONS OF LEADING HUMANOID PLATFORMS

Feature	Optimus	Figure	Unitree G1	Unitree H1	Fourier	Atlas
Height (cm)	173	173	132	180	165-175	190
Weight (kg)	57	61	35	47-70	55-63	90
DoF	28	30	23-43	19-27	44-53	56
Payload (kg)	20	20	2-3	7	50	30-50
Runtime (hr)	4-8	5	2	1.5-2	1-2	4
Hand DoF	11-22	10	7	6	6-12	7

G. Software Architecture Comparison

The software architectures of leading humanoid platforms reflect distinct design philosophies regarding autonomy, learning, control, and deployment. While all platforms combine perception, planning, and control modules, they differ significantly in their emphasis on embodied AI, reinforcement learning, openness, and cloud-scale infrastructure.

Tesla Optimus leverages Tesla's Full Self-Driving (FSD) ecosystem, reusing much of the neural network architecture, data infrastructure, simulation environment, and distributed training pipeline originally developed for autonomous vehicles. This approach provides a significant advantage in large-scale data collection, fleet learning, and continuous model improvement. Optimus emphasizes vision-based perception, transformer-based neural networks, and end-to-end learning pipelines, supported by Tesla's extensive AI computing infrastructure [32]. However, the highly integrated and proprietary nature of the software stack limits external access and research flexibility.

Figure 03 adopts an AI-first architecture centered around its proprietary Helix system. Helix integrates Vision-Language-Action (VLA) models, large language models, perception systems, and motion planning into a unified framework capable of translating natural language instructions into physical actions [54]. Compared with traditional robotics pipelines, Figure places greater emphasis on semantic reasoning, task generalization, and multimodal learning. Extensive use of teleoperation and imitation learning allows the platform to acquire manipulation skills from human demonstrations and adapt to previously unseen objects and environments.

Unitree G1 is designed as an open and accessible humanoid research platform. Its ROS2-based software architecture exposes low-level control interfaces, sensor streams, and development tools to researchers, facilitating experimentation with reinforcement learning, imitation learning, and custom control algorithms [55]. The platform's openness has led to growing adoption within academia and has accelerated community-driven software development. As a result, G1 is increasingly used as a benchmark platform for humanoid robotics research.

Unitree H1, while sharing the same ROS2-based ecosystem, places greater emphasis on high-performance locomotion and dynamic balance. Its software stack focuses heavily on reinforcement-learning-based gait generation, disturbance recovery, and sim-to-real transfer. Compared with G1, H1 is optimized more toward demonstrating agile full-body locomotion and whole-body motion control rather than serving primarily as a general-purpose research platform [42]. Consequently, H1 has become a representative example of reinforcement-learning-driven humanoid locomotion.

Fourier GR-1 and GR-2 follow a different software strategy influenced by the company's background in rehabilitation robotics and exoskeleton systems. Rather than emphasizing highly dynamic locomotion or large-scale embodied AI, Fourier prioritizes compliant control, safe human-robot interaction, and adaptive assistance. The software architecture integrates whole-body control, force-sensitive interaction, motion planning, and perception modules within a ROS-based framework[56]. Recent developments have also incorporated machine learning techniques for perception and task execution; however, the overall architecture remains more control-centric than foundation-model-centric. This design philosophy makes the platform particularly suitable for healthcare, rehabilitation, and service-oriented applications where safety and human interaction are critical.

Boston Dynamics Atlas represents the most mature example of a model-based humanoid control architecture. Atlas combines highly sophisticated whole-body control, inverse dynamics, hierarchical optimization, and Model Predictive Control (MPC) to achieve exceptional locomotion performance. Unlike AI-centric systems such as Optimus and Figure, Atlas derives much of its capability from accurate physical modeling and real-time optimization. More recently, Boston Dynamics has begun integrating learned policies and simulation-based training through NVIDIA's Isaac and GR00T ecosystems, creating a hybrid architecture that combines the robustness of classical control with the adaptability of modern machine learning methods[57].

Overall, these platforms illustrate the broader evolution of humanoid robotics software. Atlas represents the culmination of model-based control methodologies, Optimus emphasizes large-scale AI infrastructure and fleet learning, Figure advances Vision-Language-Action-based embodied intelligence, Unitree promotes open and reinforcement-learning-oriented development, and Fourier focuses on human-centered interaction and assistive robotics. Despite their differences, the industry is increasingly converging toward hybrid architectures that combine deterministic control for safety and stability with learning-based methods for perception, manipulation, and decision making.

IV. Design Challenge Taxonomy

Despite rapid progress in humanoid robotics over the past decade, achieving reliable, general-purpose humanoid systems remains a formidable challenge. Advances in actuation, perception, and machine learning have enabled increasingly capable prototypes; however, these systems still face fundamental limitations that prevent widespread deployment in real-world environments. The challenges are not confined to a single domain but instead arise from the complex interaction between mechanical design, sensing, cognition, and system-level integration. As humanoid robots must simultaneously maintain dynamic balance, perceive unstructured environments, interpret high-level task goals, and execute safe physical interactions with humans, progress in one subsystem often exposes limitations in others. To clarify these interdependent challenges and provide a structured perspective on current research directions, we organize the design and deployment challenges of humanoid robots into four major categories: (1) mechanical and hardware challenges, (2) perception and sensing challenges, (3) cognitive and planning challenges, and (4) system integration challenges. This taxonomy highlights the multi-layered nature of humanoid robotics and provides a framework for analyzing how advances across these domains collectively shape the capabilities and limitations of modern humanoid platforms. Table II maps these challenges to platform capabilities.

A. Mechanical and Hardware Challenges

1) Actuator Power Density: Actuator performance remains a fundamental constraint in humanoid robot design [58]. Human skeletal muscle can achieve mechanical power densities of up to several hundred W/kg in short, localized bursts, although coordinated whole-body motion operates at substantially lower effective power levels [59]. In contrast, typical robotic actuator systems—comprising electric motors, transmissions, and structural components—often achieve effective mechanical power densities in the tens of W/kg range, limited by gearing losses, thermal constraints, and packaging requirements [58]. This gap restricts highly dynamic behaviors such as rapid running, jumping, and disturbance recovery. Recent research explores improved actuator architectures, including axial-flux motors, high-efficiency transmissions, and artificial muscle technologies (e.g., electroactive polymers and pneumatic artificial muscles), to increase specific power while maintaining compact form factors [60] [61]. In particular, axial-flux motors are attractive for high-performance humanoid joints because their disc-like geometry increases effective rotor radius and torque density while keeping axial length short, enabling direct integration into joint modules with lower inertia and better packaging efficiency [60]. By reducing the need for bulky off-joint transmissions, axial-flux designs can improve backdrivability, responsiveness, and continuous torque delivery, which are critical for agile leg and arm motion [60]. Higher actuator power density enables greater joint acceleration, improved locomotion speed, and increased payload capacity [58].

2) Energy Storage: Energy storage remains a critical limitation for autonomous humanoid robots. State-of-the-art lithium-ion batteries typically achieve energy densities on the order of 150–260 Wh/kg, constraining the operational endurance of mobile platforms [62]. For example, a 50 kg humanoid robot designed for continuous locomotion may require approximately 2–3 kWh of onboard energy to sustain multi-hour operation, resulting in 8–12 kg of battery mass, which significantly affects payload capacity and dynamic performance [62].

In legged locomotion research, energy consumption is commonly characterized using the dimensionless Cost of Transport (CoT) metric:

$$\mathrm{CoT} = \frac{E}{mgd}$$

where E is the total energy consumed, m is the system mass, g is gravitational acceleration, and d is the distance traveled [63]. Lower CoT values correspond to more efficient locomotion.

Emerging battery technologies, including solid-state batteries and lithium–sulfur (Li–S) chemistries, promise potential energy density improvements of 2–3× over current Li-ion systems, which could significantly extend the operational endurance of future humanoid platforms [64].

3) Hand Dexterity and Tactile Sensing: The human hand comprises 27 bones, 34 muscles, and on the order of 17,000 tactile mechanoreceptors in the glabrous skin. Current robotic hands achieve approximately 11–16 DoF with limited tactile sensing, representing a critical barrier to dexterous manipulation [65]. Tactile sensor arrays must achieve spatial resolutions of 1–2 mm and force resolutions of about 0.1 N to approximate human-level manipulation capability.

4) Structural Efficiency and Materials: Structural efficiency is an important design consideration in humanoid robots, often approximated by the ratio of payload capacity to total system mass. Compared with fixed-base industrial manipulators, humanoid platforms generally exhibit lower effective payload-to-mass ratios due to the additional requirements of mobility, balance, redundancy, and human-safe operation. While industrial manipulators can achieve payload ratios approaching or exceeding unity in optimized configurations, mobile humanoid systems typically operate at substantially lower levels. This reduced structural efficiency arises from the need for lightweight limbs, distributed actuation, and compliant mechanisms that ensure safe interaction with the environment.

To mitigate these limitations, researchers increasingly employ advanced lightweight materials, including carbon-fiber composites and titanium alloys, to reduce structural mass while maintaining mechanical strength [66]. In addition, topology optimization and generative structural design techniques are being widely applied in robotic structures to minimize weight while satisfying stiffness and strength constraints [67]. These approaches enable improved load-bearing capacity and energy efficiency but often introduce higher manufacturing complexity and cost.

B. Perception and Sensing Challenges

Reliable perception is fundamental for humanoid autonomy and requires the integration of multiple sensing modalities under strict real-time constraints. Humanoid robots typically rely on sensor fusion, combining inertial measurement units (IMUs), joint encoders, and exteroceptive sensors such as cameras or LiDAR, to estimate body pose and velocity. Accurate contact state estimation is particularly challenging due to sensor noise, unmodeled contact dynamics, and measurement delays, yet it is critical for stable locomotion and disturbance rejection [68].

Visual perception further introduces challenges associated with illumination variation, occlusions, and dynamic environments, which frequently occur in real-world human environments. Simultaneous Localization and Mapping (SLAM) methods provide an important framework for enabling long-term navigation. However, humanoid robots present additional difficulties for SLAM because of rapid viewpoint changes, body oscillations during locomotion, and frequent occlusions from self-motion. Even modern visual–inertial SLAM systems can accumulate non-negligible drift over long trajectories, which may limit accuracy for tasks requiring precise end-effector positioning during extended operation [69].

In addition to vision and inertial sensing, force and torque sensing at the feet and wrists plays a crucial role in balance control and manipulation. Multi-axis force/torque sensors enable estimation of ground reaction forces, detection of contact events, and regulation of interaction forces during manipulation [70]. Modern sensors can achieve high force and torque resolution with bandwidths suitable for real-time control; however, challenges remain in sensor miniaturization, robustness, and cost reduction, particularly for deployment in large-scale humanoid systems.

C. Cognitive and Control Challenges

Despite recent advances in reinforcement learning (RL) and model-based control, robust bipedal locomotion over uneven or unstructured terrain remains a significant challenge for humanoid robots. Locomotion policies must account for complex contact dynamics, terrain uncertainty, and real-time balance constraints [30]. Similarly, manipulation planning requires reasoning about object geometry, physical constraints, and task goals, often leading to combinatorial complexity in motion and task planning [2].

Recent work explores the integration of large language models (LLMs) and other foundation models to enable high-level task reasoning and instruction following. LLM-based planners can generate task sequences from natural language descriptions, providing a promising approach for flexible robot task planning. However, bridging the gap between high-level symbolic reasoning and low-level control policies remains an open problem [71]. A related direction is the Vision–Language–Action (VLA) paradigm, which seeks to map visual observations and language instructions directly to motor actions. Architectures such as RT-2 demonstrate the ability to generalize manipulation behaviors to previously unseen objects by leveraging large-scale multimodal training [72]. While these approaches show promise for improving robot generalization, challenges remain in achieving the precision, reliability, and safety required for real-world deployment.

Human–robot interaction (HRI) introduces additional cognitive and safety considerations. Effective collaboration requires natural communication, intent recognition, and socially appropriate behavior during shared tasks [73]. Although modern language models enable robust natural language understanding, grounding language instructions in physical actions and maintaining coherent multi-turn interactions during task execution remain active research areas. Furthermore, safe physical interaction requires compliant control and reliable collision detection mechanisms. Safety standards for collaborative robots define limits on allowable contact forces to reduce injury risk during human–robot interaction [74, 75].

D. Integration Challenges

Real-time system integration presents significant challenges for humanoid robots. Stable locomotion and manipulation typically require control loops operating at frequencies of 100–1000 Hz, placing strict constraints on sensing, computation, and actuation latency [2]. Achieving reliable performance at these update rates becomes increasingly difficult as perception, planning, and control modules grow more complex. In addition, ensuring safety and reliability during unsupervised operation in public environments remains an open challenge. Large-scale deployment also requires cost-effective manufacturability, motivating design-for-manufacturing approaches that simplify mechanical assemblies and reduce component count.

The computational requirements for humanoid autonomy are substantial. Modern perception pipelines based on deep neural networks can require tens to hundreds of GFLOPs per frame, particularly for tasks such as visual perception, object detection, and scene understanding [76]. Meanwhile, whole-body control and state estimation algorithms must operate with millisecond- or microsecond-level latency to maintain stability during dynamic locomotion. To meet these requirements, current humanoid platforms increasingly employ heterogeneous computing architectures that combine CPUs, GPUs, and specialized accelerators for perception, planning, and control workloads [77].

Safety certification for humanoids operating in human environments introduces additional challenges. Unlike traditional industrial robots operating in fenced workspaces, humanoids must function in shared environments with unpredictable human behavior. Existing industrial robot safety standards such as ISO 10218 [74] define requirements for industrial robotic systems, while collaborative robot standards such as ISO/TS 15066 [75] extend these frameworks to human–robot collaboration. However, certification methodologies for robots that rely on learning-based perception and control systems remain an active area of research.

TABLE II CHALLENGE-PLATFORM MAPPING MATRIX (S: SOLVED, P: PARTIAL, U: UNSOLVED, N: NOT ADDRESSED)

Challenge	Optimus	Figure	Unitree	Fourier	Atlas
Power Density	P	S	P	P	S
Energy Storage	P	P	U	P	N
Hand Dexterity	P	S	U	P	U
Tactile Sensing	U	P	U	P	U
State Estimation	S	S	S	P	S
Visual Perception	S	S	P	P	P
Dynamic Locomotion	P	P	S	P	S
Manipulation	P	S	U	P	U
Task Planning	P	S	U	U	P
Cost/Mfg	P	U	S	U	U

V. Theoretical and Practical Gaps

A. Theoretical Gaps

A. Theoretical Gaps

Despite significant progress in humanoid robotics, several fundamental theoretical challenges remain unresolved.

1) **Unified Locomotion–Manipulation Control:** Current control architectures typically treat locomotion and manipulation as largely separate problems, often combining locomotion controllers with task-level manipulation

strategies. Developing a unified control framework that simultaneously guarantees stability, contact feasibility, and task execution remains an open theoretical challenge [2, 78, 79].

2) **Generalization Guarantees for Learned Policies:** Reinforcement learning (RL) has demonstrated impressive results in legged locomotion and manipulation, enabling robots to acquire dynamic motor skills directly from data [30, 80]. However, policies learned through data-driven approaches often lack formal guarantees outside their training distributions, making deployment in safety-critical humanoid applications challenging [81, 82]. Ensuring robustness under distribution shift and providing provable safety and stability guarantees for learned controllers remain active research problems [83, 84].
3) **Human–Robot Skill Transfer:** Learning from demonstration (LfD) and imitation learning provide promising approaches for transferring human skills to robots; however, the theoretical foundations for efficient and reliable skill acquisition from limited demonstrations remain underdeveloped, particularly in high-dimensional control tasks [85, 86, 87]. Recent large-scale robot learning systems, such as RT-2, further highlight this limitation, as they rely on extensive multimodal datasets and web-scale knowledge to achieve generalization across tasks and environments [72].
4) **Energy-Optimal Control:** Passive dynamic walking demonstrates that stable bipedal locomotion can arise from the natural dynamics of mechanical systems with little or no actuation, revealing fundamental principles of energy-efficient locomotion. Building on these insights, recent work has explored generating families of energetically optimal gaits from passive solutions and stabilizing them through hybrid dynamical control frameworks such as Hybrid Zero Dynamics (HZD) [88, 89]. Additionally, modern humanoid platforms increasingly attempt to exploit passive dynamics through learning-based controllers to reduce the cost of transport and improve locomotion efficiency [90]. Nevertheless, extending passive dynamic principles to high-degree-of-freedom humanoid robots performing complex tasks—including manipulation and dynamic interaction with the environment—remains a major theoretical challenge. Whole-body control architectures further intensify this challenge by coupling locomotion, manipulation, balance, and contact regulation within a single high-dimensional constrained optimization problem, which makes it difficult to preserve the reduced-order passive-dynamics structure typically exploited in energy-optimal gait design.
5) **Stability of Learned Controllers:** Stability analysis for learning-based control policies remains a fundamental open theoretical problem. Classical control theory provides rigorous stability guarantees through tools such as Lyapunov analysis and control-theoretic stability criteria for model-based controllers. However, extending these guarantees to neural-network-based control policies is challenging due to their highly nonlinear and high-dimensional function approximations. Recent research has explored neural Lyapunov functions, learning-based stability certificates, and verification techniques for neural network controllers to provide formal guarantees for learned policies [91, 92]. Despite these advances, scaling these approaches to complex high-degree-of-freedom systems such as humanoid robots remains difficult, particularly when controllers are trained through reinforcement learning and deployed in dynamic, real-world environments.

B. Practical Gaps

- Simulation Fidelity: Accurate modeling of contact dynamics, actuator behavior, and compliant structures remains difficult for humanoid systems.
- Dataset Availability: Large-scale datasets of humanoid behaviors and physical interactions remain limited compared with fields such as computer vision and natural language processing.
- Benchmarking Standards: The lack of standardized evaluation protocols and metrics makes systematic comparison across humanoid platforms challenging.
- Maintenance Infrastructure: Maintenance, monitoring, and support ecosystems for large-scale real-world deployment remain immature.

The sim-to-real gap remains a persistent practical challenge in humanoid robotics. Although simulation environments enable large-scale training of control policies, discrepancies between simulated and real-world physics often lead to performance degradation during hardware deployment. Domain randomization and system identification techniques can partially mitigate these issues, but accurate modeling of contact dynamics—particularly friction, compliance, and impact forces—remains difficult [29, 93]. These effects are especially pronounced in humanoid locomotion and manipulation tasks involving intermittent contacts with the environment.

Dataset availability represents another critical bottleneck. While computer vision and natural language processing benefit from internet-scale datasets, robotics data collection requires physical interaction with the environment and is therefore significantly more expensive and time-consuming. Recent efforts such as Open X-Embodiment Dataset [94] and RT-1 Dataset [95] provide large-scale robot interaction data collected across multiple platforms, but datasets specifically focused on humanoid robots remain relatively scarce. Expanding diverse, high-quality datasets for humanoid locomotion, manipulation, and human-robot interaction remains an important direction for future research.

C. Research Priorities

Based on the challenges discussed above, several research directions appear particularly critical for advancing humanoid robotics toward large-scale real-world deployment. Key priorities include the development of unified whole-body control frameworks, scalable tactile sensing technologies, energy-efficient locomotion and manipulation strategies, robust task planning in unstructured environments, and learning methods with formal safety guarantees. Progress in these areas is essential to bridge the gap between laboratory demonstrations and practical applications in human environments. From a practical deployment perspective, the following areas appear especially important for near-term research investment:

1) **Tactile sensing and dexterous manipulation:** Reliable interaction with unstructured environments requires rich tactile feedback and highly dexterous manipulation capabilities. Current robotic hands remain significantly less capable than the human hand in terms of sensing density, compliance, and control bandwidth, limiting the range of practical tasks humanoid robots can perform.
2) **Energy efficiency and power systems:** Limited battery capacity and high energy consumption significantly constrain operational runtime. Improving actuator efficiency, locomotion strategies, and next-generation battery technologies is essential for enabling extended autonomous operation.
3) **Sim-to-Real transfer methods:** Large-scale policy training in simulation is increasingly necessary due to the cost and safety constraints of real-world experimentation. However, discrepancies between simulated and real-world dynamics continue to limit reliable policy transfer, motivating further research on domain randomization, system identification, and adaptive control methods.
4) **Safety certification and regulatory frameworks:** Humanoid robots are expected to operate in close proximity to humans in shared environments. Developing safety standards, verification methods for learning-based controllers, and regulatory certification frameworks remains a critical prerequisite for widespread deployment.
5) **Cost reduction and manufacturability:** Current humanoid platforms remain prohibitively expensive due to complex mechanical designs, custom components, and limited production scale. Advances in design-for-manufacturing, modular architectures, and supply chain optimization will be necessary to enable economically viable large-scale production.

VI. Bibliometric Analysis and Research Trends

Analysis of Scopus, Web of Science, and Google Scholar databases reveals publication trends in humanoid robotics research. This section presents quantitative analysis of publication patterns, geographic distribution, citation impact, and emerging research directions.

A. Publication Trends

Publications increased from approximately 500 annually (2015) to approximately 2,500 (2024), representing a 19% compound annual growth rate (CAGR). This growth reflects technological advances enabling new research directions and increased global funding for robotics research. Fig. 1 illustrates the publication growth trajectory over the past decade.

The acceleration in publication rate since 2020 coincides with the emergence of commercially viable platforms and advances in deep reinforcement learning. The COVID-19 pandemic initially disrupted research but subsequently accelerated interest in automation and robotics.

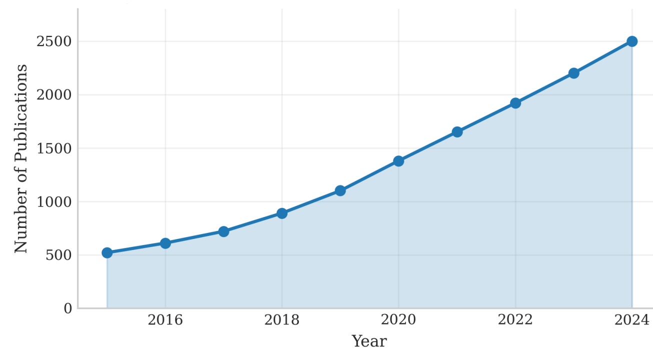

Fig. 1. Publication trends in humanoid robotics (2015-2024).

B. Research Hotspots

Keyword analysis reveals the distribution of research effort across sub-domains. Fig. 2 presents the research hotspot distribution by publication volume.

Bipedal locomotion and balance control remain the dominant research area (23%), reflecting the fundamental challenge of stable walking. Manipulation research (19%) has gained prominence as platforms achieve reliable locomotion. Human-robot interaction (15%) and learning/adaptation (14%) represent growing areas driven by deployment requirements.

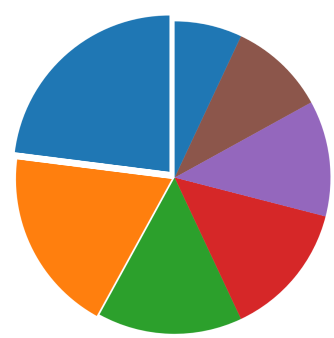

Fig. 2. Research hotspot distribution by publication volume.

C. Geographic Distribution

Research output is geographically concentrated, with East Asia leading in publication volume. Fig. 3 shows the geographic distribution of humanoid robotics research.

East Asia's leadership (38%) reflects strong government investment in robotics, particularly in Japan, China, and South Korea. Japan's aging population has motivated significant investment in care robotics. China's manufacturing sector drives demand for industrial automation. North America's contribution (31%) is concentrated in academic institutions and companies such as Boston Dynamics and Tesla.

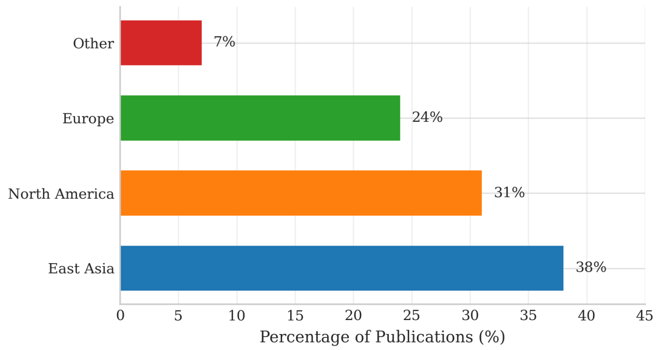

Fig. 3. Geographic distribution of humanoid robotics research.

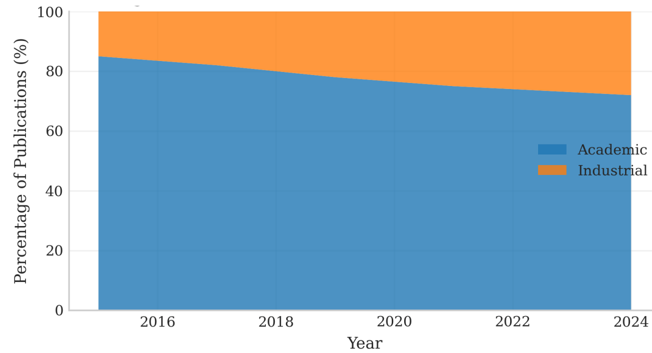

Fig. 5. Evolution of academic vs. industrial contributions.

D. Citation Impact and Foundational Works

Citation analysis reveals the foundational works that have shaped the field. Fig. 4 presents citation counts for key contributions.

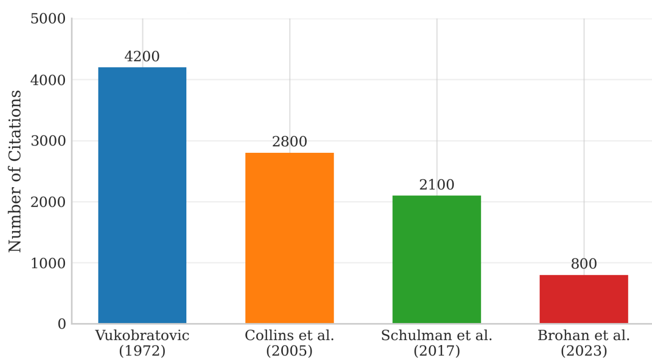

Fig. 4. Citation impact of foundational works in humanoid robotics.

The ZMP stability criterion [6] remains the most cited work, establishing the theoretical foundation for bipedal walking. Passive dynamic walking [96] demonstrated that efficient locomotion is possible without active control. Recent deep RL methods [18] show rapid citation growth, indicating the field's shift toward learning-based approaches.

E. Academic vs. Industrial Contributions

The balance between academic and industrial contributions has shifted over time. Fig. 5 illustrates the evolution of publication sources.

Academic institutions remain the primary source (72%), but industrial contributions increased from 15% (2015) to 28% (2024). Key industrial contributors include Boston Dynamics, Toyota, Honda, and emerging companies such as Tesla, Figure AI, and Unitree. This trend reflects the maturation of the field and commercial viability of humanoid platforms.

F. Emerging Research Directions

Analysis of recent publications and preprints reveals several emerging research directions in humanoid robotics: (1) foundation models for robotics, leveraging large-scale pre-training to improve generalization; (2) tactile sensing and haptic feedback for dexterous manipulation; (3) energy-efficient locomotion through optimization and learning-based gait synthesis; (4) multi-modal perception integrating vision, tactile, and proprioceptive sensing; (5) safe and robust learning-based control; (6) human–robot collaboration in shared environments; and (7) sim-to-real transfer and large-scale data generation.

Among these, foundation models for robotics have emerged as a significant new direction. Recent systems such as RT-2 and RoboCat demonstrate that large-scale pre-training across diverse tasks and modalities can improve generalization and reduce task-specific data requirements [72, 97]. In parallel, industrial initiatives such as NVIDIA's GR00T highlight efforts to scale these approaches to humanoid systems. However, substantial challenges remain, including high computational cost, data efficiency, and reliable real-world deployment on resource-constrained platforms.

In addition, the field is increasingly shaped by interdisciplinary contributions from computer vision, natural language processing, and reinforcement learning. While this convergence accelerates progress, it also introduces new challenges in integrating heterogeneous representations, aligning control and perception modules, and establishing standardized evaluation protocols across domains.

VII. Future Directions and Market Outlook

A. Technology Trajectories

Key trajectories shaping the next decade include continued advances in AI capabilities, improvements in actuation technology, progress in energy storage, increasing availability of affordable tactile sensing, and the maturation of simulation tools for large-scale training. In particular, foundation models pre-trained on large-scale datasets have shown promise in improving generalization and reducing sample complexity for robotic tasks [72].

B. Market Evolution

Industry analysts project significant market growth over the coming decade. Fig. 6 presents market projections for humanoid robots.

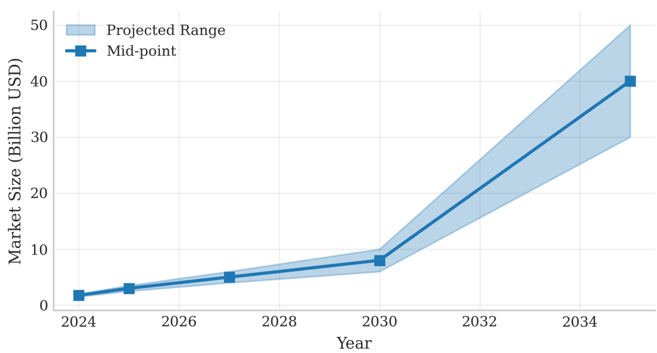

Fig. 6. Humanoid robot market projections

The humanoid robot market is projected to reach $6-10 billion by 2030 and $30-50 billion by 2035. Manufacturing applications will dominate early adoption, with service and domestic applications emerging as capabilities improve [98]. The RaaS (Robot-as-a-Service) business model is expected to accelerate adoption by reducing upfront capital expenditure.

C. Standardization and Safety

Recent updates to industrial robot safety standards, including ISO 10218 [99] and ISO/TS 15066 [75], increasingly address collaborative applications, including requirements for power and force limiting and functional safety. However, certification of learning-based systems for deployment in unstructured environments remains an open regulatory challenge. Ethical considerations are gaining prominence as humanoids approach deployment in human environments. Issues such as job displacement, privacy (particularly for vision-based systems), and autonomous decision-making in safety-critical scenarios require careful consideration.

D. Application Domains

Manufacturing remains the primary near-term application domain, with automotive manufacturers such as Tesla, BMW, and Hyundai actively piloting or evaluating humanoid systems for tasks such as parts handling, assembly, and material transport. Logistics and warehousing represent another promising domain, including applications such as truck unloading and order fulfillment. However, competition from autonomous mobile robots remains strong, requiring humanoids to demonstrate clear advantages. Healthcare and elderly care represent important long-term opportunities, particularly in aging societies such as Japan, where care robotics has been identified as a national priority.

VIII. Conclusion

This paper has presented a comprehensive review of humanoid robotics as of 2025, spanning technical, market, and research dimensions. The field appears to be approaching an inflection point, with several commercial platforms demonstrating capabilities that were largely speculative a decade ago.

The proposed taxonomy of design challenges provides a structured framework for assessing progress across the field. While significant advances have been achieved in locomotion and perception, major challenges remain in dexterous manipulation, long-duration autonomy, and cost-effective manufacturing. A comparison of existing platforms indicates that no current system fully addresses all key challenges, highlighting continued opportunities for innovation.

From a systems perspective, different platforms emphasize distinct capabilities. For example, platforms such as Unitree H1 offer a promising balance between performance and accessibility for research and development. Systems developed by Figure AI illustrate rapid progress in autonomous manipulation and industrial deployment, while Atlas continues to demonstrate leading capabilities in dynamic locomotion and whole-body control.

The goal of fully autonomous household assistance remains distant, constrained by limitations in manipulation dexterity, task planning, safety assurance, and cost. Nevertheless, current trends suggest that this objective is no longer purely aspirational. The coming decade will be critical in determining whether humanoid robots achieve widespread practical deployment.

Several factors will influence the pace of progress, including sustained investment in research and development—particularly in tactile sensing, energy systems, and embodied intelligence—as well as the establishment of safety standards and certification frameworks that enable deployment in real-world environments. Equally important are the development of viable business models and the level of public acceptance of robots operating in shared human spaces.

Finally, the current separation between high-cost industrial systems and lower-cost research platforms is likely to diminish as technologies mature and economies of scale reduce costs. Platforms that successfully balance capability, reliability, and affordability will be best positioned to shape the emerging market for general-purpose humanoid robots.

Acknowledgment

This work was partially supported by the National Science Foundation (NSF) under Grant No. #2028011. Any opinions, findings, and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] J. Wang, M. Zhou, X. Guo, and L. Qi, "Multiperiod asset allocation considering dynamic loss aversion behavior of investors," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 1, pp. 73–81, 2018.
[2] B. Siciliano and O. Khatib, Eds., *Springer Handbook of Robotics*, 2nd ed. Cham, Switzerland: Springer, 2016.
[3] S. Kuindersma *et al.*, "Design and control of atlas: A humanoid robot for the darpa robotics challenge," in *Proceedings of the IEEE International Conference on Humanoid Robots*, 2014, pp. 456–463.
[4] C. Xu, H. Wei, X. Guo, S. Liu, L. Qi, and Z. Zhao, "Human-robot collaboration multi-objective disassembly line balancing subject to task failure via multi-objective artificial bee colony algorithm," *IFAC-PapersOnLine*, vol. 53, no. 5, pp. 1–6, 2020.
[5] L. Cao, "Humanoid robots and humanoid ai: Review, perspectives and directions," *arXiv preprint arXiv:2405.15775*, pp. 1–35, 2024,

version 3, last revised 12 Oct 2025. [Online]. Available: https://arxiv.org/abs/2405.15775

[6] M. Vukobratovic and J. Stepanenko, "On the stability of anthropomorphic systems," *Mathematical Biosciences*, vol. 15, no. 1, pp. 1–37, 1972.

[7] Y. Ding *et al.*, "Robomind: Benchmarking multi-task intelligence for humanoid robots," 2024.

[8] S. Yan, H. Guo, and S. Liu, "A partition stacking classification framework with oversampling for quality prediction of aluminum alloy ingots," *International Journal of Artificial Intelligence and Green Manufacturing*, vol. 1, no. 1, pp. 36–46, March 2025.

[9] Z. Zhang, S. Dai, C. Li, W. Wang, J. Parron, and E. Herrera, "Human-robot collaborative disassembly profit maximization via improved grey wolf optimizer," *International Journal of Artificial Intelligence and Green Manufacturing*, vol. 1, no. 2, pp. 69–79, June 2025.

[10] IEEE Robotics and Automation Society, "Humanoid robotics technical committee," 2024. [Online]. Available: https://www.ieee-ras.org

[11] Figure AI, "Figure 02: Autonomous humanoid robot," 2025. [Online]. Available: https://www.figure.ai

[12] J. Denavit and R. Hartenberg, "A kinematic notation for lower-pair mechanisms based on matrices," *Journal of Applied Mechanics*, vol. 22, no. 2, pp. 215–221, 1955.

[13] M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot Modeling and Control*. Hoboken, NJ: Wiley, 2006.

[14] K. M. Lynch and F. C. Park, *Modern Robotics: Mechanics, Planning, and Control*. Cambridge, UK: Cambridge University Press, 2017.

[15] B. Colwell, "A complete review of boston dynamics' atlas robot," 2025, [Online; accessed May 24, 2026]. [Online]. Available: https://briandcolwell.com/a-complete-review-of-boston-dynamics-atlas-robot/

[16] Firgelli, "The death of hydraulics: Why every humanoid robot is going electric," 2026, [Online; accessed May 24, 2026]. [Online]. Available: https://www.firgelli.com/blogs/news/the-death-of-hydraulics

[17] C. Tang, B. Abbatematteo, J. Hu, R. Chandra, R. Martín-Martín, and P. Stone, "Deep reinforcement learning for robotics: A survey of real-world successes," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 8, no. 1, pp. 153–188, 2025.

[18] J. Schulman *et al.*, "Proximal policy optimization algorithms," 2017.

[19] L. Zhou, H. Zhu, and B. Akbari, "Multi-objective optimization of multi-factory remanufacturing process considering worker fatigue," *International Journal of Artificial Intelligence and Green Manufacturing*, vol. 1, no. 2, pp. 93–107, June 2025.

[20] R. S. Dahiya *et al.*, "Tactile sensing: From humans to humanoids," *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 1–20, 2010.

[21] G. Welch and G. Bishop, "An introduction to the kalman filter," Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, Tech. Rep. TR-95-041, 1995.

[22] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[23] C. Cadena *et al.*, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.

[24] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *Journal of Machine Learning Research*, vol. 17, no. 39, pp. 1–40, 2016. [Online]. Available: http://jmlr.org/papers/v17/15-221.html

[25] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa, "Biped walking pattern generation by using preview control of zero-moment point," in *2003 IEEE international conference on robotics and automation (Cat. No. 03CH37422)*, vol. 2. IEEE, 2003, pp. 1620–1626.

[26] J. Pratt, J. Carff, S. Drakunov, and A. Goswami, "Capture point: A step toward humanoid push recovery," in *2006 6th IEEE-RAS international conference on humanoid robots*. Ieee, 2006, pp. 200–207.

[27] P.-B. Wieber, "Trajectory free linear model predictive control for stable walking in the presence of strong perturbations," in *2006 6th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2006, pp. 137–142.

[28] O. KHATIB, "A unified approach for motion and force control of robot manipulators: the operational space formulation," *IEEE journal of robotics and automation*, vol. 3, no. 1, pp. 43–53, 1987.

[29] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, "Domain randomization for transferring deep neural networks from simulation to the real world," in *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2017, pp. 23–30.

[30] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, and M. Hutter, "Learning agile and dynamic motor skills for legged robots," *Science robotics*, vol. 4, no. 26, p. eaau5872, 2019.

[31] Wikipedia contributors, "Optimus (robot)," *Wikipedia – The Free Encyclopedia*, 2021, general overview of Optimus design, height, weight, and payload. [Online; accessed May 24, 2026]. [Online]. Available: https://en.wikipedia.org/wiki/Optimus_(robot)

[32] Tesla, Inc., *AI & Robotics - Tesla*, 2023–2026, official Tesla page on Optimus and AI-driven robotics. [Online; accessed May 24, 2026]. [Online]. Available: https://www.tesla.com/AI

[33] H. Guide, *Optimus Gen 2 - Humanoid Robot Guide*, 2026, specs, speed, weight, and performance overview of Optimus Gen 2. [Online; accessed May 24, 2026]. [Online]. Available: https://humanoid.guide/product/optimus-gen2/

[34] T. . T. Accessories, "Tesla Optimus 3 Production Timeline and the Dawn of General-Purpose Humanoids," 2026, discussion of Optimus Gen 3 production start in Summer 2026 and high-volume plans. [Online; accessed May 24, 2026]. [Online]. Available: https://www.tesery.com/blogs/news/elon-musk-reveals-aggressive-production-timeline-for-tesla-optimus-3

[35] Tesla Accessories / Third-party overview, "Tesla Optimus Evolution: Version 3 Production Timeline and the Dawn of General-Purpose Humanoids," 2026, summary of Optimus 3 roadmap and industrial deployment strategy. [Online; accessed May 24, 2026]. [Online]. Available: https://www.teslaacessories.com/blogs/news/tesla-optimus-evolution-version-3-production-timeline-and-the-dawn-of-general-purpose

[36] Figure AI, *Figure 03 – Humanoid Robot by Figure AI*, 2025, official product page. [Online]. Available: https://www.figure.ai/figure

[37] Figure AI, "Introducing Figure 03," *Figure AI News*, 10 2025, [Online; accessed May 24, 2026]. [Online]. Available: https://www.figure.ai/news/introducing-figure-03

[38] H. Guide, *Figure 03 – Humanoid Robot Guide*, 2026, technical specifications and performance overview. [Online]. Available: https://humanoid.guide/product/figure-03/

[39] R. Blog, "Figure 03 Review: Home Robot," 2026, in-depth review of Figure 03 capabilities, payload, weight, and Helix integration. [Online; accessed May 24, 2026]. [Online]. Available: https://blog.robozaps.com/b/figure-03-review

[40] BotInfo.ai, "Figure 03: Deep Dive Analysis of Figure AI's Home Robot Strategy," 2026, analysis of vertical integration, BotQ facility, and production strategy. [Online; accessed May 24, 2026]. [Online]. Available: https://botinfo.ai/articles/figure-03-deep-dive-analysis

[41] Unitree Robotics, *Unitree G1 – Open-Source Humanoid Robot*, 2023, compact manipulation-focused humanoid, 23–43 DoF, 35 kg, 2–3 kg arm payload. [Online; accessed May 24, 2026]. [Online]. Available: https://www.unitree.com/g1/

[42] Unitree Robotics, *Unitree H1 – Full-Size Universal Humanoid Robot*, 2023, [Online; accessed May 24, 2026]. [Online]. Available: https://www.unitree.com/h1/

[43] A. Robots, "Unitree G1 vs H1: Which Humanoid Robot is Better?" 2026, [Online; accessed May 24, 2026]. [Online]. Available: https://www.awesomerobots.xyz/blog/unitree-g1-vs-h1-comparison

[44] RoboStore, "Discover the Optimal Robotics Solution for R&D: Unitree G1 EDU vs H1-2," 2024, [Online; accessed May 24, 2026]. [Online]. Available: https://robostore.com/blogs/news/discover-the-optimal-robotics-solution-for-r-d-unitree-g1-edu-vs-h1-2-comparison

[45] Robozaps, "Unitree H1 Review: $90K Humanoid Specs," 2026, [Online; accessed May 24, 2026]. [Online]. Available: https://blog.robozaps.com/b/unitree-h1-review

[46] Boston Dynamics, *Atlas Humanoid Robot - Boston Dynamics*, 2024–2026, [Online; accessed May 24, 2026]. [Online]. Available: https://bostondynamics.com/products/atlas/

[47] H. Guide, *Atlas Humanoid Robot by Boston Dynamics — Capabilities & Specs*, 2026, overview of Electric Atlas: 56 DoF, 90 kg, 2.3 m reach, 50 kg lift, 4 h runtime. [Online]. Available: https://humanoid.guide/product/atlas/

[48] Sparkco, "Boston Dynamics Atlas Robot Industrial Applications," 2022, [Online; accessed May 24, 2026]. [Online]. Available: https://sparkco.ai/blog/boston-dynamics-atlas-robot-industrial-applications

[49] Fourier Intelligence, *GR-1 - FOURIER-Robotics*, 2023–2026, [Online; accessed May 24, 2026]. [Online]. Available: https://www.fftai.com/products-gr1

[50] Fourier Intelligence, "Fourier Unveils GR-1's Breakthrough in Vision Technology for Enhanced Perception," *Fourier Newsroom*, 2024, pure vision solution with 6 RGB cameras, BEV, Transformer, and occupancy networks for GR-1. [Online; accessed May 24, 2026]. [Online]. Available: https://www.fftai.com/newsroom-newintech/15

[51] H. Guide, *GR-2 - Humanoid Robot Guide*, 2026, [Online; accessed May 24, 2026]. [Online]. Available: https://humanoid.guide/product/gr-2/

[52] T. R. Report, "Fourier launches GR-2 humanoid, software

platform," 2024, [Online; accessed May 24, 2026]. [Online]. Available: https://www.therobotreport.com/fourier-launches-gr-2-humanoid-software-platform/

[53] Boston Dynamics, "Large Behavior Models and Atlas Find New Footing," *Boston Dynamics Blog*, 2025, [Online; accessed May 24, 2026]. [Online]. Available: https://bostondynamics.com/blog/large-behavior-models-atlas-find-new-footing/

[54] Figure AI, "Helix: A vision-language-action model for generalist humanoid control," https://www.figure.ai/news/helix, Feb. 2025, accessed 2026-06-01.

[55] Unitree Robotics. (2023) unitreerobotics/unitree_ros2. Accessed 2026-05-29. [Online]. Available: https://github.com/unitreerobotics/unitree_ros2

[56] Fourier Intelligence, "GR-1 Humanoid Robot Introduction," https://support.fftai.com/en/docs/GR-X-Humanoid-Robot/GR1/GR-1_Introduction/, 2024, accessed: 2026-06-09.

[57] NVIDIA. (2026, mar) Humanoid robots — use case. Accessed 2026-05-29. [Online]. Available: https://developer.nvidia.com/en-us/use-cases/humanoid-robots/

[58] X. Liu, Z. Hong, and W. Zhang, "High-performance torque feedback actuator for humanoid robots," in *2024 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2024, pp. 1048–1053.

[59] K. Kubo, T. Ikebukuro, and H. Yata, "Mechanical properties of muscles and tendon structures in middle-aged and young men," *Scientific Reports*, vol. 12, no. 1, p. 1702, 2022.

[60] X. Yu, X. Zhang, Z. Ma, Y. Zhang, T. Huang, and W. Geng, "Design and analysis of a novel modular axial flux permanent-magnet high torque density joint motor for humanoid robots with toroidal windings," in *2025 28th International Conference on Electrical Machines and Systems (ICEMS)*. IEEE, 2025, pp. 596–600.

[61] M. Shi and E. M. Yeatman, "A comparative review of artificial muscles for microsystem applications," *Microsystems & Nanoengineering*, vol. 7, no. 1, p. 95, 2021.

[62] TrendForce, "Humanoid robots move toward commercialization; solid-state batteries address power density and safety challenges," *TrendForce Industry Report*, 2026, https://www.trendforce.com/presscenter/news/20260128-12902.html.

[63] N. Kashiri, A. Abate, S. J. Abram, A. Albu-Schaffer, P. J. Clary, M. Daley, S. Faraji, R. Furnemont, M. Garabini, H. Geyer *et al.*, "An overview on principles for energy efficient robot locomotion," *Frontiers in Robotics and AI*, vol. 5, p. 129, 2018.

[64] S. Yari, A. Conde Reis, Q. Pang, and M. Safari, "Performance benchmarking and analysis of lithium-sulfur batteries for next-generation cell design," *Nature Communications*, vol. 16, no. 1, p. 5473, 2025.

[65] G. Li, R. Wang, P. Xu, Q. Ye, and J. Chen, "The developments and challenges towards dexterous and embodied robotic manipulation: A survey," *arXiv preprint arXiv:2507.11840*, 2025.

[66] H. Zong, J. Ai, K. Zhang, L. Fang, L. Jiang, J. Zhang, and B. Xu, "Lightweight and modular design of a hydraulically actuated quadruped robot with high payload-to-mass ratio," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 239, no. 9, pp. 3403–3416, 2025.

[67] L. Sha, A. Lin, X. Zhao, and S. Kuang, "A topology optimization method of robot lightweight design based on the finite element model of assembly and its applications," *Science Progress*, vol. 103, no. 3, p. 0036850420936482, 2020.

[68] S. Kolev and E. Todorov, "Physically consistent state estimation and system identification for contacts," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2015, pp. 1036–1043.

[69] R. Hartley, J. Mangelson, L. Gan, M. G. Jadidi, J. M. Walls, R. M. Eustice, and J. W. Grizzle, "Legged robot state-estimation through combined forward kinematic and preintegrated contact factors," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 4422–4429.

[70] J.-H. Kim, "Multi-axis force-torque sensors for measuring zero-moment point in humanoid robots: A review," *IEEE Sensors Journal*, vol. 20, no. 3, pp. 1126–1141, 2019.

[71] M. Ahn, A. Brohan, N. Brown, Y. Chebotar, O. Cortes, B. David, C. Finn, C. Fu, K. Gopalakrishnan, K. Hausman *et al.*, "Do as i can, not as i say: Grounding language in robotic affordances," *arXiv preprint arXiv:2204.01691*, 2022.

[72] A. Brohan *et al.*, "Rt-2: Vision-language-action models," 2023.

[73] M. A. Goodrich and A. C. Schultz, "Human–robot interaction: a survey," *Foundations and trends® in human–computer interaction*, vol. 1, no. 3, pp. 203–275, 2008.

[74] International Organization for Standardization, "Iso 10218-1:2011 robots and robotic devices – safety requirements for industrial robots," ISO, 2011.

[75] International Organization for Standardization, "Iso/ts 15066:2016 robots and robotic devices – collaborative robots," ISO, 2016.

[76] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on robotics*, vol. 32, no. 6, pp. 1309–1332, 2017.

[77] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, "Learning robust perceptive locomotion for quadrupedal robots in the wild," *Science robotics*, vol. 7, no. 62, p. eabk2822, 2022.

[78] W. Sun, L. Feng, B. Cao, Y. Liu, Y. Jin, and Z. Xie, "Ulc: A unified and fine-grained controller for humanoid loco-manipulation," *arXiv preprint arXiv:2507.06905*, 2025.

[79] Y. Wang, M. Yang, Z. Ding, Y. Zhang, W. Zeng, X. Xu, H. Jiang, and Z. Lu, "From experts to a generalist: Toward general whole-body control for humanoid robots," *arXiv preprint arXiv:2506.12779*, 2025.

[80] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, "Learning to walk in minutes using massively parallel deep reinforcement learning," in *Conference on robot learning*. PMLR, 2022, pp. 91–100.

[81] B. Recht, R. Roelofs, L. Schmidt, and V. Shankar, "Do imagenet classifiers generalize to imagenet?" in *International conference on machine learning*. PMLR, 2019, pp. 5389–5400.

[82] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané, "Concrete problems in ai safety," *arXiv preprint arXiv:1606.06565*, 2016.

[83] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *2019 18th European control conference (ECC)*. Ieee, 2019, pp. 3420–3431.

[84] F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," *Advances in neural information processing systems*, vol. 30, 2017.

[85] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.

[86] T. Osa, J. Pajarinen, G. Neumann, J. A. Bagnell, P. Abbeel, and J. Peters, "An algorithmic perspective on imitation learning," *Foundations and Trends® in Robotics*, vol. 7, no. 1-2, pp. 1–179, 2018.

[87] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, "Recent advances in robot learning from demonstration," *Annual review of control, robotics, and autonomous systems*, vol. 3, no. 1, pp. 297–330, 2020.

[88] N. Rosa, B. Katamish, M. Raff, and C. D. Remy, "An approach for generating families of energetically optimal gaits from passive dynamic walking gaits," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 8551–8557.

[89] A. B. Ghansah, J. Kim, K. Li, and A. D. Ames, "Dynamic walking on highly underactuated point foot humanoids: Closing the loop between hzd and hlip," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 12 686–12 693.

[90] B. Xia, B. Li, J. Lee, M. Scutari, and B. Chen, "The duke humanoid: Design and control for energy-efficient bipedal locomotion using passive dynamics," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2025, pp. 6579–6586.

[91] J. Wu, A. Clark, Y. Kantaros, and Y. Vorobeychik, "Neural lyapunov control for discrete-time systems," *Advances in neural information processing systems*, vol. 36, pp. 2939–2955, 2023.

[92] Y.-C. Chang, N. Roohi, and S. Gao, "Neural lyapunov control," *Advances in neural information processing systems*, vol. 32, 2019.

[93] X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel, "Sim-to-real transfer of robotic control with dynamics randomization," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 3803–3810.

[94] A. O'Neill, A. Rehman, A. Maddukuri, A. Gupta, A. Padalkar, A. Lee, A. Pooley, A. Gupta, A. Mandlekar, A. Jain *et al.*, "Open x-embodiment: Robotic learning datasets and rt-x models: Open x-embodiment collaboration 0," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 6892–6903.

[95] A. Brohan, N. Brown, J. Carbajal, Y. Chebotar, J. Dabis, C. Finn, K. Gopalakrishnan, K. Hausman, A. Herzog, J. Hsu *et al.*, "Rt-1: Robotics transformer for real-world control at scale," *arXiv preprint arXiv:2212.06817*, 2022.

[96] S. Collins *et al.*, "Efficient bipedal robots based on passive-dynamic walkers," *Science*, vol. 307, no. 5712, pp. 1082–1085, 2005.

[97] K. Bousmalis, G. Vezzani, D. Rao, C. Devin, A. X. Lee, M. Bauzá, T. Davchev, Y. Zhou, A. Gupta, A. Raju *et al.*, "Robocat: A self-improving generalist agent for robotic manipulation," *arXiv preprint*

arXiv:2306.11706, 2023.
[98] Goldman Sachs Research, "The global market for humanoid robots," 2025, market Analysis Report.
[99] International Organization for Standardization, "Robots and robotic devices—safety requirements for industrial robots—part 1/2: Robots and robot systems and integration," 2025, iSO 10218-1/2.

Nan Wang received her M.Sc. in Computer Engineering and Ph.D. in Computer Science in 2009 from Mississippi State University and is currently an Associate Professor of Computer Science at William Paterson University. Her research focuses on generative AI, data mining, machine learning, computational biology, Robotics and blockchain. She currently leads research on programming humanoid robots for classroom and industrial use.

Ali Mohammad Saghiri received his M.Sc. and Ph.D. degrees in Computer Engineering from the Amirkabir University of Technology. He is currently an Assistant Professor of Computer Science at William Paterson University. His research spans blockchain, Web 3.0, artificial intelligence, Robotics and distributed systems. He has authored four books and over 50 scientific papers, focusing on the evolution of self-adaptive systems and the societal impacts of AI.

Bohan Fan received his Ph.D. in Computer Science in 2023 from the University of Illinois Chicago. He is currently an NTTP Instructor of Computer Science at William Paterson University. His research interests include geometric algorithms, machine learning, generative AI, and humanoid robotics.