Humanoid Robotics: From Engineering to Daily Partners

Samer AlFayad[1], Maya Sleiman[2], Feilong Wang[1] and Hang Su[1]

***Abstract*—This article presents a perspective on the current state and possible future directions of humanoid robotics, with particular emphasis on the gap between recent engineering advances and the realization of humanoid robots as reliable daily partners for humans. Although rapid progress has been achieved in locomotion, manipulation, and artificial intelligence, most humanoid systems remain primarily optimized for performance-driven demonstrations and staged showcases, with limited sustained integration into everyday human environments. We argue that closing this gap requires a fundamental shift from motion-centric engineering toward embodied, perceptive, and human-centered system design. Key enabling directions include integrated mechatronic and sensing architectures, multimodal perception and interaction, emotion recognition and affect-aware behavior, safety mechanisms for close human–robot coexistence, and large language model (LLM)–driven robotic intelligence and autonomous learning under practical constraints. Rather than providing a comprehensive technical review, this article highlights current limitations, synthesizes emerging research trends, and outlines a vision in which humanoid robots are engineered not merely to demonstrate technical capability, but to earn trust, adapt to human needs, and function as dependable partners in daily life.**

I. Introduction

Humanoid robots have long been regarded as one of the most ambitious frontiers in robotics: embodied machines that can move, perceive, and interact in spaces built for human bodies and human routines [1], [2]. Over the past decade, advances in actuation and power-dense hardware, whole-body control, learning-based perception, and increasingly capable AI models have collectively pushed humanoid platforms toward levels of mobility and dexterity that were previously out of reach [3], [4]. Contemporary systems can execute dynamic locomotion, perform coordinated whole-body motions, and manipulate objects with improving precision, driving renewed academic momentum and heightened public attention [5], [6].

This renewed attention is also amplified by a wave of high-profile industrial initiatives. Tesla's Optimus program frames the humanoid as a general-purpose worker intended for unsafe, repetitive, or undesirable tasks, reflecting an industry belief that the next scale-up of embodied AI may come from vertically integrated autonomy stacks and large-scale engineering. Figure AI has pursued an explicitly workforce-oriented narrative, reporting pilot deployments in industrial settings and highlighting lessons from extended runtime and task repetition in production-like environments. In China, Unitree has attracted broad interest by commercializing comparatively accessible humanoid platforms and showcasing highly dynamic behaviors, which can accelerate ecosystem experimentation but also underscores how far the field remains from robust, long-horizon autonomy. More recently, EngineAI ("Zhongqing") and XPENG have publicized full-size humanoid platforms positioned for practical deployment—from factory-oriented "work" scenarios to a broader "Physical AI" roadmap—signaling both rapid iteration and intensifying competition. Collectively, these programs illustrate genuine progress in hardware integration and whole-body capability, while also serving as a reminder that public-facing milestones are often evaluated under curated conditions.

Yet the field faces a persistent deployment gap [7]. Despite impressive progress, most humanoid robots remain confined to laboratories, controlled settings, or carefully staged showcases. Success in short-horizon demonstrations does not necessarily translate into sustained operation in everyday human environments, where uncertainty is the norm and requirements extend beyond nominal performance. Daily deployment demands robustness to occlusion and clutter, long-horizon reliability under wear and drift, maintainability and serviceability, and interaction behaviors that are not only physically safe but also socially acceptable and predictable [7], [8].

Closing this gap requires more than incremental improvements in locomotion controllers or mechanical strength [2]. It calls for a shift in how humanoid systems are conceived, designed, and evaluated: from motion-centric optimization toward embodied, perceptive, and human-centered system design [9]. In this framing, the humanoid body is not merely a carrier of algorithms, but a sensing-and-actuation substrate that shapes what the robot can observe, how it can respond under contact and uncertainty, and whether it can coexist comfortably with people [5]. This transition places new emphasis on integrated mechatronic co-design, distributed and embedded sensing, multimodal perception and interaction, affect-aware behavior, and safety mechanisms suited to close human–robot proximity [10], [11]. Equally important, it motivates evaluation criteria aligned with daily partnership—repeatability, uptime, graceful degradation, transparency, and trust—rather than one-off success in tightly scripted tasks [12], [13].

This article presents a perspective on the technological directions and conceptual reframing needed to move humanoid robotics from engineering showcases toward dependable daily partners [2]. Section II discusses mechatronic foundations, arguing that mechanical design, actuation, compliance, sensing,

[1]Samer AlFayad, Feilong Wang and Hang Su and are with IBISC Laboratory, University of Paris Saclay, 91000 Évry, Paris, France. (e-mail: hang.su@ieee.org)

[2]Maya Sleiman is with Kalysta Actuation and with IBISC Laboratory, University of Paris Saclay, 91000 Évry, Paris, France. (e-mail: mayasleiman@ieee.org).

and control should be co-designed from the outset, with growing importance placed on integrated and embedded sensing architectures that reduce fragility and improve observability at the body level. Section III focuses on multimodal interaction as a closed-loop process spanning multimodal perception, emotion recognition and affective state inference, intention communication, and predictive safety for close human–robot coexistence. Section IV examines robotic intelligence enabled by large language models and agent-based learning, emphasizing how high-level reasoning and autonomous adaptation can be made practical only when grounded in embodiment and constrained by verification, auditability, and runtime safety envelopes [14]. Rather than a comprehensive technical survey, we synthesize emerging research trends, highlight current limitations, and outline a forward-looking vision for humanoid robots that are designed not merely to demonstrate technical capability, but to earn trust, adapt to human needs, and integrate meaningfully into daily life [12], [15].

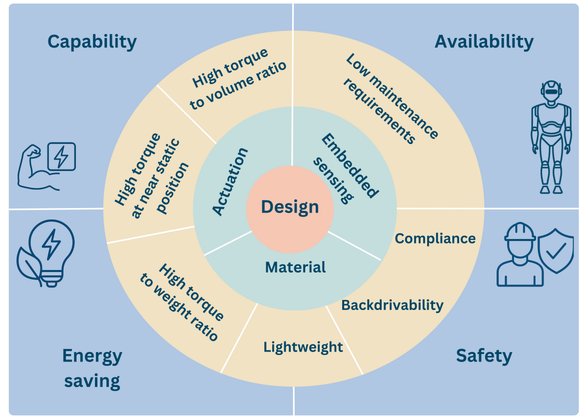

Fig. 1. Mechatronic co-design in humanoid robotics unifies actuation, materials, and embedded sensing within a single design framework, avoiding sequential integration and enabling safe, improved energy efficiency, and robust human–robot interaction.

II. Mechatronic Foundations of Humanoid Robots

Mechatronic design forms the physical and functional foundation of humanoid robotics, directly shaping system capability, safety, reliability, and the quality of human-robot interaction. Unlike industrial manipulators operating in structured environments, humanoid robots are required to move, perceive, and interact within spaces occupied for humans. This requirement places unique constraints on the mechanical architecture, actuation, sensing, and control, and demands a tightly integrated, system-level mechatronic design philosophy.

A. From Component Integration to Mechatronic Co-Design

Mechatronic design is the cornerstone of humanoid robotics, as it directly influences system safety, operational availability, and the effectiveness of human–humanoid robot collaboration. When humanoid robots are intended to operate continuously in human environments, system design cannot rely on isolated decisions in mechanics, sensing, actuation, and control followed by late-stage integration. Instead, and as represented in Fig. 1, mechanical structure, actuation principles, sensing strategies, control architectures, and computation must be developed jointly from the beginning as elements of a unified system.

B. Actuation, Compliance, and Physical Interaction

Achieving motion that is physically compatible with human movement while simultaneously ensuring efficiency, safety, and strength remains a fundamental challenge in humanoid robotics. Contemporary humanoid platforms predominantly rely on electric actuation due to its maturity, controllability, and suitability for scalable deployment. Within this context, several alternative actuation principles continue to be explored to investigate different trade-offs in compliance, power delivery, and interaction dynamics.

- Variable stiffness actuators (VSAs) modulate joint impedance at the mechanical level by adjusting stiffness independently of position or torque. In mobile robots, and in particular walking robots, VSAs can improve energy efficiency by exploiting the energy-storing properties of internal springs [16]. Pre-loaded elastic elements provide a short-duration power boost during transient events. However, the output power that can be delivered continuously and reliably remains fundamentally limited by the power capabilities of the motor [17]. Moreover, VSAs typically comprise multiple mechanical and actuation components that require a high level of integration and coordination, increasing system complexity and limiting their deployment in full humanoid platforms.
- Series elastic actuators (SEAs) introduce intentional mechanical compliance by placing an elastic element between the actuator and the load [18]. The resulting compliance enables passive energy absorption and accurate force regulation, which significantly improves robustness and safety during physical interaction and contact-rich tasks. Hence, they exhibit several advantageous properties, including high control bandwidth at moderate force levels, low output impedance, a wide dynamic range, effective rejection of internal disturbances, and inherent tolerance to shock loading [19]. As a result, SEAs are selectively adopted in humanoid subsystems where interaction quality and robustness are prioritized over peak dynamic performance [20].
- Hydraulic and electro-hydraulic systems continue to be investigated for applications requiring high power density [21] and the ability to sustain large forces at fixed or near-static positions [22]. In such quasi-static conditions, force generation is primarily governed by fluid pressure rather than continuous electrical current at the joint, which can be advantageous from an energy and battery usage perspective. From a system standpoint, hydraulic actuation can be implemented using either centralized hydraulic architectures or decentralized electro-hydraulic actuation (EHA). In centralized systems, hydraulic energy is generated by one or more central pumps and distributed through pressurized fluid lines to multiple joints. This approach enables efficient redistribution of high hydraulic

energy to different actuators when needed and supports high force density ratio at the joint level. However, centralized hydraulics introduce challenges related to fluid routing, leakage risk, noise, maintenance, and limited scalability in humanoid platforms.

Decentralized electro-hydraulic actuation adopts a hybrid architecture in which hydraulic energy is generated locally at each joint using electrically driven pumps and compact hydraulic circuits [22]. This configuration retains key advantages of hydraulic actuation, such as high force density and the ability to sustain large forces at near-static positions, while leveraging electric actuation for local control, modularity, and easier integration. By eliminating long hydraulic lines and central fluid distribution, EHA architectures improve fault isolation, simplify maintenance, and enhance system modularity, making them more compatible with humanoid robot design constraints. Given rising concerns over sustainable energy and environmental impact, hybrid technology is increasingly recognized as an effective solution to energy challenges [23]. While electric actuation remains well-suited for dynamic, efficient, and precisely controlled motion, both centralized hydraulics and decentralized EHA solutions remain attractive for prolonged load-bearing or force-intensive tasks.

Beyond actuation, lightweight materials and optimized mass distribution play a central role in improving balance, agility, and energy efficiency, directly reducing actuator loading, peak power demands, and thermal stress. By lowering inertia and impact forces, these design choices also promote safer physical interaction. Recent advances in high-strength aluminum alloys, carbon-fiber-reinforced composites [24], and additive manufacturing techniques [25] enable complex, load-efficient structures that reduce mass while maintaining stiffness and durability . Such developments support not only stable locomotion and manipulation, but also extended operational endurance and high system availability in human-centered environments.

C. Embedded and Integrated Sensing

As humanoid robots evolve toward being everyday partners, mechatronic systems must be conceived from the beginning to support perception, physical interaction, and long-term adaptive behavior, providing a robust basis for intelligent decision-making and collaboration. Beyond generating motion, humanoid robots must continuously perceive their own state, the surrounding environment, and physical contact with humans. Such continuous sensing enables not only safe regulation of interaction forces during close human collaboration, but also monitoring of internal loads, temperatures, and usage patterns over time. This information forms the basis for structural health monitoring (SHM) [26], enabling early detection of degradation, preventive maintenance, and sustained operational availability throughout prolonged deployment.

Traditional humanoid designs often treat sensing as a post hoc addition, integrated only after the mechanical architecture has been finalized. This approach restricts sensing coverage and limits optimal sensor placement choices. Externally mounted sensors also increase system bulkiness, complicate calibration and maintenance procedures, and hinder scalability and future upgrades. As a result, increasing attention is being directed toward embedded and distributed sensing, in which force, torque, tactile, strain, and thermal sensors are integrated directly within joints, links, and surface structures. A similar approach is often adopted with electronic boards, which are typically integrated only after the mechanical architecture has been finalized.

Mechanically integrated sensing requires that joint housings, limb shells, end-effectors, and load-bearing structures incorporate predefined cavities, routing channels, and interface points for diverse sensors. These include vision modules, force–torque sensors, tactile arrays, thermal elements, inertial units, and future modalities not yet widely deployed. Strategic placement, such as embedding tactile elements along compliant surfaces, distributing pressure and shear sensors across the foot sole, ensures that perceptual data corresponds closely to mechanical interactions with the environment. Designing these interfaces at the mechanical level allows sensors to be co-located with critical contact regions, reducing signal noise and improving real-time responsiveness.

Furthermore, the mechanical design should reserve modular expansion interfaces, ensuring compatibility with future advances in sensing hardware. Standardized mounting points, unified internal connector pathways, and replaceable sensor modules enable humanoids to evolve as sensing technologies mature. This is essential as emerging technologies such as chemical sensing, active tactile skins, or micro-pressure wave sensors will likely become integral to human-centered robotic function. A mechanically modular architecture ensures that sensors can be upgraded without requiring redesign of the entire robot body.

Recent advances in manufacturing technologies now make this integration feasible. Techniques such as multi-material fabrication, structural embedding of sensing elements, and integrated routing of electrical interconnects enable sensors to become intrinsic parts of load-bearing components rather than add-on attachments [27] [28]. In parallel, progress in additive manufacturing and composite material processing, particularly carbon-fiber-reinforced polymers and hybrid metal–composite structures, allows the fabrication of lightweight, load-efficient components with embedded sensing and wiring in a single manufacturing process. These developments support compact designs while preserving structural strength, durability, and accessibility.

By tightly coupling sensing and a distributed intelligence framework with mechanical design, humanoid robots can achieve improved safety margins, higher reliability, and sustained operational availability. Such sensor-oriented mechatronic design is essential for long-term autonomous operation and for humanoid robots to function as dependable, everyday partners in human-centered environments.

D. Mechanical Design as an Enabler of Embodied Intelligence

In humanoid robotics, mechatronic design is not only concerned with structural integrity or motion generation, but fundamentally shapes how intelligence emerges through interaction with the physical world. Embodied intelligence arises from the tight coupling between mechanical structure, sensing, actuation, control, and learning, whereby perception, action, and adaptation are continuously grounded in physical interaction with the environment. Mechanical design choices directly influence what a humanoid robot can perceive, how it can safely explore, and how effectively it can learn from experience and recover [29].

Equally important, mechanical and control-based compliance play a critical role beyond safety alone. Compliance enables stable physical interaction, reduces the consequences of modeling uncertainty, and allows robots to safely explore contact-rich environments [30], conditions that are widely recognized as prerequisites for autonomous learning and long-term adaptation. Without appropriate compliance, exploration becomes hazardous, and learning severely constrained.

From this perspective, mechanical design becomes inseparable from higher-level cognitive capabilities [31]. Early codesign of sensing, electronics, and compliant actuation establishes the physical conditions under which perception, learning, and decision-making can operate reliably. As humanoid robots evolve toward continuous operation alongside humans, mechatronic design must therefore be understood as a foundational enabler of intelligence, autonomy, and dependable human–robot collaboration.

III. Multimodal Interaction

To transform humanoid robots from performance-driven demonstrators into reliable daily partners, interaction must be formulated as a *closed-loop systems problem* rather than a superficial interface layer appended after mechanical and control design [32]. In human-centered environments, humanoid robots must: (i) possess sufficient multimodal observability to perceive both the world and the interaction process itself; (ii) infer action-relevant human affect and intent to avoid responses that are mechanically correct yet socially inappropriate; (iii) coordinate through multimodal communication so that their behavior is legible, predictable, and interruptible; and (iv) maintain safety in both physical and psychosocial dimensions [33], [34]. An overview of the affect-driven multimodal interaction pipeline is illustrated in Fig. 2.

Traditional humanoids have been developed within a primarily mechanical paradigm emphasizing motion control and stability [4], [5]. Although this paradigm has yielded remarkable locomotion and manipulation performance, current humanoids still lack emotional awareness and social understanding. Their interactions often appear “cold”: they fail to interpret human emotions, overlook subtle behavioral cues, and cannot adjust their responses in real time as human states evolve [9], [15]. These limitations hinder natural, safe, and socially coherent interaction in human-centered spaces.

Building on the embodied structures and embedded sensing foundations established in Section II, this section discusses multimodal perception, affective state inference, multimodal interaction and intention communication, and predictive safety in close human–robot coexistence [35], [36].

A. Multimodal Perception for Human–Robot Interaction

Vision has long served as the dominant perceptual modality in humanoid robotics due to its dense spatial information and mature learning pipelines [36]. However, vision alone is fragile under occlusion, illumination changes, and viewpoint constraints. Moreover, visual information provides limited access to variables crucial for daily interaction, such as contact state, interaction forces, and subtle cues reflecting human comfort, hesitation, or responsiveness. A true daily partner must understand not only *what* is occurring in the physical environment, but also *how* interaction unfolds and evolves over time [35].

Multimodal perception integrates heterogeneous sensing channels to address these limitations [36]. These channels include proprioception (IMU, joint encoders, motor currents), force/torque sensing, tactile arrays and electronic skin, proximity sensors, acoustic cues, thermal feedback, and emerging chemical/olfactory sensing. The primary challenge lies not in having many sensors, but in fusing fragmented and incomplete sensory streams into coherent estimates of contact configuration, slip, external disturbances, human motion intent, and scene dynamics. While learning-based multimodal representations can improve robustness, deployment-grade systems require explicit uncertainty handling, out-of-distribution detection, and failure-mode awareness to avoid brittle or opaque behavior in everyday environments [37], [38].

To support emotion- and interaction-oriented behavior, multimodal perception must evolve from traditional task- or geometry-centered approaches toward *affect- and interaction-centered* frameworks [35]. In this perspective, perception captures not only objects and motion but also dynamic emotional cues and social states that unfold throughout interaction. Beyond visual features, prosodic signals, tactile contact patterns, body posture dynamics, interpersonal distance, and environmental context all serve as meaningful indicators [39].

An important trend is whole-body somatosensation. Distributed tactile sensing and electronic skin transform the robot’s body into an active perceptual substrate, enabling earlier detection of unintended contact, faster compliance modulation, and safer near-field interaction [10], [40]. Combined with force, thermal, and inertial sensing, whole-body perception supports both physical safety and inference of user comfort, physiological state, and interaction fluency [41], [42]. Looking forward, richer human-like modalities—such as chemical or olfactory cues—may further expand situational awareness by detecting hazards or environmental changes imperceptible to vision [11].

In this sense, multimodal perception forms the perceptual substrate for affective inference, intention communication, and safety decision-making [35], [36]. Only with stable, comprehensive perceptual estimates can downstream emotional understanding and interactive control be reliable.

B. Affective State Inference for Behavior Modulation

Daily partnership requires not only geometric and physical understanding but also awareness of human emotional, psychological, and social states [9], [35]. Emotion recognition

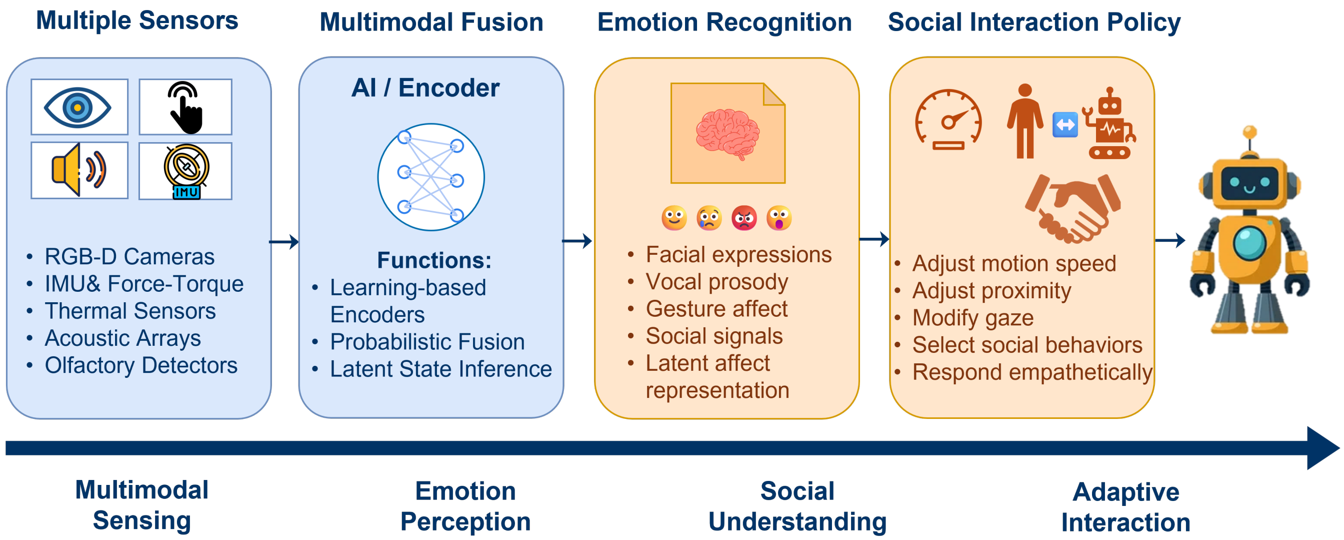

Fig. 2. Overview of the affect-driven multimodal interaction pipeline. Multiple heterogeneous sensors provide complementary information which is fused through learning-based and probabilistic multimodal models. Emotion recognition extracts latent affective variables from facial, vocal, gestural, and social signals. These affective estimates modulate the social interaction policy, enabling adaptive and socially coherent robot behavior.

in humanoid robotics is therefore better framed as *affective state inference*, as facial appearance alone is often ambiguous and natural interaction exhibits highly multimodal emotional expression. Affect emerges through prosody and speaking style, body posture and motion dynamics, gaze direction, interpersonal distance, interaction timing, and, in collaborative tasks, contact and force patterns [39], [35].

To endow robots with emotional awareness and social competence, emotion inference must move beyond single-modality recognition toward *affect-oriented multimodal perception frameworks* [35], [36]. In such frameworks, emotion recognition is not an isolated facial-expression classifier but a latent-variable estimation problem grounded in fused multimodal cues. Multimodal architectures—combining RGB-D vision, microphone arrays, tactile and electronic-skin sensing, inertial measurement, thermal cues, and physiological proxies—enable inference of human stress, uncertainty, comfort, urgency, engagement, and other socially relevant affective variables.

From a systems perspective, the most meaningful affective variables are those that are interpretable and actionable. These variables can regulate speed profiles, interpersonal distance, compliance levels, and clarification strategies in motion and dialogue policies [32], [13]. In this regard, affect inference is not an aesthetic feature but a mechanism that reduces intent misalignment, prevents startling behavior, and facilitates timely communication repair.

Learning-based multimodal fusion and affect inference—using deep encoders, temporal models, and probabilistic estimators—map raw perceptual streams to coherent affective representations [36], [35]. Robots can thus identify socially meaningful patterns and anticipate how interaction may evolve. This enables real-time modulation of behavior, such as adjusting speed and proximity, shifting gaze and posture, altering vocal tone, or reorganizing turn-taking strategies to maintain social appropriateness and psychological comfort.

Two deployment constraints arise. First, affect inference must be robust across different users, contexts, and cultures, motivating personalization, temporal modeling, and calibration. Second, affective signals can be sensitive and easily misinterpreted; systems must incorporate privacy protections, bias awareness, conservative confidence estimation, and graceful degradation strategies such as “ask-before-act” behaviors [12].

C. Coordination and Intention Communication in Multimodal Interaction

Multimodal interaction is not the mere accumulation of speech, gestures, and displays; it is the design of coordination mechanisms that ensure robot behavior is legible, predictable, and interruptible, and that align with human expectations in shared spaces [33], [32]. Practically, this requires robots to communicate intent and constraints at human-interpretable granularity, respect human response delays and turn-taking rhythms, and support safe interruption, correction, and resumption. Over time, systems should adapt to user preferences—comfort distance, acceptable speed, preferred explanation style—reducing unnecessary confirmation while preserving safety [34].

Different channels serve complementary roles: speech and text convey plans, uncertainty, and confirmation requests; gaze and head orientation provide lightweight cues regarding attention and direction; gestures support deictic reference; and haptic feedback is essential in handovers and physical assistance [9]. For daily use, interaction design should prioritize clarity and consistency rather than anthropomorphic “social performance.” Consistent motion primitives, stable confirmation protocols, and explicit action boundaries often matter more for trust than expressive fidelity [12].

Affective inference directly modulates interaction policies. Perceived hesitation may trigger slower motion and explicit confirmation; perceived stress may reduce verbosity and prioritize safety-critical actions; discomfort cues may increase distance and promote gentler trajectories. The goal is not to emulate emotion, but to regulate behavior in ways that sustain human comfort and predictability in long-term coexistence [13], [32].

D. Predictive Safety in Close Human–Robot Interaction

Safety in everyday environments cannot rely solely on force limits, collision detection, or emergency stops [32]. While reactive mechanisms are essential, they are insufficient for long-term coexistence, as discomfort and loss of trust often arise before physical thresholds are violated. Purely physical safety metrics also fail to capture socially unsafe behaviors—such as abrupt motions, intrusive proximity, or poorly timed actions—that may violate human expectations even when speeds and forces remain within limits [34].

A more appropriate framing treats safety as a *predictive, interaction-aware system property*. Multimodal perception provides early indicators of risk—proximity dynamics, posture shifts, gaze changes, vocal stress, hesitation—enabling the robot to proactively slow down, increase distance, replan trajectories, or seek confirmation before unsafe contact or social violation occurs [32], [43]. During contact, tactile sensing and whole-body compliance allow rapid redistribution of forces, enabling safer physical interaction without requiring immediate task termination [44].

Safety encompasses both physical and psychosocial dimensions. Behavior that is physically harmless may still be socially unacceptable; conversely, affect-aware regulation—maintaining appropriate distance, avoiding abrupt motions, providing explanations under uncertainty—can greatly improve perceived safety [34], [12]. Robots capable of detecting fear, hesitation, or confusion can adapt their behavior to maintain comfort and social appropriateness.

The convergence of multimodal sensing, embodied AI, and affective modeling defines a unified path for future humanoid safety: systems that not only avoid harm but actively support trust, comfort, and well-being [2], [32]. Transparency, predictability, and interruptibility thus function as safety mechanisms, enabling humans to coordinate with the robot's safety envelope rather than passively endure its actions. An overview of the proposed predictive safety pipeline for close human–robot interaction is illustrated in Fig. 3

IV. Robotic Intelligence and Learning

Advances in mechatronics, sensing, and interaction provide the physical and perceptual foundations for humanoid robots, but intelligence ultimately determines whether these systems can reason, adapt, and operate autonomously in daily environments [7]. Classical robotic intelligence has largely relied on task-specific models, manually engineered state machines, and narrowly scoped learning pipelines [45]. While effective in controlled settings, these approaches often struggle to generalize across tasks, users, and long time horizons. The emergence of large language models (LLMs) and multimodal foundation models introduces a plausible path toward more general, compositional autonomy by enabling abstraction, tool use, and high-level planning—provided that such reasoning is grounded in perception and constrained by safety and verification [2], [46].

A. From Task Execution to Cognitive Mediation

LLMs offer capabilities beyond language processing, including contextual inference, symbolic-like reasoning, and long-horizon planning [47], [48]. In humanoid systems, LLMs are best viewed as *cognitive mediators* rather than low-level controllers: they translate goals and incomplete instructions into structured task representations, invoke tools (perception modules, planners, skill libraries), and coordinate sub-skills under constraints [46].

This separation aligns naturally with embodied robotic architectures. Low-level controllers manage dynamics, stability, compliance, and physical contact, while LLM-centered modules operate at a higher level over goals, constraints, and interaction context [49]. Such hierarchical organization is particularly important in human environments, where task specifications are frequently ambiguous, underspecified, and communicated implicitly through interaction [7].

B. Vision–Language–Action Models and Multimodal Grounding

For humanoid robots, the effectiveness of LLM-driven intelligence depends critically on grounding reasoning in perception and action [48]. Recent progress in *vision–language–action* (VLA) models provides a concrete mechanism for such grounding by unifying visual observations, language instructions, and action representations within a single learning framework. A representative example is RT-2, which casts robot actions as tokens and co-fine-tunes a vision–language backbone on both web-scale vision–language data and robot trajectories, yielding improved generalization and emergent semantic behaviors [50]. More recent humanoid-oriented efforts, such as Figure's Helix, explicitly target high-rate upper-body control while maintaining a "fast/slow" reasoning structure to balance reactive dexterity with deliberative planning.

These developments suggest a practical direction: instead of treating language understanding and control as separate pipelines, humanoid autonomy can be built around unified multimodal representations that link semantic intent to physical affordances (e.g., graspability, stability under contact, reachable poses) and to interaction constraints (e.g., comfort distance, interruption) [51], [52]. However, VLA-style policies also expose deployment challenges, including distribution shift, long-horizon compounding errors, and the need for calibrated uncertainty estimates when perception is incomplete or ambiguous [7], [53].

C. Agent-Based Learning, Tool Use, and Constrained Adaptation

Beyond interpreting tasks, LLMs enable *agentic* robotic systems that can decompose goals, select tools, evaluate outcomes, and revise strategies [54], [55]. In practice, this

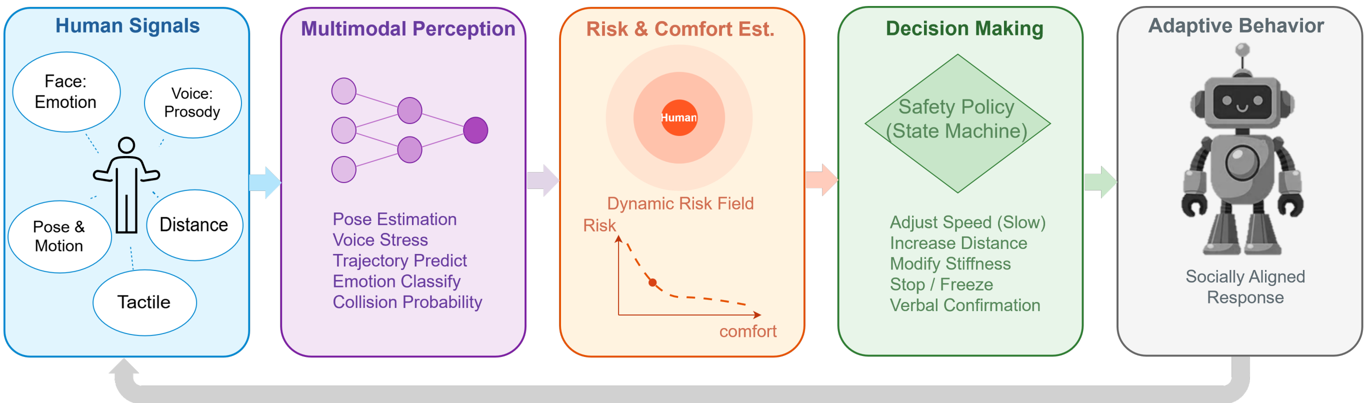

Fig. 3. Predictive safety pipeline for close human–robot interaction. Human signals such as facial emotion, prosody, body motion, distance, and tactile cues are integrated through multimodal perception to estimate dynamic risk and comfort levels. These estimates drive safety-aware decision making—including speed adjustment, distance regulation, stiffness modification, emergency stopping, and verbal confirmation—resulting in socially aligned and context-appropriate robot behavior.

agent layer can orchestrate planning, perception queries, skill invocation, memory retrieval, and self-checking routines [51], [56]. DeepMind's Gemini Robotics 1.5 and related embodied-reasoning releases further illustrate the trend toward "thinking while acting," tool use, and transfer across embodiments as ingredients for more general physical agents.

For daily deployment, however, autonomy and learning must remain bounded. Rather than unrestricted self-modification, practical humanoid systems require *constrained adaptation*: learning and policy updates occur within well-defined operational envelopes, with verification gates (simulation replay, regression tests), audit trails (model/version logging, decision traces), and runtime safety filters (action constraints, force/velocity limits, safe-stop behaviors) [57], [58]. When combined with multimodal sensing and affect-aware interaction, agent-based learning can improve behavior over time while preserving predictability, accountability, and user trust [59], [60].

D. Toward Long-Term Autonomy and Coexistence

From a perspective viewpoint, LLM- and VLA-driven intelligence marks a shift from pre-scripted functionality toward experience-driven autonomy [55], [46]. For humanoid robots to become daily partners, intelligence must support long-horizon operation, personalization to individual users, and responsiveness to changing environments—not only through better models, but through careful system integration that respects embodiment, safety, and human-centered interaction [61], [59].

LLM-based reasoning should therefore be treated as a complementary layer atop established control and learning methods [62], [56]. When grounded in multimodal perception and embodied interaction, and constrained by verification, auditability, and runtime safety envelopes, this cognitive layer provides a pathway for humanoid robots to learn from experience, adapt to human needs, and evolve over extended periods of deployment [55], [46].

V. Open Challenges and Future Directions

Despite significant progress across mechatronics, perception, interaction, and intelligence, humanoid robotics remains far from achieving widespread deployment as daily partners in human environments [55], [61]. The transition from engineering showcases to sustained, real-world operation exposes a set of open challenges that are technical, systemic, and societal in nature. Addressing these challenges requires coordinated advances across disciplines rather than isolated breakthroughs.

A. Robustness, Reliability, and Long-Term Operation

One of the most critical barriers to daily deployment is robustness over extended time horizons [61]. Many humanoid robots demonstrate impressive performance in short-duration experiments, yet struggle with long-term reliability due to mechanical wear, sensor degradation, thermal effects, and cumulative control errors [63]. Daily partners must operate continuously, tolerate uncertainty, and recover gracefully from minor faults without constant expert intervention.

Future research must prioritize durability, fault tolerance, and self-monitoring at both the hardware and software levels [64], [61]. This includes redundancy in sensing and actuation, health-aware control strategies, and diagnostic mechanisms that enable humanoid robots to assess their own operational state during prolonged use.

B. Energy Efficiency and Autonomy

Energy consumption remains a fundamental constraint for humanoid robots operating outside laboratory settings [65]. High power demands from actuation, computation, and sensing limit operational autonomy and constrain practical use in daily environments. While improvements in batteries and power electronics contribute incrementally, energy efficiency must be addressed holistically through lightweight design, compliant actuation, energy-aware control, and task-level optimization [52].

Long-term autonomy also depends on intelligent energy management, including adaptive scheduling of tasks, selective activation of sensing and computation, and coordination between physical activity and cognitive load [61]. Energy-aware intelligence will be essential for humanoid robots to function reliably over extended periods.

C. Learning, Adaptation, and Safety Guarantees

The integration of learning-based methods and LLM-driven reasoning raises important questions regarding safety, predictability, and verification [56], [62]. While adaptive behavior is essential for functioning in unstructured environments, unrestricted learning may introduce unintended behaviors or safety risks. Daily partners must therefore balance adaptability with bounded autonomy.

Future systems will require mechanisms for constraining learning within certified safety envelopes, as well as methods for monitoring and validating learned behaviors over time [57], [58]. Human-in-the-loop supervision, explainable decision-making, and conservative fallback strategies are likely to play key roles in ensuring safe long-term adaptation [60], [59].

D. Human Acceptance and Societal Integration

Beyond technical considerations, the success of humanoid robots as daily partners ultimately depends on human acceptance [66], [59]. Trust, comfort, and perceived usefulness are shaped not only by system capability, but also by behavior consistency, transparency, and alignment with social norms. Overly anthropomorphic design or exaggerated expectations may hinder acceptance rather than promote it.

Interdisciplinary research involving robotics, psychology, ergonomics, and social sciences is essential for understanding how humans perceive and interact with humanoid robots in daily contexts [67], [58]. Ethical considerations, privacy, responsibility, and accountability must be addressed proactively to ensure that technological progress aligns with societal values.

Together, these challenges highlight that the path toward daily humanoid partners is evolutionary rather than abrupt. Progress will emerge from sustained integration of engineering rigor, embodied intelligence, and human-centered design principles [55].

VI. Conclusion

Humanoid robotics is entering a pivotal phase in which engineering achievements alone are no longer sufficient to define progress. While recent advances in locomotion, manipulation, and artificial intelligence have enabled increasingly impressive demonstrations, the realization of humanoid robots as dependable daily partners remains an open and complex challenge. Bridging this gap requires a fundamental shift from performance-driven development toward systems that are embodied, perceptive, adaptive, and socially aware.

This article has presented a perspective on the key foundations and directions necessary for this transition. We have argued that mechatronic co-design, integrated and multimodal sensing, affect-aware perception, interaction-centered safety, and large language model–driven intelligence are not independent research themes, but tightly coupled components of a unified humanoid system. Only through their coordinated integration can humanoid robots move beyond staged showcases toward sustained operation in human-centered environments.

Rather than seeking to replicate human form or behavior for its own sake, future humanoid robots should be designed to complement human capabilities, respect human boundaries, and adapt to individual needs over time. In this sense, becoming a daily partner is less about achieving human likeness and more about earning trust through reliability, predictability, and responsiveness. As research continues to mature, humanoid robotics offers the opportunity to redefine how intelligent machines coexist with humans. The challenge ahead lies not in demonstrating what humanoid robots can do in controlled settings, but in determining how they can responsibly, safely, and meaningfully integrate into the fabric of everyday life.

References

[1] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[2] A. Billard, A. Albu-Schaeffer, M. Beetz, W. Burgard, P. Corke, M. Ciocarlie, R. Dahiya, D. Kragic, K. Goldberg, Y. Nagai *et al.*, "A roadmap for ai in robotics," *Nature Machine Intelligence*, pp. 1–7, 2025.

[3] 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.

[4] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, T. Koolen, P. Marion, and R. Tedrake, "Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," *Autonomous robots*, vol. 40, no. 3, pp. 429–455, 2016.

[5] G. Li and R. Gomez, "Realizing full-body control of humanoid robots," *Nature Machine Intelligence*, vol. 6, no. 9, pp. 990–991, 2024.

[6] I. Radosavovic, T. Xiao, B. Zhang, T. Darrell, J. Malik, and K. Sreenath, "Real-world humanoid locomotion with reinforcement learning," *Science Robotics*, vol. 9, no. 89, p. eadi9579, 2024.

[7] L. Kunze, N. Hawes, T. Duckett, M. Hanheide, and T. Krajník, "Artificial intelligence for long-term robot autonomy: A survey," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4023–4030, 2018.

[8] A. Cully, J. Clune, D. Tarapore, and J.-B. Mouret, "Robots that can adapt like animals," *Nature*, vol. 521, no. 7553, pp. 503–507, 2015.

[9] C. Breazeal, "Toward sociable robots," *Robotics and autonomous systems*, vol. 42, no. 3-4, pp. 167–175, 2003.

[10] C. M. Boutry, M. Negre, M. Jorda, O. Vardoulis, A. Chortos, O. Khatib, and Z. Bao, "A hierarchically patterned, bioinspired e-skin able to detect the direction of applied pressure for robotics," *Science Robotics*, vol. 3, no. 24, p. eaau6914, 2018.

[11] W. Wang, Y. Jiang, D. Zhong, Z. Zhang, S. Choudhury, J.-C. Lai, H. Gong, S. Niu, X. Yan, Y. Zheng *et al.*, "Neuromorphic sensorimotor loop embodied by monolithically integrated, low-voltage, soft e-skin," *Science*, vol. 380, no. 6646, pp. 735–742, 2023.

[12] L. Cominelli, F. Feri, R. Garofalo, C. Giannetti, M. A. Meléndez-Jiménez, A. Greco, M. Nardelli, E. P. Scilingo, and O. Kirchkamp, "Promises and trust in human–robot interaction," *Scientific reports*, vol. 11, no. 1, p. 9687, 2021.

[13] P. Slade, C. Atkeson, J. M. Donelan, H. Houdijk, K. A. Ingraham, M. Kim, K. Kong, K. L. Poggensee, R. Riener, M. Steinert *et al.*, "On human-in-the-loop optimization of human–robot interaction," *Nature*, vol. 633, no. 8031, pp. 779–788, 2024.

[14] R. Mon-Williams, G. Li, R. Long, W. Du, and C. G. Lucas, "Embodied large language models enable robots to complete complex tasks in unpredictable environments," *Nature Machine Intelligence*, pp. 1–10, 2025.

[15] D. Herath, J. Busby Grant, A. Rodriguez, and J. L. Davis, "First impressions of a humanoid social robot with natural language capabilities," *Scientific Reports*, vol. 15, no. 1, p. 19715, 2025.

[16] L. C. Visser, R. Carloni, and S. Stramigioli, "Variable stiffness actuators: A port-based analysis and a comparison of energy efficiency," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, p. 3279–3284. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5509127/

[17] S. Wolf, G. Grioli, O. Eiberger, W. Friedl, M. Grebenstein, H. Höppner, E. Burdet, D. G. Caldwell, R. Carloni, and M. G. Catalano, "Variable stiffness actuators: Review on design and components," *IEEE/ASME transactions on mechatronics*, vol. 21, no. 5, p. 2418–2430, 2015.

[18] S. V. Sarkisian, L. Gabert, and T. Lenzi, "Series-elastic actuator with two degree-of-freedom pid control improves torque control in a powered knee exoskeleton," *Wearable Technologies*, vol. 4, p. e25, Jan. 2023.

[19] D. W. Robinson, "Design and analysis of series elasticity in closed-loop actuator force control," PhD Thesis, Massachusetts Institute of Technology, 2000. [Online]. Available: https://dspace.mit.edu/bitstream/handle/1721.1/54838/45993503-MIT.pdf?sequence=2

[20] G. Pratt and M. Williamson, "Series elastic actuators," in *Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*, vol. 1, Aug. 1995, p. 399–406 vol.1. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/525827

[21] G. N. Sahu, S. Singh, A. Singh, and M. Law, "Static and dynamic characterization and control of a high-performance electro-hydraulic actuator," in *Actuators*, vol. 9, no. 2. MDPI, 2020, p. 46. [Online]. Available: https://www.mdpi.com/2076-0825/9/2/46

[22] S. AlFayad, "Hydraulic actuator with overpressure compensation," 2024. [Online]. Available: https://patents.google.com/patent/US12012947B2

[23] S. Qu, D. Fassbender, A. Vacca, and E. Busquets, "A high-efficient solution for electro-hydraulic actuators with energy regeneration capability," *Energy*, vol. 216, p. 119291, Feb. 2021.

[24] A. Albers, J. Ottnad, H. Weiler, and P. Haeussler, "Methods for lightweight design of mechanical components in humanoid robots," in *2007 7th IEEE-RAS International Conference on Humanoid Robots*, Nov. 2007, p. 609–615. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4813934

[25] A. J. Fuge, C. W. Herron, B. C. Beiter, B. Kalita, and A. Leonessa, "Design, development, and analysis of the lower body of next-generation 3d-printed humanoid research platform: Pandora," *Robotica*, vol. 41, no. 7, p. 2177–2206, 2023.

[26] P. M. Ferreira, M. A. Machado, M. S. Carvalho, and C. Vidal, "Embedded sensors for structural health monitoring: Methodologies and applications review," *Sensors*, vol. 22, no. 21, p. 8320, Jan. 2022.

[27] G. Stano, S. M. A. I. Ovy, J. R. Edwards, M. Cianchetti, G. Percoco, and Y. Tadesse, "One-shot additive manufacturing of robotic finger with embedded sensing and actuation," *The International Journal of Advanced Manufacturing Technology*, vol. 124, no. 1–2, p. 467–485, Jan. 2023.

[28] D. Chen, Z. Han, J. Zhang, L. Xue, and S. Liu, "Additive manufacturing provides infinite possibilities for self-sensing technology," *Advanced Science*, vol. 11, no. 28, p. 2400816, 2024.

[29] J. Bongard, V. Zykov, and H. Lipson, "Resilient machines through continuous self-modeling," *Science*, vol. 314, no. 5802, p. 1118–1121, Nov. 2006.

[30] A. Bicchi and G. Tonietti, "Dealing with the safety-performance tradeoff in robot arms design and control," *IEEE Robotics and Automation Magazine*, vol. 11, no. 2, p. 22–33, 2004.

[31] A. Cangelosi, J. Bongard, M. H. Fischer, and S. Nolfi, *Embodied Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, p. 697–714. [Online]. Available: http://link.springer.com/10.1007/978-3-662-43505-2_37

[32] P. A. Lasota, T. Fong, J. A. Shah *et al.*, "A survey of methods for safe human-robot interaction," *Foundations and Trends® in Robotics*, vol. 5, no. 4, pp. 261–349, 2017.

[33] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2013, pp. 301–308.

[34] J. Rios-Martinez, A. Spalanzani, and C. Laugier, "From proxemics theory to socially-aware navigation: A survey," *International Journal of Social Robotics*, vol. 7, no. 2, pp. 137–153, 2015.

[35] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Information fusion*, vol. 37, pp. 98–125, 2017.

[36] T. Baltrušaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 2, pp. 423–443, 2018.

[37] A. Kendall and Y. Gal, "What uncertainties do we need in bayesian deep learning for computer vision?" *Advances in neural information processing systems*, vol. 30, 2017.

[38] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," *Advances in neural information processing systems*, vol. 30, 2017.

[39] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. S. Narayanan, "Iemocap: Interactive emotional dyadic motion capture database," *Language resources and evaluation*, vol. 42, no. 4, pp. 335–359, 2008.

[40] S. Sundaram, P. Kellnhofer, Y. Li, J.-Y. Zhu, A. Torralba, and W. Matusik, "Learning the signatures of the human grasp using a scalable tactile glove," *Nature*, vol. 569, no. 7758, pp. 698–702, 2019.

[41] J. Ge, X. Wang, M. Drack, O. Volkov, M. Liang, G. S. Cañón Bermúdez, R. Illing, C. Wang, S. Zhou, J. Fassbender *et al.*, "A bimodal soft electronic skin for tactile and touchless interaction in real time," *Nature communications*, vol. 10, no. 1, p. 4405, 2019.

[42] Y. Yan, A. Zermane, J. Pan, and A. Kheddar, "A soft skin with self-decoupled three-axis force-sensing taxels," *Nature Machine Intelligence*, vol. 6, no. 11, pp. 1284–1295, 2024.

[43] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2016.

[44] S. Haddadin, A. Albu-Schaffer, A. De Luca, and G. Hirzinger, "Collision detection and reaction: A contribution to safe physical human-robot interaction," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 3356–3363.

[45] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238–1274, 2013.

[46] R. Mon-Williams *et al.*, "ELLMER: Embodied large language model for robotics," *Nature Machine Intelligence*, vol. 7, pp. 247–255, 2025.

[47] J. Liang, W. Huang, F. Xia, P. Xu, K. Hausman, B. Ichter, P. Florence, and A. Zeng, "Code as policies: Language model programs for embodied control," *arXiv preprint arXiv:2209.07753*, 2022.

[48] D. Driess, F. Xia, M. S. Sajjadi, C. Lynch, A. Chowdhery, A. Wahid, J. Tompson, Q. Vuong, T. Yu, W. Huang *et al.*, "Palm-e: An embodied multimodal language model," 2023.

[49] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[50] B. Zitkovich, T. Yu, S. Xu, P. Xu, T. Xiao, F. Xia, J. Wu, P. Wohlhart, S. Welker, A. Wahid *et al.*, "Rt-2: Vision-language-action models transfer web knowledge to robotic control," in *Conference on Robot Learning*. PMLR, 2023, pp. 2165–2183.

[51] C. R. Garrett, T. Lozano-Pérez, and L. P. Kaelbling, "Integrated task and motion planning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, pp. 265–293, 2021.

[52] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[53] S. Liang, Y. Li, and R. Srikant, "Enhancing the reliability of out-of-distribution image detection in neural networks," in *International Conference on Learning Representations (ICLR)*, 2018.

[54] J. Liang *et al.*, "Code as policies: Language model programs for embodied control," 2022.

[55] A. Billard *et al.*, "A roadmap for AI in robotics," *Nature Machine Intelligence*, vol. 7, pp. 818–824, 2025.

[56] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial Intelligence*, vol. 101, no. 1–2, pp. 99–134, 1998.

[57] A. D. Ames *et al.*, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2017.

[58] P. A. Lasota, T. Fong, and J. A. Shah, "A survey of methods for safe human–robot interaction," *Foundations and Trends in Robotics*, vol. 5, no. 4, pp. 261–349, 2017.

[59] L. Cominelli, F. Feri, R. Garofalo, C. Giannetti, M. A. Melendez-Jimenez, A. Greco, M. Nardelli, E. P. Scilingo, and O. Kirchkamp, "Promises and trust in human–robot interaction," *Scientific Reports*, vol. 11, 2021.

[60] P. Slade *et al.*, "On human-in-the-loop optimization of human–robot interaction," *Nature*, vol. 633, pp. 779–788, 2024.

[61] L. Kunze *et al.*, "Artificial intelligence for long-term robot autonomy: A survey," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4023–4030, 2018.

[62] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238–1274, 2013.

[63] I. Radosavovic *et al.*, "Real-world humanoid locomotion with reinforcement learning," *Science Robotics*, vol. 9, p. eadi9579, 2024.

[64] A. Cully, J. Clune, D. Tarapore, and J.-B. Mouret, "Robots that can adapt like animals," *Nature*, vol. 521, pp. 503–507, 2015.

[65] S. H. Collins, A. Ruina, R. Tedrake, and M. Wisse, "Efficient bipedal robots based on passive-dynamic walkers," *Science*, vol. 307, no. 5712, pp. 1082–1085, 2005.

[66] C. Breazeal, "Toward sociable robots," *International Journal of Human-Computer Studies*, vol. 59, no. 1–2, pp. 119–155, 2003.

[67] D. Herath, J. B. Grant, A. Rodriguez, and J. L. Davis, "First impressions of a humanoid social robot with natural language capabilities," *Scientific Reports*, vol. 15, p. 19715, 2025.

Samer AlFayad is currently a full professor in Mechatronics Systems at UEVE and Paris-Saclay University. From 2011 to 2019, he was an associate professor in humanoid robotics at UVSQ, holding an industrial excellence chair in hydraulic domestication at UVSQ & BIA company. From 2010 to 2011, he was a Post-doc researcher at TUM-Germany with an Alexander Von Humboldt Foundation scholarship. He received his Ph.D. in robotics from UVSQ in 2009. He was awarded the best Ph.D. thesis in robotics by the French CNRS in 2010 and the best Ph.D. thesis in mechatronics for the 20th anniversary of UVSQ. He has been an investigator of several French national projects. He is the leader of the mechatronics research group. He was the team leader of the Paris-Saclay team for Mohamed Bin Zayed International Robotics Challenge MBZIRC2017. He is co-coordinator of the Institute for Control and Decision (iCode Paris-Saclay). He is co-coordinator of the Human-in-the-loop Control and Decision Institute (H-CoDe Paris-Saclay). He published 51 international patents. He obtained 4 M€ funding during the last five years.

Maya Sleiman is currently CTO of Kalysta Actuation in Évry, France, where she leads mechatronic system development. Since 2024, she is also an associated researcher at the IBISC Laboratory (Université d'Évry Paris-Saclay). She received her Ph.D. in Robotics with jury congratulations from Université Paris-Saclay in 2022, focusing on new actuation technologies for humanoid robotics and assistive devices. From 2022 to 2023, she was a postdoctoral researcher at IBISC, working on lightweight electro-hydraulic actuators for safe human–robot interaction. Her earlier industrial experience includes project engineering roles in France and the Middle East in large-scale mechanical and structural projects. She is an author and co-author of multiple international journal articles and conference papers in humanoid robotics and hydraulic actuation, and co-inventor of several international patents in electro-hydraulic actuators

Feilong Wang has over four years of research experience in humanoid robot development and control. He is currently pursuing an M2 degree in Mechatronics, Machine Vision, and Artificial Intelligence at Université Paris-Saclay, France. His research interests focus on sensor-fusion-based teleoperation, humanoid robot control, and adaptive strategies for human–robot interaction. He is currently involved in a French ANR funded project entitled "Evolutionary Exoskeleton for Teenagers with Reduced Mobility". He has authored multiple first-author publications in international journals and conferences, and his work was recognized as one of the Top 10 Technological Innovations at the World Robot Competition 2023. He has strong research interests in wearable robotics and embodied intelligence.

Hang Su is currently a professor at Paris-Saclay University, France. He obtained his Ph.D. degree from Politecnico di Milano in Italy in 2019. Additionally, he serves as the Associate Editor of the IEEE SMC Society eNewsletter. His editorial roles include Associate Editor positions for prestigious IEEE journals such as the Transactions on Automation Science and Engineering (IEEE TASE) and the Transactions on Cognitive and Developmental Systems (IEEE TCDS), starting in 2022. He also contributes as an Associate Editor to leading conferences in the robotics field, including the IEEE International Conference on Robotics and Automation (ICRA) and the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), roles he has held for over last five years. He has been invited to deliver a talk at an upcoming Nature Conferences event on robotics. He has published over 120 papers in international journals and conferences. He was a recipient of the Outstanding Interaction Paper Award at the IEEE International Conference on Robotics and Automation 2022 (ICRA 2022), the 2021 Andrew P. Sage Best Transactions Paper Award on IEEE Transactions on Human-Machine Systems (IEEE THMS), the Best Conference Paper Award in Advanced Robotics at the IEEE International Conference on Advanced Robotics and Mechatronics in 2020, and the ICRA Travel Award funded by the IEEE Robotics and Automation Society in 2019. His main research interests include human-robot interaction, medical robotics, exoskeleton, sensor fusions, bilateral teleoperation, etc.