# NEXT-GENERATION AUTONOMOUS MACHINES

## APPLICATIONS AND IMPACTS OF PHYSICAL AI

AHMET DAYANÇ

FERİDUN KARAKOÇ

SERÜVEN
YAYINEVİ

**Genel Yayın Yönetmeni / Editor in Chief** • C. Cansın Selin Temana

**Kapak & İç Tasarım / Cover & Interior Design** • Serüven Yayınevi

**Birinci Basım / First Edition** • © Aralık 2024

**ISBN** • 978-625-5552-20-4



**Serüven Yayınevi / Serüven Publishing**
**Türkiye Adres / Turkey Address:** Kızılay Mah. Fevzi Çakmak 1. Sokak
Ümit Apt No: 22/A Çankaya/ANKARA
**Telefon / Phone:** 05437675765
**web:** www.seruvenyayinevi.com
**e-mail:** seruvenyayinevi@gmail.com

**Baskı & Cilt / Printing & Volume**
Sertifika / Certificate No: 47083

# NEXT-GENERATION AUTONOMOUS MACHINES: APPLICATIONS AND IMPACTS OF PHYSICAL AI

AHMET DAYANÇ[1]

FERİDUN KARAKOÇ[2]

1 Res. Asst., Kutahya Dumlupinar University, ahmet.dayanc@dpu.edu.tr, ORCID: 0000-0002-5214-9021
2 Asst. Prof. Dr., Kutahya Dumlupinar University, feridun.karakoc@dpu.edu.tr, ORCID: 0000-0002-6210-4070

SERÜVEN YAYINEVİ

**CONTENTS**

# Chapter 1:

## FOUNDATIONS OF PHYSICAL AI

Physical AI refers to the integration of AI systems into robotic platforms or tangible entities (such as physical objects) in a way that enables meaningful interactions with the physical world. The primary goal of Physical AI is not merely to employ intelligence in abstract computational processes, but also to ensure its direct engagement with the physical environment. Therefore, Physical AI differs from traditional AI applications: rather than limiting itself to analyzing and making inferences solely from datasets in virtual environments, it can also analyze data acquired from the physical world and transfer its outputs back into that environment. Through this characteristic, "Physical AI" sits at the intersection of multiple disciplines, including machine learning, robotics, mechatronics, biomimetics, materials science, computer science, and mechanical engineering. For instance, whereas traditional AI would analyze an image as purely digital data, a robot equipped with Physical AI can sense objects by touch, move them, navigate in a real room, interact with humans in the same physical space, and perceive and adapt to its environment through sensor feedback. In this manner, Physical AI equips robots with sensory and motor skills simultaneously, allowing them to perform versatile, human-like functions (Costeira & Lima, 2020).

Physical AI advances the development of both robotic systems and AI algorithms, enabling these systems to acquire much more complex, flexible, and human-like capabilities. On the one hand, it facilitates the ability of robotic arms, autonomous vehicles, and service robots to make decisions, adapt, and learn in response to changing environmental conditions. On the other hand, by processing sensor data, it drives forward the progress of AI models that interact with that data. The decision-making mechanisms of Physical AI may rely on raw data received from physical sensors, formulating motion plans based on that information. Furthermore, it can optimize its actions over time by drawing lessons from errors, thereby responding to the demands of dynamic environments.

It should be noted that Physical AI is not limited exclusively to digital or robotic systems; rather, there may be various types of systems in which Physical AI can be found, reflecting a wide range of possibilities. One such dimension involves bio-integration and the design of soft robotics and biohybrid systems inspired by the functioning of living organisms. The goal of these systems is to integrate one or more characteristics found in nature—such as adaptation, flexibility, energy efficiency, and resilience—into artificial systems, thereby creating more advanced examples of Physical AI. For instance, by replicating an octopus's flexible limbs in an artificial form, robots can benefit from this flexibility to pass through narrow spaces and move across irregular surfaces. Similarly, sensor systems inspired by living cells or biological tissues can grant AI a more natural and efficient environmental awareness (PgCert, 2024).

Physical AI opens new horizons in the field of Human-Robot Interaction (HRI). Currently, interaction with robots largely involves issuing commands and relying on predefined action sequences. With Physical AI, humanoid robots can respond to emotional expressions, body language, gestures, and changes in their surroundings, thereby rendering human-like interactions more natural and fluid. This has significant implications in caregiving, therapy, education, production line management, logistics, and many other fields—enabling humanoid robots to be employed in diverse scenarios while increasing their effectiveness and ability to address various needs (Universal Robots, 2019).

Moreover, beyond direct human interaction, Physical AI can find practical applications in numerous other sectors that affect human life, ranging from industrial manufacturing and autonomous agriculture to smart home systems and medical applications. For instance, on an industrial production line, a humanoid robot driven by Physical AI can perceive products in real time and dynamically integrate itself into the manufacturing process in place of a human operator, effectively managing that process (Automation.com, 2024). An agricultural robot could detect soil moisture, ambient temperature, and the presence of harmful insects, autonomously improving plant care and harvesting processes. In hospitals, surgical robots could process real-time feedback to perform delicate operations on tissues with high accuracy.

Naturally, there remain challenging issues to be resolved regarding Physical AI. These include cost, energy efficiency, durability, material quality, sensor sensitivity, precise motion control, data processing capacity, the continual development and optimization of AI models, and safety considerations. Additionally, the ethical and social dimensions of Physical AI are of paramount importance. Questions arise about the degree of autonomy these systems should have in their interactions with humans, the ethical norms to which they should be subject, and how their safe and beneficial usage for humanity can be ensured.

In conclusion, “Physical AI” denotes an interdisciplinary domain aimed at drawing AI out of an abstract layer and integrating it into real-world physical systems, robotic structures, and objects. This approach allows robots and other Physical AI systems to perceive, learn, adapt to, and interact more naturally with their surroundings. It is anticipated that, in the future, more flexible, capable, responsive, and resilient Physical AI systems will enable far-reaching transformations in social, industrial, and scientific fields.

The conceptual roots of intelligence reach back to philosophical investigations seeking to understand the human mind. In ancient Greek philosophy, particularly in the work of Aristotle, the effort to classify logical inference rules prompted a formal examination of “intelligence.” Aristotle developed the foundational principles of logic to provide a coherent framework for reasoning processes. His theory of “categories” classifies types of being and thereby organizes the objects of thought at a conceptual level. This approach is directly related to the mind’s capacity—what we call “reason”—to abstract and categorize. Aristotle’s development of syllogistic methods provides the mind with a systematic model for the pursuit of truth, thus clarifying the functional scope of “reason.”

According to Aristotle, the roots of knowledge begin with sensory perception; however, the generalizations obtained through induction from pure sense data rise to a conceptual level once the intellect is involved. This process positions human “intelligence” not only as dependent on sensory data but also as capable of drawing meaning from abstract concepts and generating principles. In this sense, intelligence functions at a higher mental level, molding raw data and deriving principles, definitions, and cause-effect relationships from it.

In Aristotle’s work De Anima, the mind is regarded as the highest faculty among the various layers of the soul found in living beings. Beyond vegetative and animal faculties, the intellect and intelligence unique to humans possess the power of abstract thinking and self-reflection. Here, the concept of entelecheia signifies the actualization and full operation of latent mental faculties, suggesting that intellect and intelligence are not passive receivers but faculties that develop, mature, and become active.

For Aristotle, the highest form of knowledge is sought in discovering the ultimate causes and principles of being. This endeavor calls upon the mind's power of abstraction and intelligence's capacity to grasp fundamental principles. From this standpoint, mind and intelligence are not merely technical tools but also instruments of a metaphysical inclination that seeks to uncover the nature of reality (Wikipedia, 2024).

At the foundation of artificial intelligence lies an interdisciplinary approach aimed at embedding certain aspects of human intelligence into machines. In the 20th century, the convergence of mathematics, logic, and cognitive sciences gave rise to the foundations of AI. The logic- and mathematics-based formalizations of philosopher-mathematicians such as Bertrand Russell and Alfred North Whitehead, along with Alonzo Church and Alan Turing's work on computability theory, provided a robust theoretical groundwork for AI (Mgmus, 2021). Turing's computability theory and his concept of the "Turing Machine", in particular, paved the way for defining machines capable of algorithmic computation, thereby delineating the theoretical boundaries of the AI discipline.

Pioneers—including John McCarthy, Marvin Minsky, Allen Newell, and Herbert Simon—developed symbolic AI (often referred to as "Good Old-Fashioned AI" or GOFAI), in which mental processes were modeled through symbols, logical rules, and formal languages (Petracca, 2021). These concepts enabled a structured examination of knowledge representation, symbolic logic, inductive and deductive reasoning, rule-based inference, expert systems, and problem-solving strategies under the umbrella of AI.

Nevertheless, the inadequacies of purely rule-based symbolic systems in handling uncertain, noisy, or incomplete information encouraged the adoption of statistical methods and probabilistic modeling approaches. In this context, probability theory, Bayes' theorem, Markov chains, and probabilistic graphical models (including Bayesian networks and Markov random fields) have enabled AI applications to make decisions, predictions, and inferences in uncertain environments (Almabetter, 2024). Statistical learning theory provides the theoretical framework necessary to analyze the general performance of data-driven models, assess their generalization capabilities, and formalize learning algorithms.

Within machine learning, and deep learning in particular, linear algebra (such as matrix multiplications, eigenvalues, and eigenvectors) and optimization techniques (such as gradient descent, stochastic gradient descent, and Newton or quasi-Newton methods) are crucial for parameter learning in high-dimensional data spaces and for minimizing loss functions (h2o.ai, 2024). These mathematical tools serve as the building blocks for the development of efficient learning algorithms.

In addition to logic, mathematics, and statistics, cognitive science and neuroscience have also played decisive roles in the theoretical evolution of AI. Research in cognitive psychology, which investigates perception, attention, memory, problem-solving, and language processing, has offered models that have inspired AI algorithms. This understanding of the nature of intelligence underpins both symbolic and connectionist approaches in system design.

Neuroscience stands at the forefront in establishing the conceptual foundations of artificial neural networks. The McCulloch-Pitts neuron model, a simplified abstraction of biological neurons employing parallel and distributed information processing, was among the earliest contributions to AI (He, Yang, He, & Zhao, 2021). This model laid the historical groundwork for the multi-layer neural networks we now call deep learning. Modern deep learning architectures—enhanced by advanced optimization methods, regularization techniques, and complex layer structures—have built upon that basic model to achieve near-human performance in tasks such as perception, recognition, classification, and prediction, particularly when large datasets are available.

In the realm of "Physical AI," synthetic data generation, learning, optimization, and modeling play significant roles in everything from design and prototyping to predicting the behavior of physical systems (Nvidia, 2024). At this point, Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs)—collectively known as "Generative Models"—assume a critical function (Belcic, 2024). By learning complex data distributions, these models can produce new, realistic, and creative data outputs. This capability provides a powerful infrastructure for both virtual simulations and predictions regarding physical systems.

The primary objective of Generative Models is to learn the probability distribution of a particular dataset and then generate new examples that share similar statistical properties. For instance, a model may generate realistic images, audio signals, or robot movement patterns. Such models facilitate the acquisition of high-dimensional and complex data for Physical AI.

Among the types of generative models are VAEs (Variational Autoencoders), which compress high-dimensional inputs into a lower-dimensional space, producing latent representations that preserve the meaningful information contained in the original data. Through dimension reduction, this process retains the significant content of the data. These latent space representations enable the statistical capture, manipulation, and variability of data space.

In the context of machine learning, "mathematical dimensions" do not necessarily correspond to conventional spatial dimensions of the physical world; rather, they reflect the features of data. For instance, an image of a handwritten digit in the MNIST dataset, sized at 28 × 28 pixels in grayscale, can be represented as a 784-dimensional vector, where each pixel is assigned a value between 0 (black) and 1 (white). If the same image were in color, each pixel would comprise three dimensions (red, green, and blue, or RGB), resulting in a total of 2,352 dimensions.

However, not all such dimensions carry meaningful information. Although the digit itself occupies only a small portion of the image, a substantial portion of the input space may be merely background noise. Consequently, compressing data to contain only salient information—in other words, creating a latent space—can substantially enhance the accuracy, efficiency, and efficacy of many machine learning tasks. Figure 1 below illustrates the architecture of an autoencoder neural network (Bergmann & Stryker, 2024).

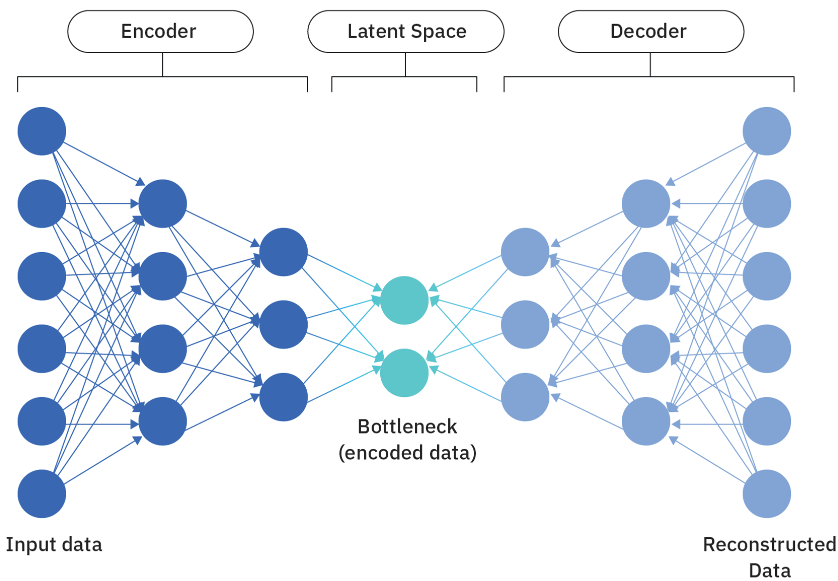


**Figure 1.** *Architecture of an autoencoder neural network*

The "VAE" model is an Autoencoder architecture composed of "Encoder" and "Decoder" components. The key distinction lies in the "Variational Inference" approach (Refaeli, 2023). Typically, the Encoder employs a Neural Network to produce a latent distribution characterized by mean and variance parameters derived from the input data, which is then mapped to a probability distribution (commonly a normal distribution). The Decoder, in turn, attempts to reconstruct the data space from latent vectors sampled from this latent space. A major advantage of VAEs is their ability to capture the underlying structural features of the data by modeling a continuous latent space. Since the "Generative" process is probabilistic, it allows for more diverse data generation. Moreover, it is possible to perform meaningful arithmetic operations in the latent space (e.g., interpolation between two face images).

On the other hand, the "GANs" model consists of two neural networks trained adversarially against each other, known as the "Generator" and the "Discriminator" (Google for Developers, 2022). This approach operates according to game theory principles (the minimax algorithm) (Suginoo, 2024). The Generator typically processes random noise vectors through a series of layers to produce meaningful data, while the Discriminator acts like a Binary Classifier, distinguishing between real data and the fake data generated by the Generator. During training, the Generator learns to produce increasingly realistic data that can fool the Discriminator. Training is performed via Gradient Descent–based methods: the Generator refines its output based on feedback from the Discriminator. By the end of training, the Generator is capable of producing high-quality, realistic synthetic samples resembling the true data distribution. GANs can thus generate highly realistic images, videos, sounds, and complex data spaces. They are particularly successful at producing high-resolution data, generally yielding sharp results because the generative process is optimized through adversarial competition. For instance, GANs can augment real robot sensor data to facilitate more extensive learning or expand limited datasets obtained from laboratory experiments, helping researchers explore the design space more efficiently.

In conclusion, "GANs" and "VAEs" represent two significant approaches within the family of "Generative Models." VAEs leverage probabilistic modeling to capture the latent structure underlying the data, whereas GANs use adversarial learning to produce highly realistic samples. For Physical AI, both approaches offer a broad spectrum of applications, from modeling physical systems and expanding design spaces to simulating complex interactions and performing data augmentation.

Below is a table (Table 1) comparing the fundamental differences and similarities between Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). Table 1 summarizes the general characteristics of both model families, their architectures, mathematical foundations, strengths and weaknesses, and typical areas of application.

**Table 1.** *Generative Adversarial Networks vs Variational Autoencoders*

| **Feature / Key Differences** | **GANs (Generative Adversarial Networks)** | **VAEs (Variational Autoencoders)** |
|---|---|---|
| Objective | To generate synthetic data that closely resembles real data | To learn a probabilistic representation of the data distribution and generate data from it |
| Architectural Structure | Two components: the Generator (producing synthetic data from noise) and the Discriminator (distinguishing real from fake) | A single encoder-decoder structure: the Encoder compresses the data into a latent space, and the Decoder reconstructs data from the latent vector |
| Learning Method | A minimax game: the Generator and Discriminator learn in mutual competition | Variational Bayesian approach: maximizing the Evidence Lower Bound (ELBO) |
| Mathematical Foundation | Game theory, based on the Jensen–Shannon divergence or a similar measure | Variational inference, minimizing the Kullback–Leibler (KL) divergence |
| Quality of Generated Data | Typically produces sharper, more realistic, and highly detailed images | Tends to produce smoother, sometimes blurrier outputs, but generally exhibits good diversity |
| Mode Collapse | Susceptible to mode collapse, potentially ignoring certain modes in the data | Less prone to mode collapse, better reflects diversity in the data distribution |
| Latent Space Representation | Does not provide an explicit latent space, the learned distribution is indirect | Offers a meaningful latent space, arithmetic can be performed on the latent vectors, and transitions are smooth |
| Training Difficulty | Training can be challenging, requiring precise hyperparameter tuning; unstable optimization is common | Generally more stable training, converges more easily when well-tuned |
| Applications | High-quality image generation (e.g., faces, synthetic photographs), style transfer, super-resolution | Understanding data distributions, data completion, anomaly detection, and variational modeling |
| Computational Cost | Often high, the Discriminator requires substantial data and computation to distinguish realistic fake samples | More controlled computational cost, ELBO optimization is direct, although complex density estimations may sometimes be required |
| Diversity of Outputs | Output diversity can be limited, especially in the case of mode collapse | Broad diversity, various examples can be easily generated from the latent space |

*Figure 2 presents a schematic representation of both VAEs and GANs (Singhal, 2023).*

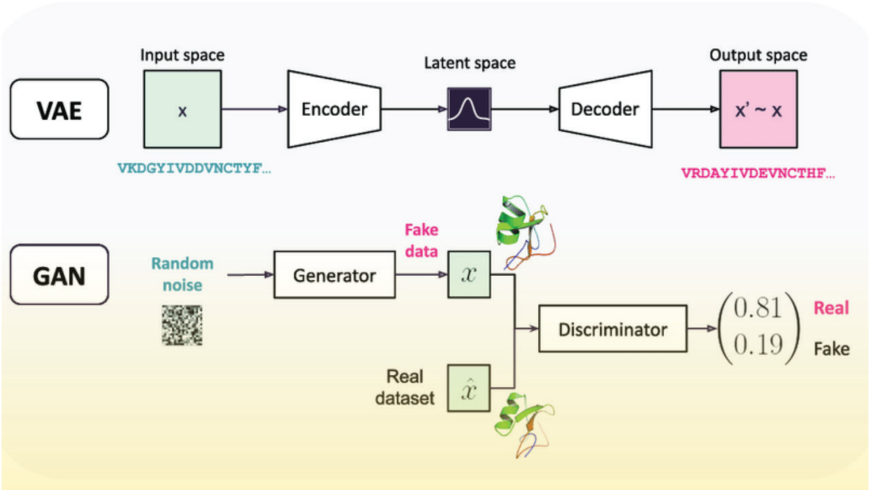


**Figure 2.** *VAEs and GANs*

In summary, even though both GANs and VAEs fall under the category of generative models, they employ different approaches in the data generation process. VAEs focus on modeling the underlying probability distribution of the data via an "Encoder–Decoder" system that learns a continuous latent space, while GANs use a "Generator–Discriminator" rivalry to produce highly realistic synthetic data. VAEs offer a more interpretable latent space representation, whereas GANs yield more realistic samples. This distinction enables both methods to serve complementary roles in multidimensional processes such as data augmentation, simulation, and design-space exploration within Physical AI applications.

## REFERENCES

Almabetter. (2024). Bayes Theorem in AI (Artificial Intelligence). Retrieved December 20, 2024, from https://www.almabetter.com/bytes/tutorials/artificial-intelligence/bayes-theorem-in-ai

Automation.com. (2024). LLMs Are Interesting, But Physical AI Is About to Reshape Our World. Retrieved December 20, 2024, from https://www.automation.com/en-us/articles/october-2024/physical-ai-reshape-our-world

Belcic, I. (2024). What is a Generative Model? | IBM. Retrieved December 21, 2024, from https://www.ibm.com/think/topics/generative-model

Bergmann, D., & Stryker, C. (2024). What is a Variational Autoencoder? | IBM. Retrieved December 21, 2024, from https://www.ibm.com/think/topics/variational-autoencoder

Costeira, J. P., & Lima, P. (2020). A simple guide to Physical AI | AI-on-Demand. Retrieved December 20, 2024, from https://www.ai4europe.eu/research/simple-guide-physical-ai

Google for Developers. (2022). The Discriminator | Machine Learning. Retrieved December 21, 2024, from https://developers.google.com/machine-learning/gan/discriminator

h2o.ai. (2024). What is Stochastic Gradient Descent? Retrieved December 20, 2024, from https://h2o.ai/wiki/stochastic-gradient-descent/

He, J., Yang, H., He, L., & Zhao, L. (2021). Neural networks based on vectorized neurons. *Neurocomputing*, *465*, 63–70. https://doi.org/10.1016/J.NEUCOM.2021.09.006

Mgmus. (2021). Alonzo Church: Unsung Hero of Modern Computing. Retrieved December 20, 2024, from https://www.mgmus.com/post/alonzo-church-unsung-hero-of-modern-computing

Nvidia. (2024). What is Physical AI? Retrieved December 21, 2024, from https://www.nvidia.com/en-us/glossary/generative-physical-ai/

Petracca, E. (2021). Simulating Marx: Herbert A. Simon's cognitivist approach to dialectical materialism. *History of the Human Sciences*, *35*(2), 101–125. https://doi.org/10.1177/09526951211031143

PgCert, A. R. (2024). AI-Driven Soft Robotics Design: Leveraging AI for Biologically Inspired Soft Robotics. Retrieved December 20, 2024, from https://www.linkedin.com/pulse/ai-driven-soft-robotics-design-leveraging-ai-inspired-ripla-pgcert-fykee

Refaeli, D. (2023). Variational Auto Encoders (VAEs). Variational Auto Encoders are an | Medium. Retrieved December 21, 2024, from https://medium.com/@david_39639/variational-auto-encoders-vaes-ccd1ffa87912

Singhal, R. (2023). GAN vs VAE: Two Creative Approaches to Generative AI - Linkedin. Retrieved December 21, 2024, from https://www.linkedin.com/pulse/gan-vs-vae-two-creative-approaches-generative-ai-rajat-singhal

Suginoo, M. (2024). Mini-Max Optimization Design of Generative Adversarial Networks (GAN) | Medium | Towards Data Science. Retrieved December 21, 2024, from https://towardsdatascience.com/mini-max-optimization-design-of-generative-adversarial-networks-gan-dc1b9ea44a02

Universal Robots. (2019). What Is Human-Robot Interaction - A Complete Guide. Retrieved December 20, 2024, from https://www.universal-robots.com/in/blog/human-robot-interaction-a-complete-guide/

Wikipedia. (2024). Aristotle. Retrieved December 20, 2024, from https://en.wikipedia.org/wiki/Aristotle

# Chapter 2

## ADVANCING PHYSICAL AI THROUGH MULTIDISCIPLINARY INTERACTIONS

Developing artificial intelligence systems in the physical environment requires a multidisciplinary approach. In a "Physical AI" system, the goal is for robots to exhibit capabilities not only in software but also in hardware. To achieve this goal, five fundamental disciplines—namely, Mechanical Engineering, Computer Science, Materials Science, Biology, and Chemistry—must work in an integrated manner, interacting with one another. The roles of these fields and their interactions in creating Physical AI vary. In Figure 3, one can see the foundational skills necessary to drive Physical AI innovation (Brogan, 2020).

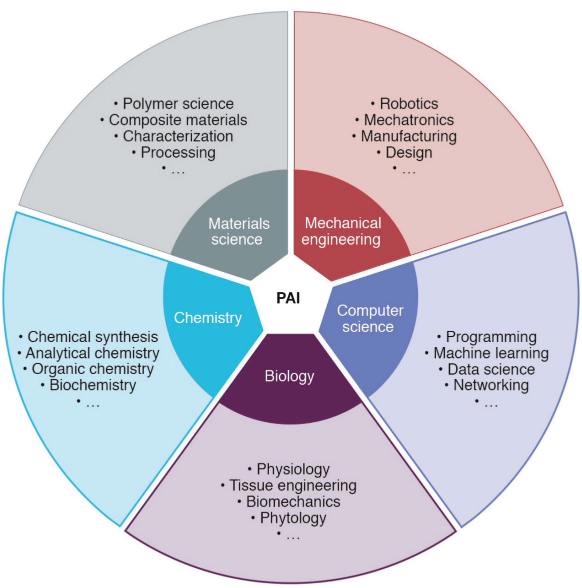


**Figure 3.** *The foundational skills necessary to drive Physical AI innovation.*

Mechanical Engineering is responsible for the kinetic and kinematic design of Physical AI systems. Robotics deals with the coordination of a robot's physical actions through components such as manipulator design, motion mechanisms, servo motors, and control systems (Reach Robotics, 2022). Articulated robotic arms or similar manipulators are driven by servo motors or stepper motors. The ability of these arms to perform precise positioning directly affects the robot's repeatability and accuracy performance.

Mechatronics focuses on the integrated design of mechanical, electronic, and control software (Babaiasl, 2022). It addresses sensors, actuators, and control circuitry of "Physical AI" robots as a whole. Mechatronic design enables Physical AI robots to make real-time decisions and take action.

Manufacturing encompasses methods for mass production or specialized production of complex robotic systems. By combining prototyping methods (e.g., 3D printing) with mass production methods (e.g., injection molding and sheet metal processing), it is possible to achieve both rapid prototyping and scalable manufacturing (3ERP, 2024).

Design involves the process of designing a robot's mechanics, taking into account both ergonomics and functionality. In designing the points of physical contact between robots and their users or environments, both safety and ergonomics are of critical importance (e.g., human-robot interaction, HRI).

In essence, Mechanical Engineering provides the fundamental infrastructure for physical movement and power transmission mechanisms in Physical AI. From a mechanical engineering perspective, "Physical AI" encompasses not only robotic arms and mobile platforms but also next-generation "soft robotics" applications, physical human-robot collaboration known as cobots (Yasar, 2024), and multi-axis motion systems.

Computer Science represents the "intelligence" and "information processing" aspects of robots. Programming includes the necessary programming languages and methods to develop robotic behaviors, control loops, and interaction protocols. A wide range of software is developed, from low-level embedded software to high-level autonomous decision-making algorithms, to manage a robot's functions.

Machine Learning develops algorithms that self-optimize by learning from data (e.g., sensor data, images, and sounds) (IBM, 2024b). Physical AI robots can adapt to their environment and improve their performance as they learn new tasks. Machine learning is employed in processes such as extracting meaningful outputs from sensor data, making predictions, and establishing adaptive control mechanisms. Applications span computer vision, natural language processing, and power optimization, among others.

Data Science provides the methods needed (e.g., data mining, statistical modeling) to store, process, and analyze large volumes of sensor and other data (IBM, 2024a). It addresses the need for real-time data processing, cloud-based analytics, and long-term storage for the significant volumes of data generated by robots or edge devices (Edge Computing).

Networking creates the infrastructure enabling Physical AI robots to communicate in real time with cloud computing resources or with one another. Technologies such as Wi-Fi, 5G, and low-latency network protocols exemplify the infrastructure supporting interactions between robots and cloud services or among robots themselves. This becomes especially critical in swarm robotics applications (Pham et al., 2021).

In summary, because it underpins "intelligent control" and "decision-making" mechanisms, computer science can be viewed as the "brain" of Physical AI projects.

Biology enables the development of "biomimicry" and "bio-inspired" designs by modeling the functioning of living systems. Efficient energy use, high adaptability, and durability in natural organisms can guide advancements in robotics and materials science. Examples include gecko-inspired climbing robots and octopus-like "soft" robotic arms (Sun, Bauman, Yu, & Zhao, 2023). In fact, a gecko-inspired robot can amputate its own limb to survive (Gunia, 2024).

Physiology is essential for understanding processes such as circulation, respiration, and nerve conduction in living organisms, thereby guiding robots in efficient energy usage and autonomous vital functions. Biological mechanisms like respiration, circulation, and nerve conduction shed light on goals such as self-regulation (akin to homeostasis) and energy efficiency in robot design.

Tissue Engineering is important for producing biological tissues in artificial environments and integrating them into robotic systems (e.g., biological sensors and "soft robotics"). The concept of using hybrid systems with biological tissues or cell-based sensors is rapidly gaining significance, particularly in medical and rehabilitation robotics (e.g., prostheses integrated with living cells).

Biomechanics examines the mechanical principles of the musculoskeletal systems of living organisms for translation into robot designs, enabling the creation of robots with more effective and efficient movements. Understanding human musculoskeletal structures and the locomotion systems of animals provides key insights into joint design, the balance between flexible/rigid structures, force distribution, and efficient movement strategies.

Phytology, especially in the realm of "plant-inspired" robotics, can help develop new robotic solutions by mimicking the growth, elongation, and environmental adaptation mechanisms of plants. Plant growth mechanisms, the transport of water and nutrients, and responses to light or gravity offer novel material and kinematic strategies for "plant-inspired" robotics.

In conclusion, biology facilitates the development of nature-inspired designs in Physical AI and even the integration of certain living components into robots. Biological approaches can be extended to the idea of using "biological modules" in Physical AI robots (e.g., bacteria-based sensors, algae-based energy production).

Chemistry addresses issues such as energy conversion, surface interactions, and the use of self-healing materials in Physical AI robots. Batteries and fuel cells, as well as systems that generate energy through chemical reactions (e.g., enzymatic fuel cells), can increase robotic autonomy.

Chemical Synthesis identifies the chemical processes needed to develop advanced polymers, composites, flexible materials, nanomaterials, and self-healing structures. In particular, the field of soft robotics focuses on the development of silicone, hydrogels, or other synthetic polymers.

Analytical Chemistry helps design and examine chemical sensing mechanisms that allow robots to detect their own status and environment. Sensor design for detecting gases, chemicals, or biological substances in a robot's surroundings is guided by this field. It plays a critical role in safety, environmental monitoring, and medical diagnostic robots.

Organic Chemistry can be employed in designing carbon-based materials. Carbon-based flexible electronics, wearable sensors, or thin-film transistors (TFTs) are pivotal applications in this context.

Biochemistry is essential for advanced applications, such as "bio-inspired" fuel cells or enzyme-based sensors, which are developed by taking inspiration from reactions inside living cells (e.g., enzymatic reactions). Robotic systems inspired by cellular and enzymatic processes are becoming increasingly prevalent.

In conclusion, chemistry offers an important layer that differentiates Physical AI from conventional robotic approaches by enabling reactive or adaptive interactions between robots and their environment.

Materials Science provides the scientific basis for the "body," "sensors," and "actuators" required in "Physical AI" systems. The choice of a robot's structural material is extremely important, as a combination of lightweight, flexible, and durable materials directly influences the robot's performance.

Polymer Science allows the design of flexible yet robust materials for use in flexible robotic systems, "soft actuators," and wearable technologies. For example, Shape Memory Polymers (SMP) can change shape with thermal or electrical stimulation.

Composite Materials involve specially engineered materials that optimize desired characteristics—such as lightweight properties, durability, and adaptability under various conditions—by combining different materials. This approach can achieve both high strength and lightness and is widely used.

Characterization analyzes the micro- and macro-scale properties of materials (e.g., mechanical, electrical, or thermal properties) to clarify design requirements. Techniques such as microscopy (SEM, TEM) and spectroscopy (FTIR, Raman) examine material microstructure, surface roughness, electrical conductivity, and other features, which is essential for selecting the right materials in design and production.

Processing covers methods of handling and shaping raw materials during manufacturing. Additionally, techniques in the manufacturing process—such as casting, extrusion, injection molding, sintering, or 3D printing—impact the quality and scalability of final robotic components. Hence, correct material selection and processing techniques enhance both the durability and functionality of Physical AI robots.

Through advances in nanotechnology and next-generation surface coating techniques (e.g., superhydrophobic or self-cleaning surfaces), materials science confers additional capabilities to robots under various environmental conditions.

In conclusion, the interdisciplinary interactions can be summarized as follows:

Þ Interaction between Materials and Mechanics: New-generation composite materials and processable polymers optimize the balance among weight, durability, and performance in robotic design.

Þ Interaction between Biology and Machine Learning: By analyzing biological data with machine learning methods, robots can form behavioral models that enable them to better understand and adapt to their environment.

Þ Interaction between Chemistry and Materials Science: Self-healing materials, electrochemically activated actuators (Yang et al., 2024), or thermally responsive polymers can be developed to incorporate innovative functionalities into robotic systems.

Þ Interaction between Mechatronics and Biomechanics: Actuation mechanisms resembling the human musculoskeletal system can enable robots to move more naturally and flexibly.

Þ Interaction between Networking and Robotics: Coordinated operation among numerous Physical AI robots (e.g., on a production line) is made possible by networking technologies, facilitating real-time data sharing across robotic systems.

The success of Physical AI relies on an integrated and interdisciplinary approach that combines fundamental fields such as Mechanical Engineering, Computer Science, Biology, Chemistry, and Materials Science. Equipping robots not only with "artificial intelligence" but also with intelligent materials, biologically inspired components, and advanced chemical processes where necessary will shape future AI solutions in physical environments. This holistic approach paves the way for robots to become more flexible, more durable, environmentally adaptive, and continuously capable of learning. Consequently, bringing together expertise from all these diverse domains is key to maximizing the potential of Physical AI.

All these concepts are indeed critical components of the Physical AI paradigm. Furthermore, the interdisciplinary nature of these interactions highlights why Physical AI systems necessitate such a broad scope of research and development activities. In addition, issues such as control theory, human-robot interaction (HRI), effective user interfaces, and sociotechnical dimensions (e.g., ethics, safety, and regulatory frameworks) are becoming increasingly important in Physical AI projects.

Another application that differs among Physical AI approaches is Swarm Intelligence, which draws inspiration from collective behavioral patterns observed in nature. It investigates how a large number of simple agents can solve complex problems through distributed, self-organizing interactions. Examples from nature include ants arranging their colonies, bees coordinating their hives, and fish swimming in schools to evade predators—these are the most commonly known natural cases in the field of swarm intelligence. This approach involves large groups of robots or software agents that interact with one another or the environment through simple signals.

One of the core principles of swarm intelligence is Distributed Control, where each robot or agent makes decisions independently based on simple rules, using locally gathered information. Through Local Interaction, the agents do not execute a global plan but only interact with nearby agents and the environment, thus facilitating scalability. Self-organization emerges from each unit adhering to simple rules, leading to emergent behavior at the system level. In subsequent stages, more complex tasks can be collectively accomplished. With Adaptation, the swarm can swiftly respond to environmental changes; other agents can take over the tasks of those that fail or become inoperative, thereby increasing robustness.

Depending on the nature of the task, using a single robot or employing a swarm intelligence approach may have advantages and disadvantages:

Complexity: A single robot approach requires incorporating complex tasks into one robot's architecture, which demands a powerful and often more expensive design, both in hardware and software. If the robot fails, the entire system loses functionality. In contrast, a swarm agent is typically simpler and cheaper to produce, although developing distributed control algorithms can be significantly more challenging. Still, a failure in one agent does not paralyze the entire system, providing high fault tolerance.

Scalability: With a single robot, performance improvements often involve adding more hardware or software modules, yet physical size and power constraints limit these enhancements. By contrast, swarm intelligence allows for scaling by adding more robots to the group, which can carry out tasks in parallel to boost overall performance. However, designing decentralized algorithms, software, and protocols for controlling large swarms requires advanced expertise.

Cost: A single robot can have high production costs—especially if it requires multiple capabilities, such as a diverse sensor array, a powerful processor, etc. Maintenance is simpler because there is just one system to maintain, though adding advanced features can rapidly raise maintenance costs. In swarm intelligence, each robot may be cheaper, but using hundreds or thousands of units can increase total expenditure. Thanks to the distributed architecture, defective robots can be replaced easily, and maintenance or backup strategies are more flexible. In the long term, especially with mass production or the use of simple hardware components, the total cost can be lower than that of one large, complex robot.

Reliability and Fault Tolerance: In a single-robot system, any hardware or software malfunction can halt the entire operation (i.e., a single point of failure). In swarm intelligence, even if one robot fails, others continue to fulfill the task. Depending on the level of task allocation and collaboration, the overall functionality can be maintained at a high level of certainty.

Software and Algorithm Development Challenges: A single-robot control system might initially seem simpler, but increasing the robot's versatility also increases algorithmic complexity. However, since testing and validation occur on a single platform, it is generally easier to detect errors. In swarm intelligence, multiple interactions among simple robots produce emergent behavior, requiring the development of collective intelligence algorithms. Additionally, designing localization, communication, and consensus mechanisms can be complex. Testing and simulation can become computationally heavy, making it difficult to generate accurate results.

Application Areas: A single robot can be used in large, complex operations that require a general-purpose design. Examples include medical robotics (surgical robots), humanoid robots, or industrial robots, often developed around a well-defined purpose and function, with established safety standards and control procedures. Swarm intelligence, however, is advantageous for parallel search operations in search and rescue missions (e.g., under rubble or in hazardous environments). In agricultural robotics, multiple robots can simultaneously monitor large areas for plant health or pest detection. In exploration, surveillance, and mapping, swarms can provide more efficient and cost-effective solutions.

Hence, swarm intelligence aims to transfer the distributed decision-making and self-organizing mechanisms found in large numbers of simple organisms into technological applications. Choosing to accomplish a task with a single robot versus a swarm of robots offers significant advantages in fault tolerance, scalability, and collective interaction. Yet, designing, controlling, and managing such systems can be more complex than developing a single-robot solution.

If high adaptability and parallel task execution are required, swarm intelligence becomes an appealing option. On the other hand, a single-robot solution, when specifically tailored to a given problem, can provide lower software complexity and easier management. However, the risk of a single point of failure and limitations in scalability must be carefully considered.

In conclusion, whether to use a single-robot or a swarm intelligence-based approach depends on the nature of the task, cost objectives, acceptable fault tolerance, and expected performance. Particularly as task scale and complexity grow, swarm intelligence is increasingly important within physical AI, robotics, AI, and multi-agent systems, taking on a more prominent role in practical applications.

## REFERENCES

3ERP. (2024). Robotics Prototyping & Part Manufacturing. Retrieved December 21, 2024, from https://www.3erp.com/industries/robotics-and-automations/

Babaiasl, M. (2022). What is Mechatronics Engineering? | Mecharithm. Retrieved December 21, 2024, from https://mecharithm.com/learning/lesson/what-is-mechatronics-22

Brogan, C. (2020). Skills development in Physical AI could cultivate lifelike intelligent robots | Imperial News. Retrieved December 21, 2024, from https://www.imperial.ac.uk/news/208053/skills-development-physical-ai-could-cultivate/

Gunia, A. (2024). This robot amputates its own limbs to survive | CNN. Retrieved December 21, 2024, from https://edition.cnn.com/science/yale-self-amputating-soft-robot-hnk-spc/index.html

IBM. (2024a). What is Data Science? Retrieved December 21, 2024, from https://www.ibm.com/think/topics/data-science

IBM. (2024b). What Is Machine Learning (ML)? Retrieved December 21, 2024, from https://www.ibm.com/think/topics/machine-learning

Pham, Q.-V., Nguyen, D. C., Mirjalili, S., Hoang, D. T., Nguyen, D. N., Pathirana, P. N., & Hwang, W.-J. (2021). Swarm intelligence for next-generation networks: Recent advances and applications. *Journal of Network and Computer Applications*, *191*, 103141. https://doi.org/10.1016/J.JNCA.2021.103141

Reach Robotics. (2022). What is a robotic manipulator? Retrieved December 21, 2024, from https://reachrobotics.com/blog/what-is-a-robotic-manipulator/

Sun, J., Bauman, L., Yu, L., & Zhao, B. (2023). Gecko-and-inchworm-inspired untethered soft robot for climbing on walls and ceilings. *Cell Reports Physical Science*, *4*(2), 101241. https://doi.org/10.1016/J.XCRP.2022.101241

Yang, L., Zhang, Y., Cai, W., Tan, J., Hansen, H., Wang, H., … Mu, J. (2024). Electrochemically-driven actuators: from materials to mechanisms and from performance to applications. *Chemical Society Reviews*, *53*(11), 5956–6010. https://doi.org/10.1039/D3CS00906H

Yasar, K. (2024). What is a Collaborative Robot (Cobot)? | Definition from TechTarget. Retrieved December 21, 2024, from https://www.techtarget.com/whatis/definition/collaborative-robot-cobot

# Chapter 3

## HUMANOID APPLICATIONS OF PHYSICAL AI

The term "humanoid" refers to advanced robotic systems in which Physical AI is implemented in a human-like body structure (i.e., a humanoid form) and employs Artificial Intelligence (AI) technologies in a physical environment (Huang, 2024). Essentially, these systems are created by combining mechanical designs and control engineering approaches that model the anatomy and functionality of the human body with high-level artificial intelligence techniques. Figure 4 below shows some humanoids produced in 2024, presented in sequential order (Thompson, 2024).

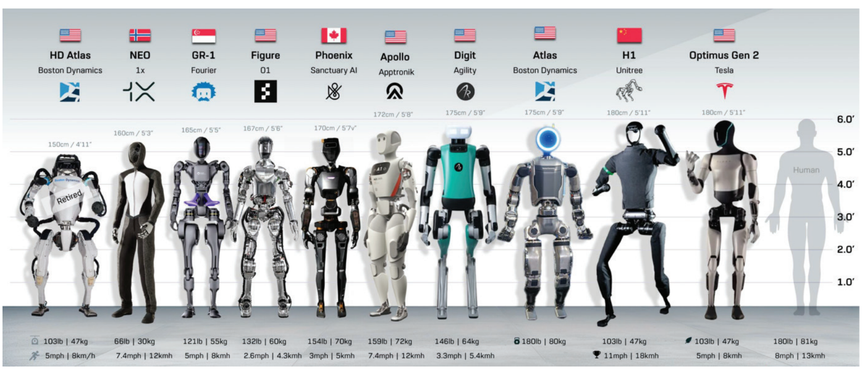


**Figure 4.** *Some Humanoid Robots of 2024*

A humanoid is designed to resemble a human, particularly in appearance and mobility. Depending on the situation, the head, torso, arms, and legs are human-like, and the degrees of freedom are generally close to those of human anatomy. This ensures a more natural human-machine interaction and improves the robot's ability to integrate into social environments.

For these robots, high levels of autonomy and learning capacity are provided through advanced methods such as Machine Learning, Deep

Neural Networks, and Reinforcement Learning. For instance, capabilities like face recognition, object detection, speech command interpretation, and natural dialogue enable humanoid robots to interact with humans more naturally.

Humanoid robots are equipped with a large number of sensors (tactile sensors, cameras, LiDAR, microphones, etc.). The data are processed using sensor fusion and real-time processing techniques, enabling the humanoid to quickly adapt to its dynamic environment. At this stage, Computer Vision and Natural Language Processing approaches play a critical role.

Human-like walking, running, jumping, and even climbing stairs can be achieved through precise kinematic and dynamic modeling of the robot. Methods such as Inverse Kinematics, Trajectory Planning, and Optimal Control are crucial for achieving flawless movements.

Control in humanoid robots is generally handled at two main levels: low-level control (instantaneous balance and movement at the motor level) and high-level control (task planning, decision-making, and learning). This architecture provides the robot with multi-tasking capabilities and flexibility for executing complex tasks. Since humanoids bring together different engineering fields (e.g., mechanical, electronics, software, and control theory) in a holistic approach, their technical building blocks are also highly diverse.

As a technical component of the mechanical structure, the most commonly used actuators for joint movement in humanoid robots are Servo Motors and Brushless DC Motors (BLDC). When high torque, low noise, and precision requirements are paramount, BLDC motors are preferred. In special designs where linear motion is required, linear actuators may be used. Series Elastic Actuators (SEA) incorporate a compliant element (such as a spring) between the motor and the load, providing more precise torque measurement and force feedback (Lee & Oh, 2019). Their usage is increasingly common in environments demanding safe human-robot interaction and smooth transitions. Meanwhile, Variable Stiffness Actuators (VSA) represent a new generation of actuators that can adjust the joint's stiffness and damping parameters in real time to achieve more natural, human-like movement and compliance (Sun, Xiong, Chen, Chen, & Yang, 2024). Passivity-Based Control and Compliance Control approaches aim to keep the robot stable and interaction-friendly by regulating joint stiffness settings and force/position feedback. For example, in manufacturing environments or medical applications, it is critically important that a robot safely interacts with humans. Adaptive Control and Robust Control methods automatically adjust the robot's behavior in unknown or changing environments, minimizing performance loss.

Regarding transmission, gearboxes can increase torque and reduce speed, or vice versa. For precise control, mechanisms such as planetary gearboxes or harmonic drives with low backlash and high gear ratios are preferred. In some joints or limbs, belt or chain-based transmission systems may be used to reduce weight or provide smoother transitions, especially if spacing in the construction requires it. In certain precise joints, direct drive is used; here, the motor shaft is directly connected to the joint, minimizing mechanical backlash in the actuator.

From the perspective of robot joints, revolute joints allow rotational movement similar to human joints. Prismatic joints, which provide linear motion and are commonly seen in industrial robots, might be used less frequently in humanoid designs, depending on the needs. Multi-axis joints enable multiple rotation axes at a single joint, facilitating more complex movements (Beard, 2024).

In terms of materials, aluminum alloys and titanium offer advantages due to their strength and light weight. Carbon fiber and other composite materials reduce overall robot mass, improving energy efficiency and mobility. Certain plastics and polymers may be used for low-cost, flexible parts or protective casings. From a design standpoint, employing cage structures and methods like topology optimization can yield lightweight designs (nTopology, 2024). Figure 5 below shows some optimized parts (Bernardino, 2022).

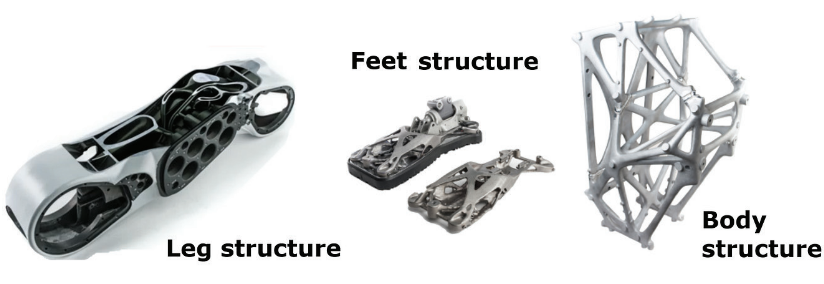


**Figure 5.** *Optimized Parts*

Moreover, for thermal management, fans, heatsinks, or liquid cooling systems may be utilized to control the heat generation of actuators and electronic components. With careful cable harness design, sensor and power cables inside the robot can be managed in an organized, flexible, and safe manner.

Regarding the vision system, an RGB camera is necessary for capturing color images and performing basic tasks such as object detection and face recognition. Stereo/Depth cameras provide 3D depth information that enables the robot to perceive its environment more accurately. Computer Vision algorithms then process the data from these sensors to support the robot's decision-making and navigation processes.

From the sensor perspective, the Inertial Measurement Unit (IMU) contains accelerometers and gyroscopes, which measure orientation, acceleration, and angular velocity, playing a critical role in balance control (SBG Systems, 2023). Force/Torque sensors, located either at the joint level or on end-effectors, measure the forces applied by the robot and the reactions they provoke during interaction. These sensors form the basis of Compliant Control and Haptic Feedback applications. Tactile sensors, especially on the robot's hands and soles, measure contact pressure and friction, enhancing the robot's ability to grip objects sensitively and maintain balance (Tacterion, 2023). Microphones enable the robot to detect voice commands and environmental sounds, which is important for Speech Recognition and Natural Language Understanding (NLU). Temperature and humidity sensors provide additional environmental data, while proximity sensors help prevent collisions and provide early warnings of obstacles in the robot's vicinity.

From a low-level control standpoint, the motor control loop employs algorithms such as Proportional-Integral-Derivative (PID) or Model Predictive Control (MPC) to regulate instantaneous position, speed, and torque at each actuator (Mortenson, 2024). Additionally, real-time processing ensures that critical functions, such as balance calculations in experimental biped robots—ZMP (Zero Moment Point) and Inverse Kinematics—can be performed within milliseconds.

At the high-level control layer, task planning and decision-making modules handle various robot tasks (e.g., walking, carrying objects, conversing). Behavior Tree or State Machine-based algorithms facilitate logical transitions between different states. Cognitive architectures like SOAR and ACT-R are employed for more "human-like" reasoning, problem-solving, and long-term learning (Kilicay-Ergin & Jablokow, 2012).

Regarding middleware, the Robot Operating System remains the most widely used platform, owing to its broad plugin support and modular structure (ros.org, 2024). It provides numerous packages for sensor fusion, navigation, simulation, and visualization. Alternative middleware frameworks such as YARP, LCM, and OROCOS may also be chosen based on specific needs.

When considering the integration of Artificial Intelligence (AI), methods such as Object Detection (YOLO, Faster R-CNN), Semantic Segmentation (U-Net, SegNet), and Pose Estimation (OpenPose) within the scope of Computer Vision assist the robot in interpreting visual information. In terms of Natural Language Processing (NLP), technologies such as Speech Recognition (Google Speech API, CMU Sphinx) and Language Understanding (BERT, GPT) enable spoken interaction with humans. Text-to-Speech (TTS) algorithms (e.g., WaveNet and Tacotron) allow the robot to speak in a natural manner (Oord & Dieleman, 2016). Within the scope of Reinforcement Learning (RL), Deep Reinforcement Learning methods such as PPO, DDPG, and SAC enable the robot to learn tasks related to walking, balance, or complex movements on a "trial-and-error" basis. Moreover, strategies exist for training robots initially in simulation environments and subsequently transitioning them to operate successfully on real hardware. Ontology-based approaches and Graph Database methods facilitate access to conceptual information (Graph.build, 2024), while high-level logical inference and planning (e.g., Hierarchical Task Network Planning – HTN) increase the system's flexibility (Geeksforgeeks, 2024).

From the perspective of battery systems, Lithium-ion (Li-ion) and Lithium Polymer (Li-Poly) batteries are frequently preferred due to their light weight and high energy density. Additionally, a Battery Management System (BMS) monitors the health and charge–discharge balance of the battery cells, extending battery life and enhancing safety. Some advanced designs may also support wireless charging pads and technologies such as inductive charging. In addition, research on energy harvesting (e.g., solar power, waste heat recovery) holds potential for extending the operating time of robots.

Regarding power optimization, Dynamic Power Scaling can be employed to dynamically adjust the actuator power or processor frequency according to the robot's tasks. Through Sleep/Wake Mechanisms, unused system components can be temporarily powered down, thus improving energy efficiency.

In terms of communication and networking, robots typically possess both wired and wireless connectivity. Utilizing high-speed serial communication in wired connections between robot components or sensors ensures that sensor data is transmitted with minimal latency. In wireless solutions, using Wi-Fi, 5G, or Bluetooth enables remote access, cloud computing, and data sharing. For processing control signals that require low latency, Real-Time Ethernet solutions (e.g., EtherCAT) or Time-Sensitive Networking (TSN) may be preferred (Botek Otomasyon, 2020).

Concerning safety and fault tolerance, redundancy is crucial; having multiple and backup versions of critical sensors and components allows rapid intervention in the event of a malfunction and ensures the continuation of the system's essential functions. Failsafe Mechanisms, designed to protect against issues such as overload, overheating, or electrical short circuits, include both hardware and software solutions that shut down the system to prevent damage. Soft Actuators and Torque Limiting methods reduce impact force during human–robot interaction, thereby lowering the risk of injury. Vision-Based Collision Avoidance, facilitated by cameras and proximity sensors, provides predictive obstacle detection and avoidance. Manual intervention by an operator or maintenance personnel in critical situations is also of great importance, particularly in enclosed or crowded areas, to ensure safety.

When examining the development of humanoid robots, a particular emphasis is placed on Whole-Body Control. The ability for a humanoid robot to control all of its joints—including arms, legs, and torso—in a synchronized manner allows for more efficient and balanced execution of complex movements (e.g., climbing up and down stairs). For Dynamic Balancing and stability control, methods based on the Zero Moment Point (ZMP) concept enable the robot to remain upright and walk steadily in real time. Another emerging area concerns sensors. Improved visual, auditory, and tactile data collection and integration enable robots to better perceive humans and their environment and respond more appropriately. In addition, there have been advances in cloud robotics. Offloading part of the computationally intensive AI algorithms to the cloud can reduce the onboard hardware requirements of the robot while taking advantage of continuously updated databases in real time. However, in situations where latency and bandwidth are concerns, powerful processors and GPU units on the robot itself can enhance real-time decision-making capabilities. Figure 6 below illustrates the changes in a humanoid over the years (Bernardino, 2022).

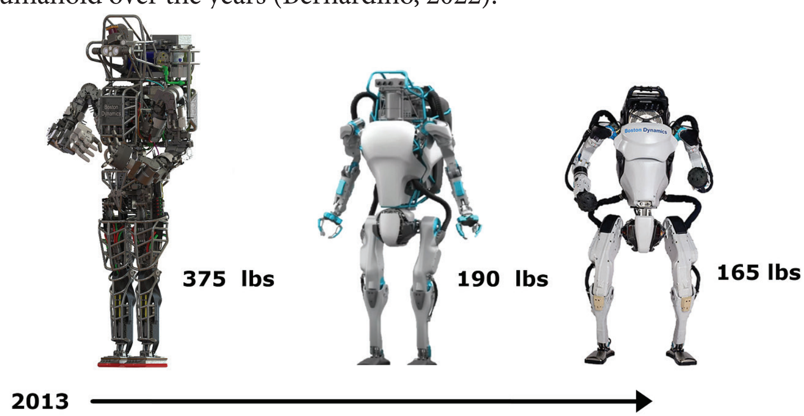


**Figure 6.** *Atlas' s design evolution (from 2013 to 2023)*

Furthermore, Figure 7 below depicts some of the technical components of a humanoid (Bernardino, 2022).

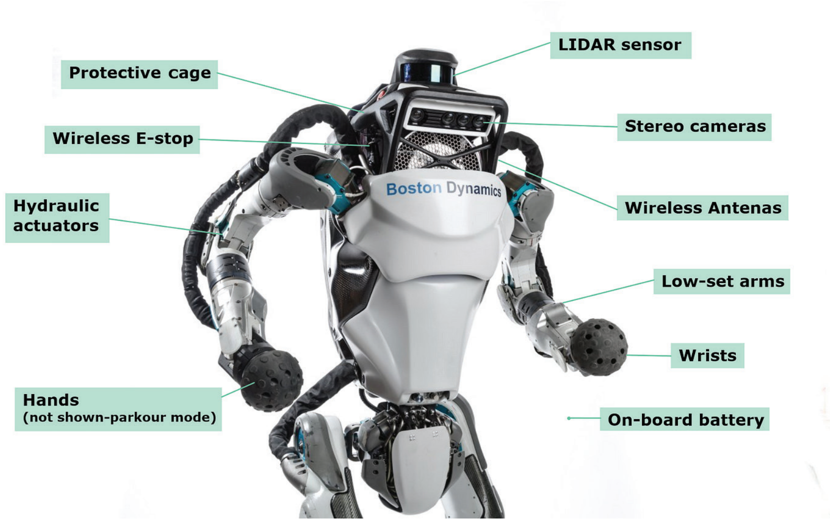


**Figure 7.** *Some Components of Atlas Humanoid Robot*

For instance, the Atlas Humanoid Robot, first introduced by Boston Dynamics in 2013 and continuously updated, is a humanoid robot whose primary goal is to provide an agile and dynamic platform that mimics human movement. Atlas is currently being developed to operate in search and rescue missions, hazardous tasks, and operations that require collaboration with humans.

The robot can run, jump, and even perform backflips in a remarkably competitive manner, a feat made possible by advanced motion control algorithms and sensor systems. For environmental sensing and navigation, Atlas employs lidar technology and stereo cameras. Lightweight structural components and the use of advanced design techniques such as lattice structures and topology optimization, in addition to durable materials, support the robot's agility and flexibility. Atlas also utilizes advanced hydraulic systems for precision motion and balance. Equipped with AI-assisted systems that continuously compute and adjust its movements and balance, Atlas can maintain its footing under challenging terrain and unstable conditions through real-time feedback mechanisms. Inspired by parkour, Atlas can perform complex maneuvers to overcome obstacles, showcasing running, jumping, and balancing on a single foot with remarkable stability and speed.

When interacting with humans, Atlas can easily imitate complex motions, thereby enhancing human–robot interaction. Observing Atlas's evolution suggests that humanoid robots could have even broader applications in the future: assisting humans in hazardous or hard-to-access areas during search and rescue missions, handling heavy loads in industrial settings such as factories or construction sites, or providing social support in homes and public spaces. It is anticipated that Atlas will be deployed in more extensive projects in the coming years.

In line with developments in 2024, Atlas's design has been updated. Additionally, a fully autonomous technical demonstration was presented to highlight potential industrial applications. Figure 8 below shows the current design of Atlas (Boston Dynamics, 2024a).

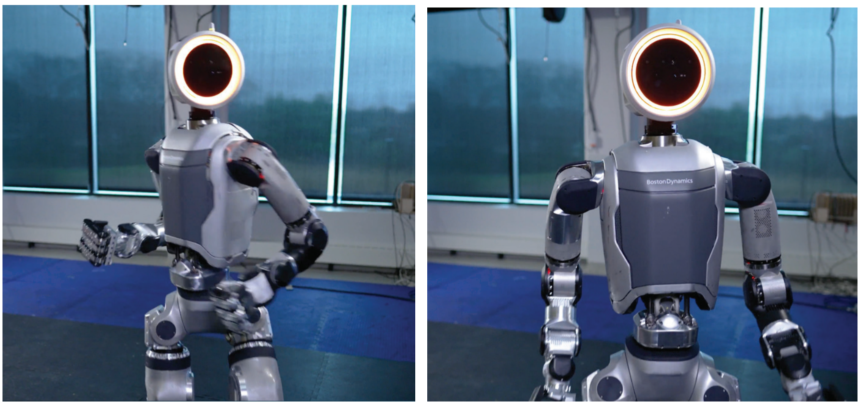

**Figure 8.** *New Atlas Humanoid Robot (2024)*

Moreover, Figure 9 below depicts a task performed entirely autonomously by Atlas in an industrial setting, wherein it carries motor covers (Boston Dynamics, 2024b).

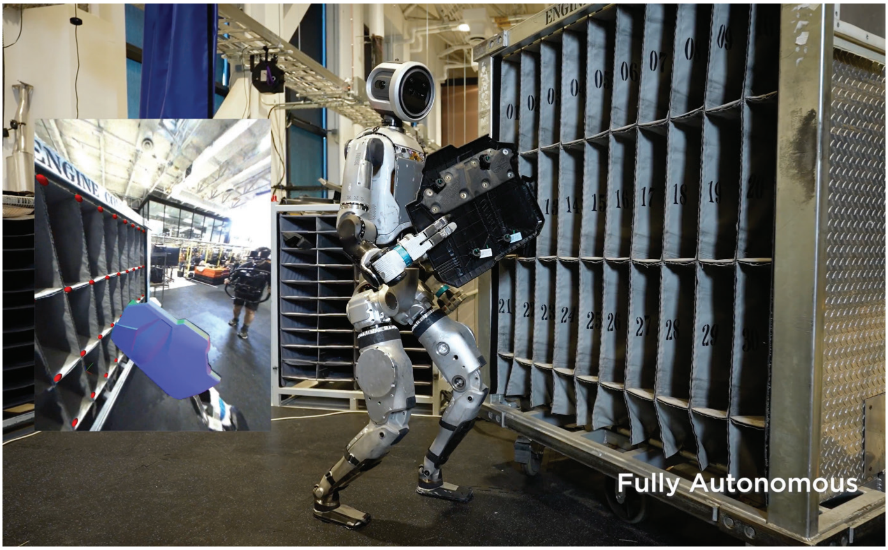


**Figure 9.** *Atlas Performing a Fully Autonomous Task*

In conclusion, the technical components of a humanoid robot must be addressed across a wide spectrum, ranging from the mechanical structure to the AI architecture, and from energy management to communication protocols. Human-like mobility and the potential for natural interaction can only be realized if these components operate in high alignment with one another.

Looking ahead, rapid developments in robotics, AI, and human–computer interaction (HCI) indicate that far smarter, more precise, safer, and more cost-effective versions of these technologies are likely. Robots with autonomy and learning capacities in the physical world are poised to have transformative impacts in numerous sectors, including healthcare, education, manufacturing, and even space exploration. However, in addition to technical progress, ongoing oversight, regulation, and scrutiny are required with respect to ethical and societal dimensions (e.g., privacy, security, workforce transformation). These considerations not only underscore the foundational structure of humanoid technology but also illuminate future research and potential application areas. When each of these technical components—each of which requires in-depth research and expertise—is brought together in an integrated manner, humanoid robots become capable of effective operation under real-world conditions.

## REFERENCES


Beard, E. (2024). Robot joints: The 6 different types explained - Standard Bots. Retrieved December 22, 2024, from https://standardbots.com/blog/types-of-robot-joints?srsltid=AfmBOopClV8LRx0S_TFzkyISubCZhKSKdVitntqK32Rc6QdbDoAOWTwo

Bernardino, M. (2022). Atlas, and the secrets behind the most dynamic humanoid in the world! | Medium. Retrieved December 21, 2024, from https://medium.com/@martabernardino1/atlas-and-the-secrets-behind-the-most-dynamic-humanoid-in-the-world-fcd79fc91a25

Boston Dynamics. (2024a). An Electric New Era for Atlas. Retrieved December 22, 2024, from https://bostondynamics.com/blog/electric-new-era-for-atlas/

Boston Dynamics. (2024b). Atlas Goes Hands On. Retrieved December 22, 2024, from https://bostondynamics.com/video/atlas-goes-hands-on/

Botek Otomasyon. (2020). EtherCAT Nedir? Neden İhtiyaç Duyulmuştur? Retrieved December 22, 2024, from https://botekotomasyon.com/ethercat-nedir-neden-ihtiyac-duyulmustur/

Geeksforgeeks. (2024). Hierarchical Task Network (HTN) Planning in AI. Retrieved December 22, 2024, from https://www.geeksforgeeks.org/hierarchical-task-network-htn-planning-in-ai/

Graph.build. (2024). Ontology in Graph Models and Knowledge Graphs. Retrieved December 22, 2024, from https://graph.build/resources/ontology

Huang, M. (2024). Physical AI Accelerated by Three NVIDIA Computers for Robot Training, Simulation and Inference | NVIDIA Blog. Retrieved December 21, 2024, from https://blogs.nvidia.com/blog/three-computers-robotics/

Kilicay-Ergin, N. H., & Jablokow, K. W. (2012). Problem-solving variability in cognitive architectures. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, *42*(6), 1231–1242. https://doi.org/10.1109/TSMCC.2012.2201469

Lee, C., & Oh, S. (2019). Development, analysis, and control of series elastic actuator-driven robot leg. *Frontiers in Neurorobotics*, *13*, 450201. https://doi.org/10.3389/FNBOT.2019.00017/BIBTEX

Mortenson, T. (2024). PID vs. Other Control Methods: What's the Best Choice? - RealPars. Retrieved December 22, 2024, from https://www.realpars.com/blog/pid-vs-advanced-control-methods nTopology. (2024). Next-Gen Engineering Design Software. Retrieved December 22, 2024, from https://www.ntop.com/

Oord, A. van den, & Dieleman, S. (2016). WaveNet: A generative model for raw audio - Google DeepMind. Retrieved December 22, 2024, from https://deepmind.google/discover/blog/wavenet-a-generative-model-for-raw-audio/

ros.org. (2024). Robot Operating System: Home. Retrieved December 22, 2024, from https://ros.org/

SBG Systems. (2023). IMU - Inertial Measurement Unit. Retrieved December 22, 2024, from https://www.sbg-systems.com/inertial-measurement-unit-imu-sensor/

Sun, X., Xiong, X., Chen, W., Chen, W., & Yang, G. (2024). Design and control of a novel variable stiffness actuator based on antagonistic variable radius principle. *ISA Transactions*, *147*, 567–576. https://doi.org/10.1016/J.ISATRA.2024.02.011

Tacterion. (2023). Tactile Sensing: Definition, Examples & Sensors. Retrieved December 22, 2024, from https://www.tacterion.com/wiki/tactile-sensing

Thompson, A. D. (2024). Humanoid robots ready for LLMs – Dr Alan D. Thompson – LifeArchitect.ai. Retrieved December 21, 2024, from https://lifearchitect.ai/humanoids/

# Chapter 4

## VARIOUS APPLICATIONS, IMPACTS AND FUTURE TRENDS OF PHYSICAL AI

Physical AI may have numerous transformative effects on humanity in the future, viewed from multiple perspectives. The primary focus of this technology is enabling artificial intelligence (AI) systems with cognitive functions to directly interact with the physical world, make autonomous decisions, and possess the capacity to move. In this context, examples such as “humanoid robots” and “autonomous vehicles” are among the most striking future applications that could have a far-reaching transformative impact. It is important to assess the social, economic, and ethical dimensions of these applications, as well as the technical opportunities and challenges they may present.

Physical AI’s “sensing” and “actuating” capabilities not only allow the system to perceive the external world but also enable it to intervene in its surroundings. Consequently, “Perception,” “Decision-making,” “Motion Planning,” and “Control” mechanisms should be seamlessly integrated into people’s lives due to their significant potential to influence daily routines.

Since humanoid robots feature a physical design that resembles the human body, they can more easily adapt to the tools, environments, and tasks encountered in daily life. For example, by virtue of their human-like arms, humanoids can conveniently use devices designed for human ergonomics. Through human-like facial expressions and structural posture, they can add a more “natural” dimension to social interactions with humans. More importantly, areas such as care for the elderly or individuals with disabilities represent important potential application domains for humanoid robots. These robots can perform functions such as daily assistance (e.g., meal preparation and medication reminders), physical support (e.g., aiding individuals with restricted mobility), and “cognitive assistance” (e.g., memory enhancement and the monitoring of mental stimulation activities). By doing so, they can serve supportive roles both in institutional care facilities and in personal environments, alleviating the staffing burden on healthcare services and enhancing personalized care opportunities. For instance, Tesla Bot appeared at a Tesla event on October 10, 2024, where it was present among a large group of people for social purposes and interacted with them. Figure 10 below shows a moment from this event (Tesla, 2024b).


**Figure 10.** *The Tesla Bot amidst the Crowd*

Additionally, Figure 11 illustrates Tesla Bot serving people in the crowd during the same event (Tesla, 2024b).


**Figure 11.** *Tesla Bot in the Act of Serving Humans*

On the other hand, in the context of "factories of the future," humanoid robots and other Physical AI elements can offer more flexible and intelligent automation on production lines. Today, humanoid robots that can replace or work in tandem with robotic arms are capable of undertaking specialized or complex assembly tasks in place of humans. Within the framework of "Human-Robot Collaboration (HRC)," humans and robots can share common tasks to maximize efficiency.

For example, Figure AI and BMW have achieved significant progress in the capabilities of the Figure 02 humanoid robot used on production lines. This robot has increased its speed by 400% and improved its success rate by sevenfold. A video published by the company demonstrates that Figure 02 completes 1,000 placement operations per day. In particular, its ability to handle sheet metal placement tasks requiring millimeter-level precision is highlighted. In this task, the robot must place sheet metal parts into a narrow space without collisions and with precision—ten times more challenging than previous tabletop manipulation tasks. Figure 02 has been tested at BMW's Spartanburg plant and is planned to continue in use for years to come. This collaboration showcases the potential of humanoid robots in automotive manufacturing processes. Moreover, Figure AI has received an investment of 675 million dollars from major investors such as OpenAI, Nvidia, Microsoft, and Jeff Bezos, bringing the company's valuation to 2.6 billion dollars. The company plans to deploy more humanoid robots in real-world environments. These developments reveal the potential of humanoid robots to enhance productivity by working alongside humans on production lines and taking on repetitive tasks. Figure 12 below shows the "Figure 02" humanoid robot at the BMW facility (Sinha, 2024).

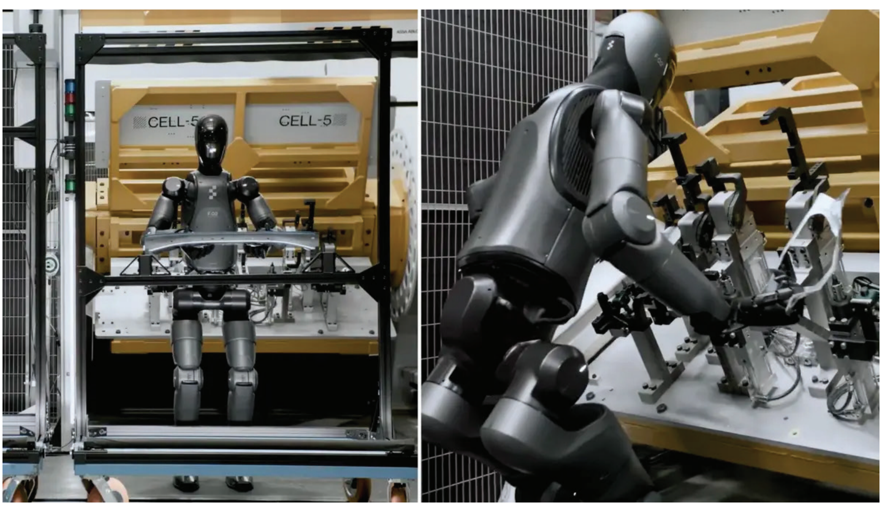


**Figure 12.** *Humanoid Robot (known as "Figure 02") in BMW' s Facility*

In the case of autonomous vehicles, "autonomous driving" technologies have the potential to radically transform people's transportation practices. Beginning with semi-autonomous systems such as "Advanced Driver

Assistance Systems (ADAS)," the process is evolving toward a fully driverless experience. Generally, this development may lead to fewer human-related accidents, reduced traffic congestion, and improved energy efficiency—effects that can be considered highly positive. Furthermore, in contrast to ride-hailing, the widespread adoption of ride-sharing could help alleviate urban transportation burdens (Ecolane, 2022). Figure 13 below shows Tesla's fully autonomous vehicle called Cybercab, which lacks a steering wheel and pedals (Edkins, 2024).


**Figure 13.** *New Tesla Cybercab*

Autonomous trucks and drones could revolutionize critical logistics processes such as "last-mile delivery" (Onfleet, 2023). By partially reducing the need for human drivers or delivery personnel, they offer the possibility of shorter delivery times and lower costs in the e-commerce industry. Similarly, in smart warehouses, operations could be carried out more rapidly and with fewer errors by employing autonomous forklifts and robotic couriers equipped with self-navigation systems. Figure 14 below shows Tesla's fully electric semi truck, which has the potential to be fully autonomous with a "Full Self Driving" option (Tesla, 2024a).

**Figure 14.** *Tesla Semi Truck*

From an urban planning and infrastructure standpoint, developments in smart traffic lights, "vehicle-to-infrastructure" (V2I), and "vehicle-to-vehicle" (V2V) communication can help optimize traffic flow (Dey, Rayamajhi, Chowdhury, Bhavsar, & Martin, 2016). In this regard, the design of next-generation highways and intersection systems, the integration of shared autonomous vehicle fleets, and the development and expansion of electric vehicle charging infrastructure will profoundly influence future urban planning strategies. Figure 15 illustrates an intelligent city structure where smart traffic lights, vehicles, and pedestrians communicate with each other, and traffic flow is autonomously managed and optimized (Aster Fab, 2021).

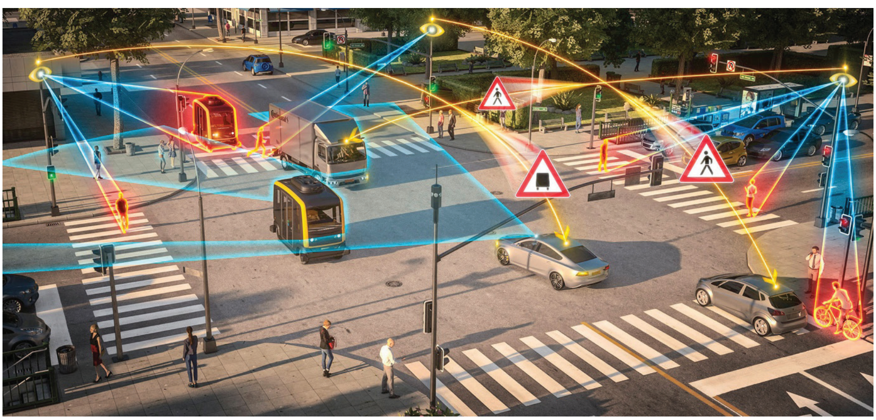

**Figure 15.** *Illustration of Autonomous Traffic Flow Management*

When considering societal and economic impacts, the widespread adoption of Physical AI-based systems may lead to profound transformations in the labor market. The takeover of certain routine or hazardous tasks by robots has the potential to enhance both safety and efficiency, while simultaneously allowing human labor to be directed toward more creative and strategic areas. On the other hand, during transitional periods in which skill sets do not align with emerging requirements, issues such as rising unemployment, income disparities, and the need for "reskilling" (i.e., acquiring new skills) may arise (Tamayo, Doumi, Goel, Ondrejkovic, & Sadun, 2023).

With respect to economic growth and innovation, improvements in productivity and efficiency, along with the emergence of new sectors and business models, may create additional momentum. The growth of the "service robots" market could stimulate entrepreneurs and investors to pursue new opportunities in this domain (Intel, 2024). Moreover, it is possible for the ecosystems of autonomous vehicle and robot manufacturers to expand, thereby generating employment. In the long term, AI-driven technologies—which form the backbone of Physical AI—will become one of the critical factors shaping global competition.

In terms of social acceptance and ethical responsibility, errors made by systems that interact with the physical environment can result not only in digital but also in tangible harm. Consequently, safety and accountability gain importance regarding societal acceptance. For instance, if an autonomous vehicle is involved in a traffic accident, ambiguity surrounding liability can be problematic. Likewise, if a humanoid robot providing caregiving services is involved in an unwanted incident, questions about how legal processes will be carried out become highly significant. These negative scenarios make it imperative to introduce new regulations and ethical guidelines.

Regarding technical challenges and areas for development, there is a pressing need for significant progress to ensure consistency in perception and decision-making processes. For Physical AI systems to perform flawlessly, extensive advancements are required in areas such as "sensor fusion," "real-time data processing," and "reinforcement learning (RL)." Autonomous vehicles that utilize multiple sensors (e.g., LIDAR, radar, and cameras) must reliably collect data under all conditions, including adverse weather, traffic, and road situations. Tesla, for example, has removed radar and ultrasonic sensors from its vehicles under the Tesla Vision system and aims to collect reliable data using only camera-based systems (Tesla Support, 2024).

In terms of durability, one example is that autonomous electric vehicles are built with more robust chassis designs (Singh, 2020). This approach addresses the potential damage a battery might incur under adverse conditions. In addition, humanoid robots—featuring a number of degrees of freedom comparable to that of humans—require complex control algorithms to maintain dynamic stability across multiple axes of motion, while also necessitating structural robustness. Developments in battery technology, materials science, and mechanical design—particularly the creation of lightweight yet sturdy structures—further contribute to making robots more practical and long-lasting.

From the perspective of communication and security, technologies such as “5G/6G” are essential for the efficient operation of autonomous vehicles, due to the need for high-speed, low-latency connections. However, cybersecurity vulnerabilities pose significant risks for Physical AI systems. If a robot or vehicle is hacked, large-scale damage could ensue. Therefore, innovations in “secure communication protocols” and “cryptography” are of critical importance.

In the near future, it is plausible that urban environments will encompass a broad ecosystem, including driverless public transportation, delivery robots, humanoid assistants, and industrial manufacturing robots. Fields that address social needs such as entertainment, companionship, and education are expected to expand rapidly. To meet the growing demands for regulation and legal oversight, countries will likely accelerate efforts to establish “standardization” and licensing processes that ensure Physical AI technologies are used safely and within well-defined boundaries of responsibility.

Physical AI represents one of the most tangible forms of artificial intelligence, with the potential to affect every sector of society. Humanoid robots, offering human-like interaction and service capabilities, could spearhead revolutionary transformations in fields such as healthcare, caregiving, and manufacturing. Autonomous vehicles, meanwhile, may reshape transportation and logistics systems, prompting sweeping changes in urban planning and economic models. Nevertheless, it is essential to address ethical, legal, security, and social acceptance concerns throughout the development and implementation of these technologies.

As these new generations of autonomous machines become increasingly integrated into daily life, it is vital to anticipate potential negative consequences and develop solutions accordingly. The labor market will likely undergo rapid restructuring. Especially for positions categorized as repetitive tasks, the displacement of such jobs by machines may further widen the gap between skilled and unskilled labor. This could deepen social inequalities and create substantial employment challenges for certain occupational groups. Societies have faced comparable challenges during past transitions, particularly under Industry 4.0, when automation technologies became widespread. However, because Physical AI partially automates decision-making processes, the scale and velocity of this transformation may be significantly larger.

At this juncture, controlling and monitoring the operation of systems capable of autonomous decision-making raises important technical and ethical questions. For example, "On what value judgments are machine learning algorithms trained?" or "To what extent should human intervention be permitted?" may become increasingly common inquiries in the future. Additionally, the evolution of these technologies may heighten the risks related to data privacy. The proliferation of technologies such as facial recognition could threaten individual privacy and dangerously increase the capacity for digital surveillance.

Closely related to these developments is the issue of algorithmic bias, which can reproduce existing prejudices and inequalities through technology (Jonker & Rogers, 2024). Physical AI systems trained primarily on data from specific demographic groups may exhibit higher error rates or unfair treatment toward underrepresented populations. For example, an autonomous patient-monitoring robot used in hospitals could pose direct health risks if it misclassifies or incorrectly diagnoses individuals from certain groups. Such scenarios are not limited to healthcare; similar inequities may arise in sectors such as education, security, and business.

Another significant negative impact lies in the realm of information pollution and the potential for public opinion to be manipulated. Physical AI tools may be exploited by malicious actors to generate and disseminate manipulative content. The rapid proliferation of data—whether genuine or fabricated—could escalate social tensions and exacerbate political polarization. Under such circumstances, public discourse and trust among citizens might deteriorate, adversely affecting democratic processes and social harmony.

Furthermore, the expansion of security risks must be emphasized. Even though various security protocols are being developed, the broad adoption of autonomous systems inherently enlarges the pool of potential targets. Advanced hacking methods could be used to disable critical infrastructure, including transportation, energy, and healthcare systems, leading not only to financial losses but also to threats against human lives. In military applications, the development and deployment of lethal autonomous weapons could transform the very concept of warfare and result in ethical and humanitarian crises.

Additionally, the production and continuous operation of Physical AI-based systems necessitate substantial energy and raw material consumption. This may exacerbate existing problems related to global warming and ecological degradation. Moreover, although it is foreseen that robots will be produced in significantly higher numbers in the future, strategies for disposing of or recycling these machines once they reach the end of their lifecycle remain insufficiently developed.

Finally, society's growing dependence on technology may present a grave risk. Delegating human responsibilities and competencies to machines could, over time, weaken both individual and collective decision-making capacities. As technological systems dominate all decision-making processes, humans might lose creativity and flexibility in their thinking. Furthermore, if people become complacent or incapable of responding to technological failures, large-scale societal problems could ensue. In short, the potentially adverse impacts of Physical AI range from shifts in the labor force and socio-economic structures to ethical and privacy dilemmas, environmental damage, security breaches, and the erosion of human autonomy. Hence, it is imperative for technology developers, policymakers, and all segments of society to take early-stage preventive and regulatory measures to address these risks.

In summary, while Physical AI—signifying the seamless integration of artificial intelligence into the physical world—presents significant opportunities for a sustainable and human-centric future, robust oversight and multidisciplinary approaches to governance and control are indispensable. Technological progress is likely to yield positive long-term outcomes only if it unfolds in harmony with societal values.

## REFERENCES


Aster Fab. (2021). V2X: Why should we accelerate adoption? Retrieved December 22, 2024, from https://aster-fab.com/v2x-why-should-we-accelerate-adoption/

Dey, K. C., Rayamajhi, A., Chowdhury, M., Bhavsar, P., & Martin, J. (2016). Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication in a heterogeneous wireless network – Performance evaluation. *Transportation Research Part C: Emerging Technologies*, *68*, 168–184. https://doi.org/10.1016/J.TRC.2016.03.008

Ecolane. (2022). Ride-Hailing vs. Ride-Sharing: The Key Difference. Retrieved December 22, 2024, from https://www.ecolane.com/blog/ride-hailing-vs.-ride-sharing-the-key-difference-and-why-it-matters

Edkins, J. (2024). New Tesla Cybercab makes its UK debut - a close-up with the new autonomous EV | Carwow. Retrieved December 22, 2024, from https://www.carwow.co.uk/tesla/news/8001/new-tesla-cybercab-robovan-price-specs-pictures

Intel. (2024). What Are Service Robots? Retrieved December 22, 2024, from https://www.intel.com/content/www/us/en/robotics/service-robot.html

Jonker, A., & Rogers, J. (2024). What Is Algorithmic Bias? | IBM. Retrieved December 22, 2024, from https://www.ibm.com/think/topics/algorithmic-bias

Onfleet. (2023). What is Last Mile Delivery? Everything You Need to Know. Retrieved December 22, 2024, from https://onfleet.com/blog/what-is-last-mile-delivery/

Singh, S. (2020). Simple tech: EV chassis design - Overdrive. Retrieved December 22, 2024, from https://www.overdrive.in/news-cars-auto/features/simple-tech-ev-chassis-design/

Sinha, S. (2024). BMW's Figure 02 humanoid robot gets 400% faster in manufacturing tasks | Interesting Engineering. Retrieved December 22, 2024, from https://interestingengineering.com/innovation/humanoid-robot-figure-02-400-speed

Tamayo, J., Doumi, L., Goel, S., Ondrejkovic, O. K., & Sadun, R. (2023). Reskilling in the Age of AI | Harvard Business Review. Retrieved December 22, 2024, from https://hbr.org/2023/09/reskilling-in-the-age-of-ai

Tesla. (2024a). Semi | Tesla. Retrieved December 22, 2024, from https://www.tesla.com/semi

Tesla. (2024b). Tesla on X: "The future will be streamed live 10/10, 7pm PT. Retrieved December 22, 2024, from https://x.com/Tesla/status/1843922599765590148

Tesla Support. (2024). Tesla Vision Update: Replacing Ultrasonic Sensors with Tesla Vision. Retrieved December 22, 2024, from https://www.tesla.com/support/transitioning-tesla-vision