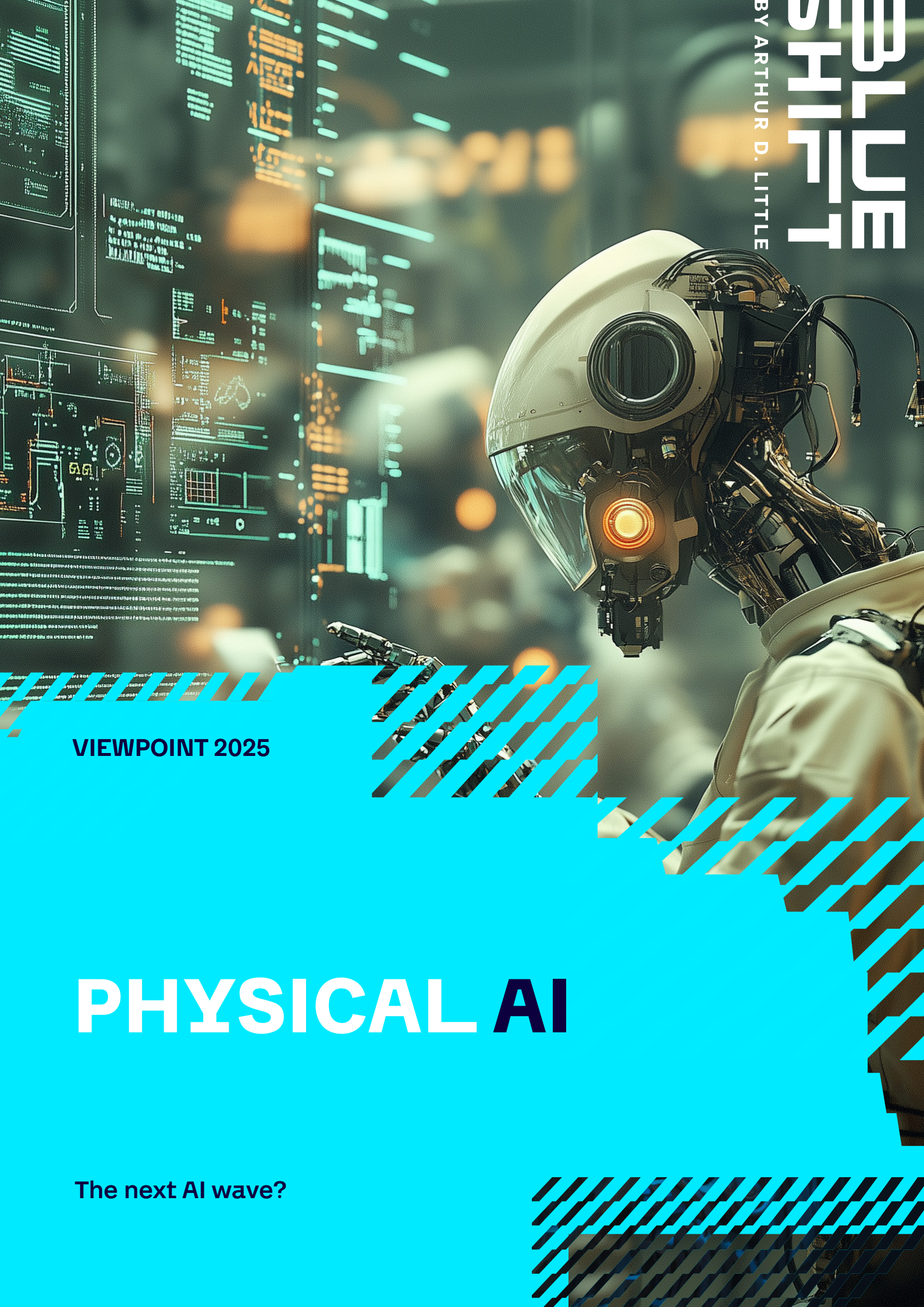VIEWPOINT 2025

PHYSICAL AI

The next AI wave?

PHYSICAL AI

The next AI wave?

AUTHORS

Zoe Huczok
Albert Meige
Rick Eagar

Physical AI is all about interactions between AI and the physical world. In combination with robotics technology, physical AI promises to revolutionize the capabilities of intelligent physical devices for applications from industrial process optimization to healthcare and personal robotic assistants.

However, there are still some difficult challenges to overcome. In this Viewpoint, we take a brief look at physical AI, examining its current state, potential use cases/challenges, and where it's heading. Is it really the next wave for AI?

WHAT IS PHYSICAL AI?

At the 2025 *Consumer Electronics Show*, Nvidia CEO Jensen Huang hailed physical AI as the next big thing for AI. Basically, physical AI is all about productive interactions between AI and the physical world. Classical machine learning (ML) and generative AI (GenAI) are primarily trained on data sourced from the publicly available Internet. Their outputs are provided in digital form (text, images, and sound) for human use. In contrast, physical AI directly captures data from the real world; for example, through sensors and Internet of Things (IoT) devices or through dedicated text, image, and sound training content. Its outputs act directly on the physical world; for example, by controlling actuators or other devices. Typical applications of physical AI are therefore in areas like robotics and automation, where AI gathers and processes real-world data, makes decisions, and acts on those decisions.

Physical AI differs from conventional robotics in one fundamental way. In conventional robotics, sensors feed data linearly into a processing engine that has been programmed to apply a set of rules. Based on these rules, the robot commands and controls its output devices and actuators accordingly. In physical AI, the processing engine communicates bidirectionally with a multimodal AI system rather than just applying rules. This means that a physical AI-enabled device can be much more responsive and adaptable to changing input data. It can apply reason, prioritize actions, change its processing path, and improve its performance continuously through learning without reprogramming. Because of this, physical AI is much better at dealing with complex, changing, and unpredictable situations than a conventional robot.

Physical AI differs from conventional robotics in one fundamental way.

Like all AIs, physical AI must be designed and trained to understand the physical environment. There are already some well-established approaches to achieving this. **Model-based reinforcement learning,** where the AI develops understanding through experimentation, has been used for some time, especially in robotics. **Simulation technologies,** such as factory or plant digital twins, can be used to develop a model of an environment. These are complemented by **physics-informed approaches** that involve constraining purely probabilistic AI models with known physical laws. Other methods include **graph neural networks,** which are particularly suited to the analysis of complex physical or biological systems (e.g., molecular structures or weather patterns), and symbolic AI, which relies on explicit rules and logic, in contrast to ML. Symbolic AI is useful for constrained environments where specific tasks must be executed and for imposing rule-based safety requirements.

WHY DO WE NEED IT?

Physical AI has come to the fore in recent years, partly due to continuing technological evolution and convergence. While GenAI excels at producing digital content based on vast online datasets, its lack of real-world sensory input limits its ability to interact effectively with dynamic environments. This gap is what physical AI aims to bridge. GenAI has become increasingly skilled at low-grade reasoning and information retrieval, resulting in greater integration with a range of systems and evolving into what's now referred to as "agentic AI." Agentic AI can execute entire workflows by breaking a problem into its component steps, fetching the appropriate data, analyzing it, deciding what next step to take based on an outcome, and reporting back to the user. These capabilities are fueling the adoption of AI in many enterprise applications (including R&D — see the Blue Shift report "Eureka! On Steroids").

At the same time, sensor and IoT technology has become more sophisticated and cheaper, and computing power has continued to evolve, leading to advances in simulation technology. Understanding and modeling of complex systems and their behaviors have also improved. Complexity and unpredictability are common characteristics of many real-world applications (e.g., industrial processes, business management, transport/travel systems, natural systems, and human networks). There is a natural convergence between AI and robotics for performing increasingly complex and unpredictable tasks. For AI to be applicable in all these situations, it must be able to reason and learn from experience, beyond the specific domains for which it was designed and trained.

Many important real-world applications require an understanding of multiple modes of data. For example, a doctor diagnosing a patient's condition will process a combination of text (medical records), speech (dialogue with the patient), and visual signals (visible symptoms).

Even for simpler tasks, such as responding effectively to a customer complaint, an agent will need to understand text (customer records), recognize speech, and be sensitive to other audible and/or visual signals about how the customer is feeling. Multimodality, the ability of an AI to recognize multiple data modes, is therefore key for AI to be able to function effectively in the physical world. While AIs are becoming increasingly multimodal, even the most advanced AIs and robots are still very limited in what they can achieve. This is true not only when compared to human capabilities, but also to animals — for instance, even a common house cat can plan and perform far more complex and adaptive actions than any robot.

Large language models (LLMs) such as ChatGPT, respond to questions based on their training datasets. While they can generate answers that seem knowledgeable, they do so based on statistical patterns in the data, not because they "understand" the concepts. They lack any real sensory information and knowledge about the world and how it works conceptually. As computer scientist Yann LeCun notes, it would be difficult for AI to function effectively as a physical-world agent without integrating sensory data alongside text, images, and sounds. He points out that a four-year-old child has already taken in more data through sensory inputs than even the largest LLM today. Conversely, it would take a human nearly half a million years to read all the textual data used to train a large LLM.

AI must be grounded in sensory data from the physical world if it is ever to move beyond narrow, task-specific functions and truly augment human capabilities in everyday, reasoning-intensive situations. This is why physical AI matters.

WHAT ARE THE USE CASES?

The use cases for physical AI are almost endless and cover virtually every sector. Some of the most likely applications relevant to the business world include:

- **Manufacturing and industrial processes.** Significant progress has already been made in developing digital twins for manufacturing plants, factories, and warehouses, enabling more responsive and data-driven optimization of operations. Physical AI will enable even faster, real-time optimization, without the need to reprogram or reconfigure due to changing operating conditions or unexpected events. Digital twins and simulations that have a stronger grounding in the physical world will be much more powerful — enabling higher performance and expanding their scope beyond individual plants to encompass entire supply chains and even the broader environment (see the Blue Shift report "The Industrial Metaverse"). This will ultimately enable the creation of "whole-system" digital models, allowing executives to harness AI for business decision-making and strategic management.

- **New robotics applications.** Physical AI will enable robotics to reach new levels of skill, multitasking, and adaptability. It will allow robot assistants to learn new tasks and adapt to new product lines or specifications without incurring downtime for reprogramming. It will also enable robots to work alongside humans in new, more uncontrolled environments, such as agriculture, construction, and mining.

> The use cases for physical AI are almost endless and cover virtually every sector.

- **Healthcare.** The applications of physical AI-enabled devices and robots in healthcare are vast, ranging from assisting with diagnoses to conducting complex surgical interventions, improving treatment precision, and providing a range of nursing, care, and support services.

- **Smart cities and homes.** Physical AI could enable real-time optimization of traffic flows, mobility provision, and improvement of public safety, with robots helping to provide essential urban services. It could also allow for adaptive control of the home environment, making a reality of the long-promised science fiction vision of domestic humanoid robots helping with day-to-day tasks.

Looking further into the future, some commentators have highlighted the potential of physical AI-enabled microbots and nanotechnology. In healthcare, microbots could travel through the bloodstream to targeted sites, enabling more precise diagnoses and effective treatments. More broadly, we could see the manufacturing of physical AI-enabled materials capable of self-adaptation and repair in response to changing conditions and environments.

WHAT ARE THE CHALLENGES?

Among the challenges that physical AI must overcome to deliver on its potential, four stand out as particularly significant.

1. The need for a world model

The primary challenge for physical AI lies in the vast amount of data required to construct a comprehensive "world model" — an abstract representation of the physical world essential for effective functioning. Finding more economical and energy-efficient methods than those used in conventional ML is therefore a key priority. One approach involves developing specific physical foundation models, which are pretrained on a diverse range of physical interactions such as navigating environments and interacting with people. Nvidia's Cosmos is one example. Described as a "world model development platform," it uses transformer-based architecture in the form of an auto-encoder and a diffusion algorithm that is "physics-informed." Another approach is Meta's Joint-Embedding Predictive Architecture (JEPA) model, which learns abstract representations of data by predicting missing or distorted information. Unlike traditional generative models, JEPA focuses on capturing high-level structures without reconstructing all the unpredictable details of a potential world, making it more efficient and scalable and reducing the amount of training data needed. More radical approaches, such as the "liquid network" AI model (see "Case study: Liquid networks") may have the potential to do "more for less" by mimicking more closely how neurons work in nature.

Physical AI interacts directly with and performs actions in the physical environment.

2. The need for decentralized computing power

Unlike agentic AI systems, which generate outputs for humans to act upon, physical AI interacts directly with and performs actions in the physical environment. This means it needs to respond immediately without latency, requiring embedded, decentralized computational power running on the device itself. Therefore, further progress in offerings for miniature computing and connectivity is important. All the premier chip manufacturers — Nvidia, Intel, AMD, IBM, Cerebras — and the hyperscalers (Google and Amazon) have released hardware for robotic model training or inference. Simulation-based approaches will also require specific computing capabilities. Decentralized architectures imply high power consumption — a development at odds with the general industry trend toward cheaper and less energy-intensive technologies. Developing "smaller brains that do not make mistakes" is therefore one of the holy grails of physical AI.

3. Reducing development times & costs

One of the biggest barriers to physical AI-enabled robotic devices is development time and cost. Robotic system training takes longer than GenAI training. For example, Covariant released its first world foundation model, RFM-1, in 2024, a full seven years after its founding. Testing and validation in real-world scenarios also greatly extend development times. Unlike pure software AI applications, which can be shipped with imperfections and improved upon during initial operation, physical AI devices must be error-free before launch. For example, Waymo robotaxis operated in pilot mode for over two years in the streets of San Francisco, California, USA — the last stage of testing — before they were commercialized as a service in February 2024. In practice, this means that large firms with deep pockets tend to dominate development, and partnerships are especially prevalent to share the costs.

4. Assuring safety

AI that acts directly on the physical environment can cause harm to humans. In safety-critical applications — such as train control systems, medical devices, industrial control systems, and aerospace systems — conventional robotics and automation ensure safety through rigorous processes of checking, verification, and validation during design, development, and commissioning, all of which are closely defined in regulatory standards. These processes and methods cannot be simply applied to AI-enabled systems, which are characterized by a lack of transparency and the propensity for errors, bias, and unpredictability. Ethical considerations also arise whenever a physical AI system must make decisions involving the avoidance of different types of harm. Notably, true Level 5 autonomy in vehicles (where systems can operate under all conditions without human intervention) has yet to be achieved. Therefore, a key challenge is developing new safety assurance approaches suitable for physical AI systems. Important aspects include improving transparency/explainability of AI decisions/actions; using rule-based models for safety-critical functions; developing new testing, verification, and validation regimes; and establishing new ethical frameworks/guidelines.

Case study: Liquid networks — Using inspiration from nature to address physical AI challenges

Daniel Rus, director of the Massachusetts Institute of Technology (MIT) Computer Science and Artificial Intelligence Laboratory, leads research aimed at overcoming some of physical AI's challenges by closely mimicking the way brains work in nature. The lab created an AI model that mimics the brain of the *C. elegans* worm, which contains only 302 neurons but can conduct many complex tasks (by comparison, the human brain contains over 80 billion neurons). The model, referred to as a "liquid network," contains far fewer neurons, but each neuron conducts more sophisticated mathematical operations than a conventional artificial neuron. This might hold significant advantages for physical AI applications for several reasons: (1) the limited number of neurons makes the system easier to understand and its decisions more explainable; (2) its compact size results in lower power consumption, which is critical for decentralized robotic systems; (3) development times are greatly reduced due to the modified approach used for testing; and (4) liquid networks can continue to adapt after initial training based on what they perceive. Further development is still ongoing.

CONCLUSION

Physical AI is key for the future

Physical AI could help us address some of the most pressing challenges of productivity, aging, and the environment and is already reshaping robotics. To effectively augment human capabilities in everyday situations that require reasoning and judgment, AI must be integrated with the physical world.

Physical AI's potential use cases are almost endless: key areas include manufacturing/industrial process optimization; robotics applications with new levels of skill, multitasking, and adaptability; healthcare; and smart cities/homes.

However, some challenges must be addressed — including the vast quantity of data needed to develop physical world models, the long development timelines/high costs of robotic systems, the need to embed substantial computational power into decentralized devices, and the complexity of ensuring safety. It remains unclear whether the "sledgehammer" approach of conventional AI — with its massive data requirements and high energy consumption — will be capable of overcoming these challenges, or if radically different approaches will be the key to moving AI into its next big wave.

Blue Shift, by Arthur D. Little, explores the impact of technologies on business, society, and humans. The Blue Shift Viewpoint covers these topics in depth, inviting guest authors, academics, and artists to contribute to the conversation.

BLUE SHIFT

BY ARTHUR D. LITTLE

ARTHUR D LITTLE