

NVIDIA Launches Cosmos World Foundation Model Platform to Accelerate Physical AI Development

- *New State-of-the-Art Models, Video Tokenizers and an Accelerated Data Processing Pipeline, Optimized for NVIDIA Data Center GPUs, Are Purpose-Built for Developing Robots and Autonomous Vehicles*
- *First Wave of Open Models Available Now to Developer Community*
- *Global Physical AI Leaders 1X, Agile Robots, Agility, Figure AI, Foretellix, Uber, Waabi and XPENG Among First to Adopt*

CES—NVIDIA today announced [NVIDIA Cosmos™](#), a platform comprising state-of-the-art generative [world foundation models](#), advanced tokenizers, guardrails and an accelerated video processing pipeline built to advance the development of [physical AI](#) systems such as [autonomous vehicles \(AVs\)](#) and [robots](#).

Physical AI models are costly to develop, and require vast amounts of real-world data and testing. Cosmos world foundation models, or WFMs, offer developers an easy way to generate massive amounts of photoreal, physics-based [synthetic data](#) to train and evaluate their existing models. Developers can also build custom models by fine-tuning Cosmos WFMs.

[Cosmos models](#) will be available under an open model license to accelerate the work of the robotics and AV community. Developers can preview the first models on the [NVIDIA API catalog](#), or download the family of models and fine-tuning framework from the [NVIDIA NGC™ catalog](#) or [Hugging Face](#).

Leading robotics and automotive companies, including 1X, Agile Robots, Agility, Figure AI, Foretellix, Fourier, [Galbot](#), [Hillbot](#), [IntBot](#), [Neura Robotics](#), Skild AI, Virtual Incision, Waabi and XPENG, along with ridesharing giant Uber, are among the first to adopt Cosmos.

“The ChatGPT moment for robotics is coming. Like large language models, world foundation models are fundamental to advancing robot and AV development, yet not all developers have the expertise and resources to train their own,” said Jensen Huang, founder and CEO of NVIDIA. “We created Cosmos to democratize physical AI and put general robotics in reach of every developer.”

Open World Foundation Models to Accelerate the Next Wave of AI

NVIDIA Cosmos’ suite of open models means developers can [customize](#) the WFMs with datasets, such as video recordings of AV trips or robots navigating a warehouse, according to the needs of their target application.

Cosmos WFMs are purpose-built for physical AI research and development, and can generate physics-based videos from a combination of inputs, like text, image and video, as well as robot sensor or motion data. The models are built for physically based interactions, object permanence, and high-quality generation of simulated industrial environments — like warehouses or factories — and of driving environments, including various road conditions.

In his opening [keynote at CES](#), NVIDIA founder and CEO Jensen Huang showcased ways physical AI developers can use Cosmos models, including for:

- **Video search and understanding**, enabling developers to easily find specific training scenarios, like snowy road conditions or warehouse congestion, from video data.
- **Physics-based photoreal synthetic data generation**, using Cosmos models to generate photoreal videos from controlled 3D scenarios developed in the [NVIDIA Omniverse™](#) platform.
- **Physical AI model development and evaluation**, whether building a custom model on the foundation models, improving the models using Cosmos for reinforcement learning or testing how they perform given a specific simulated scenario.
- **Foresight and “multiverse” simulation**, using Cosmos and Omniverse to generate every possible future outcome an AI model could take to help it select the best and most accurate path.

Advanced World Model Development Tools

Building physical AI models requires petabytes of video data and tens of thousands of compute hours to process, curate and label that data. To help save enormous costs in data curation, training and model customization, Cosmos features:

- **An NVIDIA AI and CUDA®-accelerated data processing pipeline**, powered by [NVIDIA NeMo™ Curator](#), that enables developers to process, curate and label 20 million hours of videos in 14 days using the NVIDIA Blackwell platform, instead of over three years using a CPU-only pipeline.
- [NVIDIA Cosmos Tokenizer](#), a state-of-the-art visual tokenizer for converting images and videos into tokens. It delivers 8x more total compression and 12x faster processing than today’s leading tokenizers.
- The [NVIDIA NeMo](#) framework for highly efficient model training, customization and optimization.

World's Largest Physical AI Industries Adopt Cosmos

Pioneers across the physical AI industry are already adopting Cosmos technologies.

1X, an AI and humanoid robot company, launched the [1X World Model Challenge](#) dataset using Cosmos Tokenizer. XPENG will use Cosmos to accelerate the development of its humanoid robot. And Hillbot and Skild AI are using Cosmos to fast-track the development of their general-purpose robots.

"Data scarcity and variability are key challenges to successful learning in robot environments," said Pras Velagapudi, chief technology officer at Agility. "Cosmos' text-, image- and video-to-world capabilities allow us to generate and augment photorealistic scenarios for a variety of tasks that we can use to train models without needing as much expensive, real-world data capture."

Transportation leaders are also using Cosmos to build physical AI for AVs:

- Waabi, a company pioneering generative AI for the physical world starting with autonomous vehicles, is evaluating Cosmos in the context of data curation for AV software development and simulation.
- Wayve, which is developing AI foundation models for autonomous driving, is evaluating Cosmos as a tool to search for edge and corner case driving scenarios used for safety and validation.
- AV toolchain provider Foretellix will use Cosmos, alongside [NVIDIA Omniverse Sensor RTX APIs](#), to evaluate and generate high-fidelity testing scenarios and training data at scale.
- Global ridesharing giant Uber is partnering with NVIDIA to accelerate autonomous mobility. Rich driving datasets from Uber, combined with the features of the Cosmos platform and [NVIDIA DGX Cloud](#)[™], can help AV partners build stronger AI models even more efficiently.

"Generative AI will power the future of mobility, requiring both rich data and very powerful compute," said Dara Khosrowshahi, CEO of Uber. "By working with NVIDIA, we are confident that we can help supercharge the timeline for safe and scalable autonomous driving solutions for the industry."

Developing Open, Safe and Responsible AI

NVIDIA Cosmos [was developed](#) in line with NVIDIA's [trustworthy AI](#) principles, which prioritize privacy, safety, security, transparency and reducing unwanted bias.

Trustworthy AI is essential for fostering innovation within the developer community and maintaining user trust. NVIDIA is committed to safe and trustworthy AI, in line with the White House's voluntary AI commitments and other global AI safety initiatives.

The open Cosmos platform includes guardrails designed to mitigate harmful text and images, and features a tool to enhance text prompts for accuracy. Videos generated with Cosmos [autoregressive](#) and [diffusion](#) models on the NVIDIA API catalog include invisible watermarks to identify AI-generated content, helping reduce the chances of misinformation and misattribution.

NVIDIA encourages developers to adopt trustworthy AI practices and further enhance guardrail and watermarking solutions for their applications.

Availability

Cosmos WFMs are [now available](#) under NVIDIA's open model license on Hugging Face and the NVIDIA NGC catalog. Cosmos models will soon be available as fully optimized [NVIDIA NIM](#) microservices.

Developers can access [NVIDIA NeMo Curator](#) for accelerated video processing and customize their own world models with [NVIDIA NeMo](#). [NVIDIA DGX Cloud](#) offers a fast and easy way to deploy these models, with enterprise support available through the [NVIDIA AI Enterprise](#) software platform.

NVIDIA also announced new [NVIDIA Llama Nemotron large language models and NVIDIA Cosmos Nemotron vision language models](#) that developers can use for enterprise AI use cases in healthcare, financial services, manufacturing and more.

About NVIDIA

[NVIDIA](#) (NASDAQ: NVDA) is the world leader in accelerated computing.

Certain statements in this press release including, but not limited to, statements as to: the benefits, impact, performance and availability of NVIDIA's products, services, and technologies, including NVIDIA Cosmos, NVIDIA API catalog, NVIDIA Omniverse platform, NVIDIA AI, NVIDIA CUDA, NVIDIA NeMo Curator, NVIDIA Blackwell platform, NVIDIA Cosmos Tokenizer, NVIDIA NeMo framework, NVIDIA DGX Cloud, and NVIDIA AI Enterprise software platform; third parties adopting NVIDIA's products and technologies, and the benefit and impact thereof; and the ChatGPT moment for robotics coming are forward-looking statements that are subject to risks and uncertainties that could cause results to be materially different than expectations. Important factors that could cause actual results to differ materially include: global economic conditions; our reliance on third parties to manufacture, assemble, package and test our products; the impact of technological development

and competition; development of new products and technologies or enhancements to our existing product and technologies; market acceptance of our products or our partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of performance of our products or technologies when integrated into systems; as well as other factors detailed from time to time in the most recent reports NVIDIA files with the Securities and Exchange Commission, or SEC, including, but not limited to, its annual report on Form 10-K and quarterly reports on Form 10-Q. Copies of reports filed with the SEC are posted on the company's website and are available from NVIDIA without charge. These forward-looking statements are not guarantees of future performance and speak only as of the date hereof, and, except as required by law, NVIDIA disclaims any obligation to update these forward-looking statements to reflect future events or circumstances.

Many of the products and features described herein remain in various stages and will be offered on a when-and-if-available basis. The statements above are not intended to be, and should not be interpreted as a commitment, promise, or legal obligation, and the development, release, and timing of any features or functionalities described for our products is subject to change and remains at the sole discretion of NVIDIA. NVIDIA will have no liability for failure to deliver or delay in the delivery of any of the products, features or functions set forth herein.

© 2025 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, CUDA, DGX, NGC, NVIDIA Cosmos, NVIDIA NeMo, and NVIDIA Omniverse are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated. Features, pricing, availability and specifications are subject to change without notice.

Janette Ciborowski
+1-734-330-8817
jciborowski@nvidia.com