

Run AI on Your PC? GeForce Users Are Ahead of the Curve

Generative AI is no longer just for tech giants. With GeForce, it's already at your fingertips.

Author: Jesse Clayton

Gone are the days when AI was the domain of sprawling data centers or elite researchers.

For GeForce RTX users, AI is now running on your PC. It's personal, enhancing every keystroke, every frame and every moment.

Gamers are already enjoying the benefits of AI in over 300 RTX games . Meanwhile, content creators have access to over 100 RTX creative and design apps , with AI enhancing everything from video and photo editing to asset generation.

And for GeForce enthusiasts, it's just the beginning. RTX is the platform for today and the accelerator that will power the AI of tomorrow.

NVIDIA pioneered the integration of AI and gaming with DLSS , a technique that uses AI to generate pixels in video games automatically and which has increased frame rates by up to 4x.

And with the recent introduction of DLSS 3.5, NVIDIA has enhanced the visual quality in some of the world's top titles, setting a new standard for visually richer and more immersive gameplay.

But NVIDIA's AI integration doesn't stop there. Tools like RTX Remix empower game modders to remaster classic content using high-quality textures and materials generated by AI.

With NVIDIA ACE for Games , AI-powered avatars come to life on the PC, marking a new era of immersive gaming.

Creators use AI to imagine new concepts, automate tedious tasks and create stunning works of art. They rely on RTX because it accelerates top creator applications, including the world's most popular photo editing, video editing, broadcast and 3D apps.

With over 100 RTX apps now AI-enabled, creators can get more done and deliver incredible results.

The performance metrics are staggering.

RTX GPUs boost AI image generation speeds in tools like Stable Diffusion by 4.5x compared to competing processors. Meanwhile, in 3D rendering, Blender experiences a speed increase of 5.4x.

Video editing in DaVinci Resolve powered by AI doubles its speed, and Adobe Photoshop's photo editing tasks become 3x as swift.

NVIDIA RTX AI tech demonstrates a staggering 10x faster speeds in distinct workflows when juxtaposed against its competitors.

NVIDIA provides various AI tools, apps and software development kits designed specifically for creators. This includes exclusive offerings like NVIDIA Omniverse , OptiX Denoiser , NVIDIA Canvas , NVIDIA Broadcast and NVIDIA DLSS .

Beyond gaming and content creation, RTX GPUs bring AI to all types of users.

Add Microsoft to the equation and 100 million RTX-powered Windows 11 PCs and workstations are already AI-ready.

The complementary technologies behind the Windows platform and NVIDIA's dynamic AI hardware and software stack are the driving forces that power hundreds of Windows apps and games.

Gamers : RTX-accelerated AI has been adopted in more than 300 games, increasing frame rates and enhancing visual fidelity.

Creators : More than 100 AI-enabled creative applications benefit from RTX acceleration — including the top apps for image generation, video editing, photo editing and 3D. AI helps artists work faster, automate tedious tasks and expand the boundaries of creative expression.

Video Streamers : RTX Video Super Resolution uses AI to increase the resolution and improve the quality of streamed video, elevating the home video experience.

Office Workers and Students : Teleconferencing and remote learning get an RTX boost with NVIDIA Broadcast. AI improves video and audio quality and adds unique effects to make virtual interactions smoother and collaboration more efficient.

Developers : Thanks to NVIDIA's world-leading AI development platform and technology developed by Microsoft and NVIDIA called CUDA on Windows Subsystem for Linux , developers can now do early AI development and training from the comfort of Windows, and easily migrate to servers for large training runs.

Generative AI enables users to quickly generate new content based on a variety of inputs — text, images, sounds, animation, 3D models or other types of data — bringing easy-to-use AI to more PCs.

Large language models (LLMs) are at the heart of many of these use cases.

Perhaps the best known is ChatGPT, a chatbot that runs in the cloud and one of the fastest growing applications in history.

Many of these LLMs now run directly on PC, enabling new end-user applications like automatically drafting documents and emails, summarizing web content, extracting insights from spreadsheet data, planning travel, and powering general-purpose AI assistants.

LLMs are some of the most demanding PC workloads, requiring a powerful AI accelerator — like an RTX GPU.

What's fueling the PC AI revolution?

Three pillars: lightning-fast graphics processing from GPUs, AI capabilities integral to GeForce and the omnipresent cloud.

Gamers already know all about the parallel processing power of GPUs. But what role did the GPU play in enabling AI in the cloud?

NVIDIA GPUs have transformed cloud services. These advanced systems power everything from voice recognition to autonomous factory operations.

In 2016, NVIDIA hand-delivered to OpenAI the first NVIDIA DGX AI supercomputer — the engine behind the LLM breakthrough powering ChatGPT.

NVIDIA DGX supercomputers , packed with GPUs and used initially as an AI research instrument, are now running 24/7 at businesses worldwide to refine data and process AI. Half of all Fortune 100 companies have installed DGX AI supercomputers.

The cloud, in turn, provides more than just vast quantities of training data for advanced AI models running on these machines.

But why run AI on your desktop when the cloud seems limitless?

GPU-equipped desktops — where the AI revolution began — are still where the action is.

Availability : Whether a gamer or a researcher, everyone needs tools — from games to sophisticated AI models used by wildlife researchers in the field — that can function even when offline.

Speed : Some applications need instantaneous results. Cloud latency doesn't always cut it.

Data size : Uploading and downloading large datasets from the cloud can be inefficient and cumbersome.

Privacy : Whether you're a Fortune 500 company or just editing family photos and videos, we all have data we want to keep close to home.

RTX GPUs are based on the same architecture that fuels NVIDIA's cloud performance . They blend the benefits of running AI locally with access to tools and the performance only NVIDIA can deliver.

NPUs, often called inference accelerators, are now finding their way into modern CPUs, highlighting the growing understanding of AI's critical role in every application.

While NPUs are designed to offload light AI tasks, NVIDIA's GPUs stand unparalleled for demanding AI models with raw performance ranging from a 20x-100x increase.

AI isn't just a trend — it will impact many aspects of our daily lives.

AI functionality will expand as research advances and user expectations will evolve. Keeping up will require GPUs — and a rich software stack built on top of them — that are up to the challenge.

NVIDIA is at the forefront of this transformative era, offering end-to-end optimized development solutions.

NVIDIA provides developers with tools to add more AI features to PCs, enhancing value for users, all powered by RTX.

From gaming innovations with RTX Remix to the NVIDIA NeMo LLM language model for assisting coders, the AI landscape on the PC is rich and expanding.

Whether it's stunning new gaming content, AI avatars, incredible tools for creators or the next generation of digital assistants, the promise of AI-powered experiences will continuously redefine the standard of personal computing.

Learn more about GeForce's AI capabilities .

Original URL: <https://blogs.nvidia.com/blog/2023/09/21/ai-on-local-rtx-windows-pc/>