

Vereinfachte GenAI-Entwicklung

Dell Precision-KI-Workstations und
NVIDIA AI Workbench

[Jetzt erkunden](#)



Inhaltsverzeichnis

Herausforderungen bei der Entwicklung im Bereich generative KI (GenAI)	3
Die Lösung: NVIDIA AI Workbench auf Dell Precision-Workstations	4–7
Erfolgsgeschichte: Invoke und Invoke Enterprise	8
NVIDIA AI Workbench und NVIDIA AI Enterprise	9
Nächste Schritte und Ressourcen	10

Herausforderungen bei der GenAI-Entwicklung

Von der Einrichtung von GPUs und Jupyter Notebooks über die Verschiebung von Workloads bis zum Management von Workflows – die alltäglichen Herausforderungen bei der GenAI-Entwicklung können den Projektfortschritt beeinträchtigen. Der erforderliche Zeitaufwand und das benötigte technische Fachwissen können Data Scientists, TechnikerInnen und EntwicklerInnen entmutigen. Glücklicherweise können die richtigen Technologiepartner und -tools dazu beitragen, Hindernisse zu überwinden und die Entwicklung zu beschleunigen.

Gängige GenAI-Hürden



Komplexe
GPU-Einrichtung



Begrenzte Workload-
Portabilität



Zeitaufwendiges
Workflowmanagement



Die Lösung: NVIDIA AI Workbench auf KI-fähigen Dell Precision- Workstations

Mit NVIDIA AI Workbench auf Dell Precision-Workstations können erfahrene und neue EntwicklerInnen schnell und einfach GenAI-Projekte auf lokalen Workstations erstellen, testen und anpassen und diese dann per Scale-up oder -down auf eine beliebige Plattform übertragen. Für all das können Sie die KI-fähigen Dell Precision-Workstations verwenden, die die erforderliche hohe Performance für die Verarbeitung großer Datensätze und komplexer neuronaler Netze bieten.

NVIDIA AI Workbench

- ✓ Rationalisierte Einrichtung
Schnellere Programmierung und Anpassung von Modellen mit automatisierter GPU-Systemeinrichtung/-konfiguration
- ✓ Ortsunabhängige Ausführung von Projekten
Lokale Entwicklung und einfache Verschiebung von Workloads, um sie auf der in puncto Kosten, Geschwindigkeit und Skalierung besten Plattform auszuführen
- ✓ Schnellere KI-Workflows
Höhere Produktivität dank automatisiertem Container- und Versionsmanagement

Dell Precision-Workstations

- ✓ Lokale GenAI-Entwicklung
Lokale Anpassung und Entwicklung von GenAI-Modellen auf GPU-beschleunigten Workstations
- ✓ Skalierbar
Scale-up von einer auf 4 NVIDIA RTX™-GPUs der Ada-Generation
- ✓ KI-fähig
Entwickelt und konfiguriert zur Erfüllung Ihrer KI-Workload-Anforderungen

Rationalisierte Hardwareeinrichtung

Die Produktivität leidet, wenn EntwicklerInnen sich um den komplexen Prozess der Einrichtung eines GPU-Systems für Deep Learning und GenAI-Projekte kümmern müssen. NVIDIA AI Workbench beseitigt gängige Einrichtungsprobleme wie BS-/Treiberdiskrepanzen sowie CUDA-Treiber- und -Firmwareinkompatibilitäten. EntwicklerInnen können dank automatisierter Installation und Konfiguration von Software und Treibern schnell mit dem Programmieren beginnen. NVIDIA AI Workbench sorgt außerdem dafür, dass Container für eine optimale GPU-Effizienz basierend auf Ihrer Workload konfiguriert werden.



Höhere Produktivität dank automatisierter Einrichtung



Beseitigung gängiger Probleme bei der Hardwareeinrichtung



Sicherstellung, dass Container korrekt konfiguriert sind

Ortsunabhängige Ausführung von Projekten

Mit NVIDIA AI Workbench können EntwicklerInnen dank nahtloser Rechenlastverteilung flexibel entwickeln, anpassen und zusammenarbeiten. Verschieben Sie NVIDIA AI Workbench-Projekte auf das in puncto Kosten, Geschwindigkeit und sicherer Skalierung beste GPU-gestützte System. Realisieren Sie all das ohne den Zeitaufwand und die technischen Kompetenzen, die üblicherweise erforderlich sind. Umgebungs-, Daten-, Modell- sowie Standortänderungen und -versionen können mühelos bewältigt werden. Gleichzeitig sorgt NVIDIA AI Workbench für ein sicheres Management der Authentifizierung und geheimen Schlüssel, die NutzerInnen für den Zugriff auf verschiedene Ressourcen benötigen.



Einfache Workload-
Portierung



Sichere Verschiebung von Zugangs-
daten und geheimen Schlüsseln



Mühelose Bewältigung von
Veränderungen und Versionen

Schnellere KI-Workflows



Automatisierte Installation und
Verwaltung von Projekten



Automatische Nachverfolgung
von Projektabhängigkeiten



Vereinfachtes Projektmanage-
ment mit intuitiver
Benutzeroberfläche und CLI

Mit NVIDIA AI Workbench können neue und erfahrene EntwicklerInnen schnell Elemente von KI- und Deep-Learning-Projekten identifizieren, installieren und managen. Steigern Sie die Produktivität mit der automatisierten Installation mehrerer Projektressourcen. Verfolgen Sie mühelos Projektabhängigkeiten für bestimmte Modelle und GPUs, damit sich EntwicklerInnen auf die Ausführung konzentrieren können. NVIDIA AI Workbench vereinfacht das Workflowmanagement mit einer intuitiven Benutzeroberfläche und einer leistungsstarken Befehlszeilenschnittstelle.

Sehen Sie sich Workflowbeispiele auf GitHub an:

- ➔ [Hybrid-RAG-Projekt](#)
- ➔ [NIM Anywhere-Projekt](#)
- ➔ [Stable Diffusion-Projekt](#)
- ➔ [Feinabstimmungsprojekt](#)

Erfolgsgeschichte: Invoke Enterprise

Invoke hat ein NVIDIA AI Workbench-Projekt verwendet, um Unternehmen ein sicheres, auf Zusammenarbeit basierendes Trainieren von KI-Bildmodellen zu ermöglichen. Invoke hat sowohl Invoke Enterprise als auch NVIDIA AI Workbench in einer On-Premise-Umgebung bereitgestellt und die KI-Rechenlösung lokal auf [Dell Precision 7780 Mobile Workstations](#) mit NVIDIA RTX™ 5000-GPUs der Ada-Generation ausgeführt.

„Die Bereitstellung von GenAI-Modellen auf KI-fähigen Dell Precision-Workstations – insbesondere der Precision 7780 – hat neue Leistungsmaßstäbe gesetzt. Ältere Hardware kann Schwierigkeiten mit den Compute-Anforderungen von KI haben, aber die Ausführung von Invoke auf der 7780 ist nahtlos und leistungsstark. Die Workstation unterstützt alle unserer fortschrittlichsten KI-Workflows.“

– Kent Keirse, Gründer und CEO, Invoke

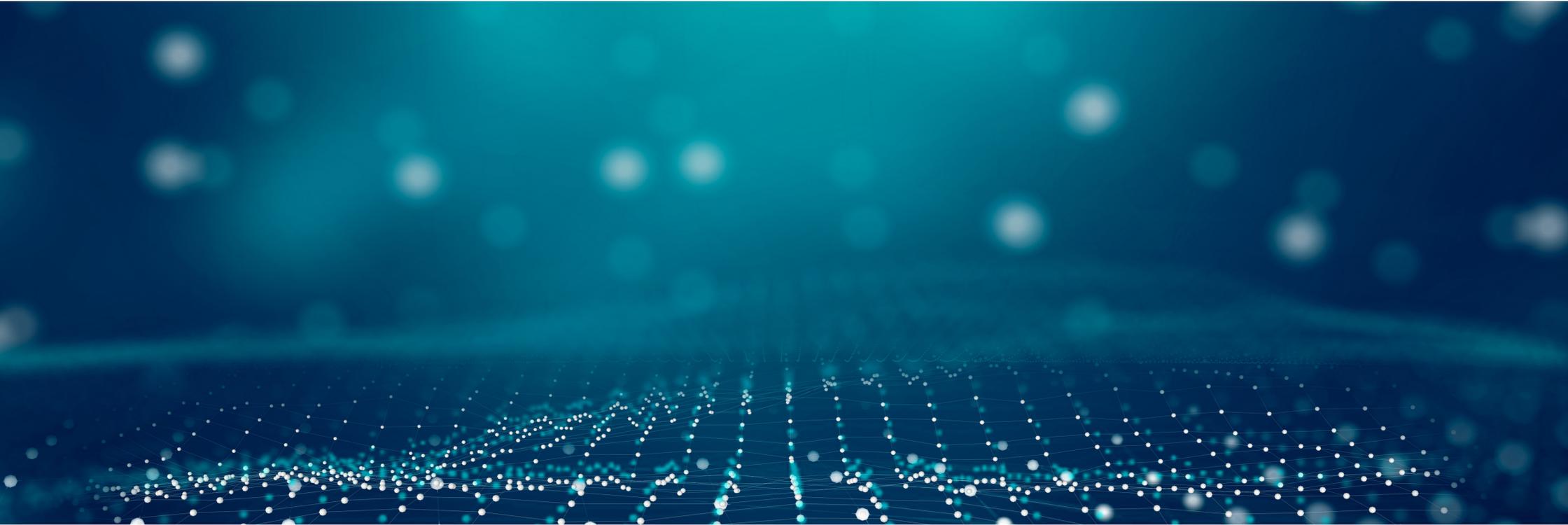
[Fallstudie lesen](#)





NVIDIA AI Workbench und NVIDIA AI Enterprise

Kombinieren Sie NVIDIA AI Workbench mit NVIDIA AI Enterprise, um Zugang zu einer cloudnativen End-to-End-Softwareplattform zu erhalten, die Data-Science-Pipelines beschleunigt sowie die Entwicklung und Bereitstellung von KI-Anwendungen der Produktionsklasse rationalisiert. NVIDIA AI Workbench ist kostenlos und umfasst Support der Enterprise-Klasse für NVIDIA AI Enterprise-LizenznehmerInnen, einschließlich Sicherheitsfunktionen, NIM und Produktionszweigen.



Nächste Schritte und Ressourcen

Jetzt erkunden

- ➔ AI Workbench kostenlos herunterladen
- ➔ KI-fähige Workstationkonfigurationen suchen
- ➔ Kontaktieren Sie uns, um NVIDIA AI Enterprise über Dell zu erwerben
- ➔ NVIDIA AI Enterprise 90 Tage lang kostenlos testen

Mehr erfahren

- ➔ Blog: „Harnessing NVIDIA Tools for Developers with Precision AI-Ready Workstations“
- ➔ Blog: „Decoding How NVIDIA AI Workbench Powers App Development“
- ➔ Blog: „Optimize AI Model Performance and Maintain Data Privacy with Hybrid RAG“
- ➔ Blog: „Develop and Deploy Scalable Generative AI Models Seamlessly with NVIDIA AI Workbench“