

# Das Potenzial von KI: Schnellere KI-Implementierungen und Innovationen mit HPE Private Cloud AI

## VERFASSER

**Steven Dickens**

Vice President und Practice Lead | The Futurum Group

## IN ZUSAMMENARBEIT MIT



**Hewlett Packard  
Enterprise**



**nVIDIA.**

**JUNI 2024**



## Übersicht

Das transformative Potenzial generativer KI (GenAI) hat in verschiedenen Branchen erhebliches Interesse und Wachstum ausgelöst. Allein im Jahr 2023 ist die Zahl der öffentlichen GenAI-Projekte auf GitHub um 248 % gestiegen, was die zunehmende Akzeptanz von KI-Technologien verdeutlicht. Aus diesem schnellen Wachstum ergibt sich jedoch auch eine Reihe von Herausforderungen. Unternehmen sehen sich mit komplexen Technologien konfrontiert und haben bei ihren KI-Projekten oft Probleme, Pilotprojekte in Produktionsumgebungen zu überführen. Für Unternehmen, die Bedenken haben, dass ihre geschützten Daten über öffentliche Modelle offengelegt werden könnten, ist der Schutz geistigen Eigentums ein wichtiges Thema. HPE geht in Zusammenarbeit mit NVIDIA auf diese Herausforderungen ein und bietet skalierbare, sichere und leistungsoptimierte Private Cloud-KI-Lösungen an. Mit diesen Lösungen können Unternehmen die Leistung von KI nutzen und gleichzeitig die Kontrolle über ihre Daten und Infrastruktur behalten, um schneller Zugang zu KI-gestützten Innovationen zu erhalten.

## Übersicht über HPE Private Cloud AI

HPE Private Cloud AI ist Bestandteil des NVIDIA AI Computing by HPE Portfolios und liefert eine sofort nutzbare, skalierbare und KI-optimierte Private Cloud zur Beschleunigung der Implementierung von KI-Projekten. Gleichzeitig wird sichergestellt, dass Unternehmen die Kontrolle über ihre Daten behalten. Die Lösung kombiniert Technologien für beschleunigtes Computing, Networking und Software von NVIDIA mit High Performance Computing und Storage von HPE und der HPE GreenLake Cloud, zusammen mit Datenpipeline-Funktionen, Orchestrierung und MLOps von HPE. Damit finden Unternehmen jeder Größe einen schnellen und flexiblen Zugang zur Entwicklung und Bereitstellung von GenAI-Anwendungen. Sie können diese Funktionen nutzen, um Hindernisse bei der Einführung von KI zu überwinden, ihre Time-to-Value zu beschleunigen und nahtlose Skalierbarkeit für zukünftiges Wachstum zu unterstützen.

## Absichtserklärung für das Researchdossier

Dieses Researchdossier enthält eine eingehende Analyse der Herausforderungen, mit denen Unternehmen bei der Implementierung von KI-Workloads konfrontiert sind, und legt dar, wie diese Herausforderungen mit HPE Private Cloud AI effektiv zu bewältigen sind. Im Dossier werden die wichtigsten Erkenntnisse und Einblicke in die Eigenschaften der Lösung in Bezug auf schnelle Nutzbarkeit, Wiederholbarkeit und Skalierbarkeit hervorgehoben. Darüber hinaus werden Empfehlungen für Unternehmen angeboten, die ihre KI-Implementierungen optimieren möchten, um den maximalen Nutzen aus KI-Technologien sicherzustellen aber gleichzeitig die volle Kostenkontrolle und eine umfassende Sicherheitsleistung wahren zu können.

# Die wichtigsten Erkenntnisse und Einblicke

1. **Schnellerer Nutzen:** Die Lösung von HPE und NVIDIA verkürzt die Markteinführungszeit bei KI-Projekten mit im Vorfeld integrierten und getesteten Tools, Modellen und Infrastrukturen erheblich. Dieses Konzept ermöglicht einen schnellen Übergang von der Pilot- zur eigentlichen Produktionsphase, sodass Unternehmen die Vorteile der KI schnell praktisch nutzen können.

2. **Wiederholbare Lösungen:** Die sofort nutzbare, standardisierte Infrastruktur von HPE Private Cloud AI unterstützt konsistente, leistungsfähige und skalierbare KI-Implementierungen. Durch die Bereitstellung definierter, integrierter und getesteter Tools können Unternehmen KI-Workloads schnell und einfach implementieren, oft mit nur einem einzigen Klick.

3. **Das Beste der Cloud:** Die Lösung von HPE ist so konzipiert, dass sie zusammen mit den Geschäftsanforderungen skaliert werden kann. Sie bietet flexible Bereitstellungsoptionen und Pay-as-you-go-Finanzierungsmodelle, die das Wachstum unterstützen und gleichzeitig die Kosten effektiv unter Kontrolle halten. Das CapEx-Modell und die Managementfunktionen der Lösung verbessern die Skalierbarkeit zusätzlich und tragen zu einer nahtlosen Erweiterung von KI-Funktionen im Unternehmen bei.

Nachfolgend finden Sie eine aktualisierte Übersicht über die Argumente dafür, dass Hybrid oder Private Clouds gegenüber Public Clouds für KI-Implementierungen besser geeignet sind, begleitet von den damit verbundenen Herausforderungen:

## Übersicht über die unternehmerischen Herausforderungen bei KI-Implementierungen

Wenn Unternehmen die Implementierung von KI in Erwägung ziehen, müssen Sie eine Reihe spezifischer Herausforderungen bewältigen, die Hybrid oder Private Cloud-Lösungen oft attraktiver machen als Public Cloud-Optionen:

**1. Definition, Entwicklung und Operationalisierung der richtigen Anwendungsfälle:** Unternehmen müssen sorgfältig die KI-Anwendungsfälle identifizieren, die ihnen den größten Nutzen bringen. Dafür werden fundierte Kenntnisse der geschäftlichen Anforderungen und die Fähigkeit gebraucht, diese in KI-Projekte umzusetzen. Sobald diese Anwendungsfälle identifiziert sind, ist die Entwicklung und Operationalisierung sehr komplex. Dazu gehört auch, die KI-Modelle in bestehende Workflows zu integrieren und sicherzustellen, dass sie greifbare Geschäftsergebnisse liefern.

**2. Erfolgsmaximierung mit Daten bei gleichzeitigem Schutz des geistigen Eigentums (IP):** Um aussagekräftige Einblicke mit KI-Modellen zu gewinnen, ist eine effektive Nutzung von Datenbeständen unerlässlich. Dabei darf aber kein sensibles geistiges Eigentum gefährdet werden. Public Cloud-Lösungen geben oft Anlass zu Bedenken hinsichtlich der Sicherheit und Kontrolle über Daten, weil es unter Umständen schwierig ist, rechtlich geschützte Algorithmen und sensible Daten vor unbefugtem Zugriff oder Datenschutzverletzungen zu schützen.

**3. Skalierung von KI-Projekten ohne Kostenexplosion:** Die Skalierung von KI-Initiativen passend zu wachsenden Anforderungen kann zu prohibitiven Kosten führen, insbesondere wenn die Lösungen auf Ressourcen der Public Cloud angewiesen sind. Unternehmen müssen den Bedarf an Rechenleistung und Speicherplatz mit den finanziellen Auswirkungen abwägen und sicherstellen, dass KI-Implementierungen bei der Skalierung kosteneffizient bleiben.

**4. Optimierung der Technologie für bestimmte Anwendungsfälle:** Herausforderung: Für unterschiedliche KI-Anwendungen werden jeweils andere Technologien und Infrastrukturen benötigt. Die Optimierung des Technologie-Stacks für bestimmte Anwendungsfälle ist extrem wichtig, aber auch komplex. Public Cloud-Lösungen bieten möglicherweise nicht immer die notwendige Individualisierbarkeit oder Leistungsoptimierung, die für bestimmte KI-Workloads gebraucht werden.

**5. Rückschlüsse und Abstimmungen nahe bei Datenquellen:** Analysefunktionen müssen in der Nähe der Daten ausgeführt werden, um Latenzen zu verkürzen und die Leistung zu steigern. Bei Public Cloud-Lösungen werden Daten häufig über große Entfernungen übertragen, was zu Verzögerungen und Ineffizienzen führt. Die Geschwindigkeit und Genauigkeit von KI-Modellen kann erheblich verbessert werden, wenn Rückschlüsse und Abstimmungen in der Nähe der Datenquellen ausgeführt werden.

Diese Herausforderungen verdeutlichen, wie wichtig es ist, Hybrid und Private Cloud-Lösungen für KI-Implementierungen in Betracht zu ziehen, da sie im Vergleich zu Public Cloud-Alternativen bessere Möglichkeiten in Bezug auf Steuerung, Individualisierung und Kostenkontrolle bieten können.



## Übersicht über die Ergebnisse

HPE Private Cloud AI bietet mehrere wichtige Merkmale und Funktionen, mit denen sich die wesentlichen Herausforderungen für Unternehmen bei der Bereitstellung von KI-Workloads bewältigen lassen:



**1. Schnellerer Nutzen:** HPE beschleunigt die Entwicklung und Bereitstellung von KI-Lösungen mit im Vorfeld integrierten und getesteten Infrastrukturen und Tools, die die Markteinführungszeit verkürzen.



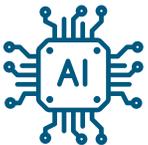
**2. Wiederholbarkeit:** Die Lösung bietet konsistente, leistungsfähige und skalierbare KI-Implementierungen anhand von standardisierten Infrastruktur-Stacks, die Zuverlässigkeit und Effizienz sicherstellen.



**3. Das Beste der Cloud:** Die Lösung von HPE unterstützt Wachstum und Flexibilität mit der HPE GreenLake-Plattform und einer skalierbaren Infrastruktur, die veränderten Geschäftsanforderungen gerecht wird.



**4. Sicherheit:** Umfassende Sicherheitsvorkehrungen gewährleisten den Datenschutz und die Einhaltung von Vorschriften, wobei die mit der Einführung von KI verbundenen Schwachstellen berücksichtigt werden.



**5. KI-Governance und Management:** Die Lösung vereinfacht die KI-Governance durch einen einheitlichen Zugriffspunkt und eine einheitliche Kontrolle, die das Management in hybriden Umgebungen optimieren.



**6. Strategische Partnerschaften:** HPE nutzt die Kooperation mit NVIDIA und mit Systemintegratoren, um die Funktionen zu verbessern. Gleichzeitig wird dadurch ein leistungsfähiger Support für KI-Implementierungen angeboten.



## Empfehlungen: Argumente für Hybrid/Private Clouds gegenüber der Public Cloud für KI-Implementierungen

Bei der Implementierung von KI bieten Hybrid/Private Cloud-Lösungen erhebliche Vorteile für Unternehmen. Um die KI-Anwendungsfällen zu identifizieren, die den größten Mehrwert bieten, werden fundierte Kenntnisse der Geschäftsanforderungen und die Fähigkeit benötigt, diese in KI-Projekte umzusetzen. Bei der Entwicklung und Operationalisierung dieser Anwendungsfälle müssen KI-Modelle in bestehende Workflows integriert werden, um sicherzustellen, dass sie greifbare Geschäftsergebnisse liefern.

Erfolgsentscheidend ist die effektive Nutzung von Daten in KI-Modellen für aussagekräftige Einblicke, bei denen kein sensibles geistiges Eigentum gefährdet wird. Indem Unternehmen sich eine strenge Kontrolle über die Daten sichern, können sie proprietäre Algorithmen und sensible Informationen vor unbefugtem Zugriff oder Verstößen schützen und so die üblichen Sicherheitsbedenken im Zusammenhang mit Public Cloud-Lösungen vermeiden.

Bei der Skalierung von KI-Initiativen in Reaktion auf wachsende Anforderungen die Kosten effektiv kontrollieren zu können, ist extrem wichtig. Unternehmen müssen den Bedarf an Rechenleistung und Speicherplatz mit den finanziellen Auswirkungen abwägen und sicherstellen, dass KI-Implementierungen bei der Skalierung kosteneffizient bleiben.

Die Optimierung des Technologie-Stacks für spezifische KI-Anwendungen ist entscheidend für den Erfolg. Unternehmen müssen ihre Technologie anpassen und abstimmen, um die Anforderungen verschiedener KI-Workloads zu erfüllen und die bestmöglichen Lösungen für verschiedene Anwendungsfälle zu finden.

KI-Inferenz und Modellanpassung in der Nähe der Datenquellen durchzuführen, ist ein weiterer wichtiger Faktor. Dieser Ansatz reduziert Latenzen und verbessert die Leistung, weil die Geschwindigkeit und die Genauigkeit der KI-Modelle zuverlässig verbessert werden.

KI-Lösungen in einer Private Cloud bieten mehrere entscheidende Vorteile für Unternehmen, die KI-Lösungen sicher an Fremd- oder Edge-Standorten sowie in Rechenzentren implementieren möchten. Die Nähe zur Datenquelle gewährleistet einen schnelleren Zugriff und eine beschleunigte Verarbeitung, was die Effizienz und Leistung von KI-Modellen verbessert. Eine umfassende Governance vereinfacht das Management, klärt Verantwortungsverhältnisse und unterstützt die Sicherheit kooperativer Workflows. Die Kostenkontrolle wird durch flexible Finanzierungsmodelle verbessert, sodass die hohen Kosten, die mit Public Cloud-Lösungen einhergehen, vermieden werden. Darüber hinaus wird durch die Optimierung des Technologie-Stacks für bestimmte Anwendungsfälle sichergestellt, dass KI-Implementierungen auf maximale Wirkung und Effizienz zugeschnitten sind.



## Fazit

Die rasante Entwicklung der Generativen KI birgt für Unternehmen sowohl Chancen als auch Herausforderungen. Wenn sie das Potenzial von KI voll ausschöpfen wollen, müssen Unternehmen den Überblick über komplexe Technologien gewinnen und auf robuste Sicherheit, Skalierbarkeit und eine effektive Kostenkontrolle achten. NVIDIA AI Computing by HPE bietet mit seiner skalierbaren, sicheren und leistungsoptimierten KI-Infrastruktur eine umfassende Lösung für diese Herausforderungen. Unternehmen können die sofort nutzbare Private Cloud für KI und die strategischen Partnerschaften von HPE für sich nutzen, um ihre KI-Initiativen zu beschleunigen, eine kürzere Markteinführungszeit zu erreichen und gleichzeitig die volle Kontrolle über ihre Daten und Infrastrukturen zu behalten. HPE wird auch künftig Innovationen vorantreiben und Unternehmen dabei unterstützen, das transformative Potenzial von KI für sich zu erschließen. Kontinuierliche Entwicklungen und strategische Kooperationen unterstützen das Ziel von HPE, mit innovativen Lösungen in der KI-gestützten Zukunft zum Erfolg zu verhelfen.

# Wichtige Informationen zu diesem Bericht

## MITWIRKENDE

### Steven Dickens

Research Analyst | The Futurum Group

## HERAUSGEBER

### Daniel Newman

CEO | The Futurum Group

## ANFRAGEN

Kontaktieren Sie uns, wenn Sie Fragen, Kommentare oder Anregungen zu diesem Bericht haben. The Futurum Group wird zeitnah antworten.

## ZITIERUNG

Dieses Dokument darf von akkreditierten Pressevertretern und Analysten zitiert werden. Die Zitierung muss jedoch im Kontext erfolgen und den Namen sowie Titel des Autors/der Autorin und „The Futurum Group“ enthalten. Andere Personenkreise benötigen für Zitierungen eine vorherige schriftliche Genehmigung von The Futurum Group.

## LIZENZIERUNG

Dieses Dokument, einschließlich aller Begleitmaterialien, ist Eigentum von The Futurum Group. Diese Publikation darf ohne vorherige schriftliche Genehmigung durch The Futurum Group in keinerlei Form vervielfältigt, verbreitet oder geteilt werden.

## HINWEISE

The Futurum Group bietet Dienstleistungen wie Forschung, Analyse, Beratung und Consulting für viele Hightech-Unternehmen an, auch für die in diesem Dokument genannten. Kein Mitarbeiter von Moor Insights & Strategy besitzt Aktienanteile an den in diesem Dokument genannten Unternehmen.



## Hewlett Packard Enterprise

### ÜBER HPE

HPE Financial Services kombiniert technologische Erkenntnisse, Finanzwissen und einen starken Fokus auf Nachhaltigkeit, um intelligentere IT-Lebenszyklen für Kunden und Partner jeder Größe zu schaffen. Unser kollaborativer Ansatz umfasst die gesamte technische Infrastruktur, vom Edge über die Cloud bis zum Endbenutzer, und liefert Lösungen für das Asset Management, die nicht nur Kapital freisetzen und die Kapazität maximieren, sondern auch nachhaltige Praktiken global und konsequent fördern. Hintergrundinformationen finden Sie unter: [hpe.com/hpefinancialservices](https://hpe.com/hpefinancialservices)



## NVIDIA ÜBER NVIDIA

[NVIDIA](https://www.nvidia.com) entwickelt die leistungsfähigsten Chips, Systeme und Software für die KI-Factories der Zukunft. Wir entwickeln neue KI-Services, mit denen Unternehmen ihre eigenen KI-Factories aufbauen können. Seit mehr als 30 Jahren nutzen Wissenschaftler, Forscher, Entwickler und Kreative die Technologie von NVIDIA, um Erstaunliches zu leisten. Mehr als 4 Millionen Entwickler erstellen inzwischen Tausende von Anwendungen für Accelerated Computing. Mehr als 40.000 Unternehmen nutzen die KI-Technologien von NVIDIA und 15.000 Startups weltweit sind Mitglieder von NVIDIA Inception.



### ÜBER THE FUTURUM GROUP

[The Futurum Group](https://www.thefuturumgroup.com) ist ein unabhängiges Forschungs-, Analyse- und Beratungsunternehmen, das sich mit Fragen der digitalen Innovation und marktverändernden Technologien und Trends beschäftigt. Jeden Tag unterstützen unsere Analysten, Marktforscher und Berater Wirtschaftsführer aus aller Welt dabei, grundlegende Veränderungen in ihrer Branche zu antizipieren und disruptive Innovationen zu nutzen, um Wettbewerbsvorteile in ihren Märkten zu gewinnen oder zu behaupten.



## KONTAKTINFORMATIONEN

The Futurum Group LLC | [futurumgroup.com](https://www.futurumgroup.com) | (833) 722-5337 |