



**HPE Private Cloud
AI: Unterstützung
für ITOps und Data
Scientists**

Unternehmen auf der ganzen Welt sind sich der transformativen Kraft der KI bewusst und befinden sich mitten in einem strategischen Wettlauf, diese zu ihrem Vorteil im Wettbewerb einzusetzen. Doch für viele ist der Weg zum effektiven Einsatz von KI mit Herausforderungen verbunden. Sie kämpfen mit erheblichem Gegenwind: der schwindelerregenden Geschwindigkeit, mit der sich die KI-Technologie weiterentwickelt, ihrer inhärenten Komplexität und, was vielleicht am kritischsten ist, der wachsenden Kluft zwischen der Nachfrage nach hochspezialisierten KI-Talenten und deren geringer Verfügbarkeit.

Das Gebot der Private Cloud

Die Bewältigung dieser Herausforderungen führt uns zu der zentralen Frage des KI-Einsatzes: Wo wird sich Ihre Intelligenz befinden? Die Wahl ist eindeutig: Entweder riskieren Sie, wichtige Daten und geistiges Eigentum des Unternehmens in der Public Cloud preiszugeben, oder Sie sichern sie in einer privaten Umgebung, die auf Kontrolle und Compliance ausgelegt ist. Der Wechsel hin zur Private Cloud für KI ist nicht bloß ein Trend, sondern eine rasante Beschleunigung, die von Kunden vorangetrieben wird, die sich der Bedenken hinsichtlich der Datensouveränität und der oft unvorhersehbaren, steigenden Kosten sehr wohl bewusst sind, die mit der Public Cloud verbunden sind. Doch der wahre Erfolg Ihrer Private Cloud hängt von ihrer Architektur und ihren Funktionen ab.

Einer der Schlüssel zu einer erfolgreichen Private Cloud-Strategie liegt eindeutig in der zugrunde liegenden Infrastruktur. Sie ist die entscheidende Grundlage für jede erfolgreiche KI-Initiative und bietet die unverzichtbaren Komponenten, die für die effektive Entwicklung, Trainierung, Bereitstellung und Skalierung von KI erforderlich sind. Bedenken Sie die Kosten, die entstehen, wenn Sie ein Jahr oder mehr damit verbringen, Ihre KI-Infrastruktur von Grund auf neu oder als Referenzarchitektur aufzubauen. Ein solches Unterfangen wird zu einem erheblichen Engpass und verzögert direkt die Integration der Softwareschicht, die für greifbare KI-Ergebnisse erforderlich ist. Im Wesentlichen bedeutet jede Verzögerung beim Aufbau Ihrer KI-Infrastruktur eine Verlangsamung oder völlige Blockierung auf dem Weg zu einer wirkungsvollen KI.

Beschleunigen Sie die KI-Bereitstellung und verringern Sie die Komplexität

Möchten Sie eine einfachere und schnellere KI-Bereitstellung? Ihre Antwort ist HPE Private Cloud AI, gemeinsam mit NVIDIA entwickelt. Diese integrierte End-to-End-Plattform ist darauf ausgelegt, Ihre Private-KI-Cloud in nur einem Tag zum Laufen zu bringen. Es handelt sich um eine werkseitig integrierte Lösung aus Hardware, Software und Services, die alles bietet, was Sie für gängige KI-Anwendungsfälle benötigen. Nach der Installation aktualisiert HPE die Lösung mit den neuesten GPU-, CPU- und Hardware-Innovationen. Durch Upgrades mit einem Klick wird sichergestellt, dass Kunden Zugriff auf die neueste KI-Software und die aktuellsten KI-Tools und -Modelle haben.

KI-Lösungen für jede Phase und Größenordnung

Jedes Unternehmen befindet sich in einer anderen Phase der KI-Einführung. Aus diesem Grund ist HPE Private Cloud AI keine Einheitslösung. Die Lösung wird basierend auf Ihren Geschäftsanforderungen an einem von vier skalierbaren Startpunkten bereitgestellt. Von der Entwicklung bis hin zu unternehmensweiten Produktionsanwendungsfällen über Inferenz, Retrieval Augmented Generation (RAG) und Feinabstimmungsmodelle. Mit HPE Private Cloud AI können Sie Ihren Weg zu KI voller Zuversicht beginnen, denn Sie wissen, dass HPE Ihnen Investitionsschutz bietet, während sich die Roadmaps für KI-Technologien weiterentwickeln.

Unterstützung für ITOps

Für ITOps und Cloud-Administratoren transformiert das Management von HPE Private Cloud AI Ihre Rolle und ermöglicht Ihnen den Aufstieg vom KI-Neuling zum KI-Helden. Über eine einheitliche Schnittstelle erhalten Sie die vollständige Kontrolle: Definieren Sie mühelos Cluster, fügen Sie Benutzer und ihre Rollen hinzu, legen Sie Zugriffskontrollen fest, sorgen Sie für zuverlässige Sicherheit, überwachen Sie die Ressourcenzuweisung und erweitern, beobachten und verwalten Sie nahtlos mehrere HPE Private Cloud AI Cluster.

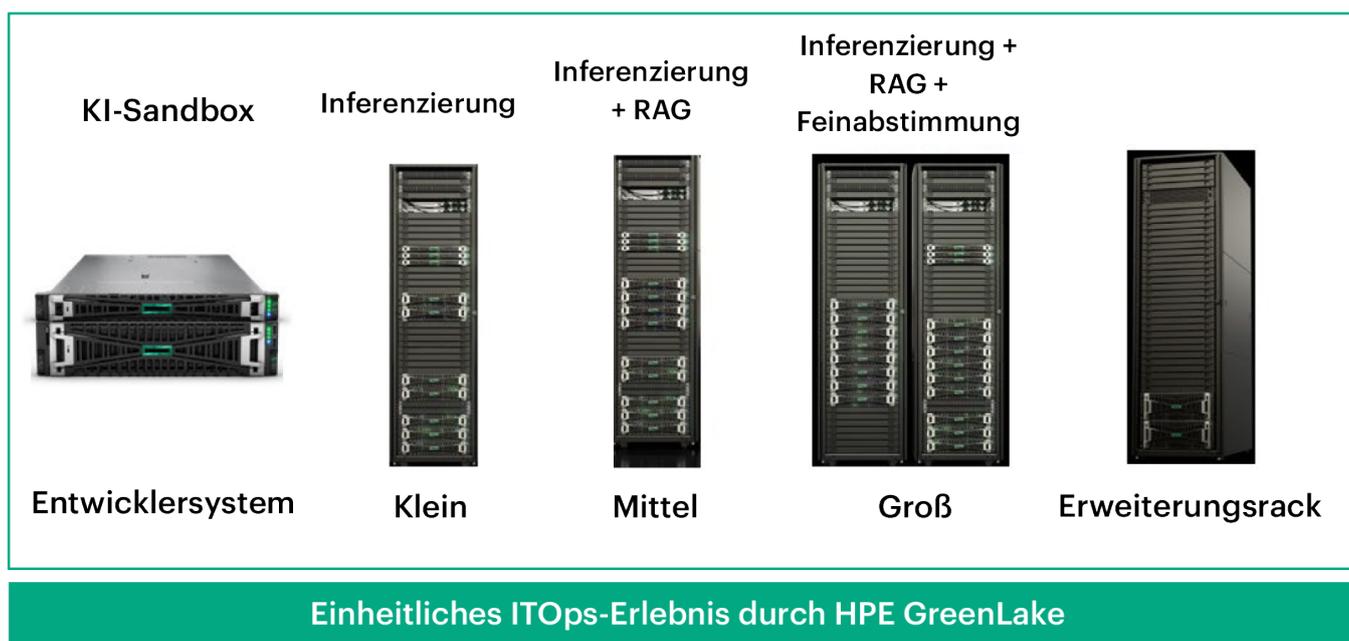


Abbildung 1. HPE Private Cloud AI: die richtige Größe für jede Phase auf Ihrem Weg zu KI

Die Leistungsfähigkeit dieser Einfachheit beruht auf automatisierten Back-End-Prozessen. Diese intelligenten Systeme richten Konfigurationen sogar auf der komplexen Kubernetes-Ebene genau nach Ihren Vorgaben ein, ohne dass Sie ein Experte sein müssen. Definieren Sie in der HPE GreenLake Cloud einfach Ihre Richtlinien, die sofort an die integrierte Softwareebene, die HPE AI Essentials Software, weitergegeben werden. Dieser Prozess erstellt plattformweite Konfigurationen, reduziert die mühsame Einrichtung einzelner Anwendungen für jeden Benutzer und ermöglicht Konsistenz in Ihrer gesamten KI-Umgebung.

Durch Air Gap geschützte KI für ultimative Sicherheit

Für Kunden, die in dunklen oder hochsicheren Umgebungen arbeiten, steht HPE Private Cloud AI in einer durch einen Air Gap geschützten Version zur Verfügung. Ihre vertraulichen Daten bleiben vollständig geschützt, sodass Sie Modelle in einer Umgebung anpassen und inferenzieren können, die vollständig von externen Netzwerken isoliert ist.

Wie das vernetzte Gegenstück verfügt auch die durch einen Air Gap geschützte Version über zuverlässige Sicherheits- und Compliance-Funktionen. Dadurch wird sie zu einer idealen Lösung für Unternehmen, die strengsten regulatorischen und Souveränitätsanforderungen unterliegen, beispielsweise im Militär-, staatlichen, Gesundheits- und Finanzsektor.

Herzstück der KI-Innovation

Das Herzstück von HPE Private Cloud AI ist die HPE AI Essentials Software, eine integrierte und äußerst leistungsfähige Lösung, die Ihren Weg zu KI beschleunigen soll. Diese Software bietet hochmoderne, sofort einsatzbereite KI- und Open-Source-Tools und ermöglicht es KI-Fachleuten, KI-Workloads in beispielloser Geschwindigkeit zu erstellen. Mit der HPE AI Essentials Software erhalten KI-Anwender unterschiedlichster Art nahtlosen Zugriff auf integrierte Daten, KI-Tools und -Modelle sowie Low-Code-Assistenten. Dies ermöglicht ihnen die mühelose Erstellung, Bereitstellung und Governance von ML-Pipelines sowie das Management von Modellen. Und das Beste: Wartung und Updates werden von HPE automatisch durchgeführt. So können sich Ihre Teams auf Innovationen und die Erzielung konkreter KI-Ergebnisse konzentrieren, anstatt sich mit komplexen Updates herumzuschlagen.

Der integrierte Modellkatalog bietet personenbasierten Zugriff auf Ihre bevorzugten Tools, einschließlich NVIDIA® NIM™ als Teil von NVIDIA AI Enterprise. Sie können ein bestimmtes Tool, das Sie bevorzugen, nicht finden oder haben eine bewährte Drittanbieteranwendung? Unser intuitiver Assistent hilft Ihnen, Helm-basierte Tools, Anwendungen oder NVIDIA Blueprints in nur wenigen Schritten zu importieren. Nach Konfiguration in die Plattform unterliegen diese importierten Anwendungen dem robusten Sicherheits- und Governance-Framework der Plattform.

Bringen Sie KI zu Ihren Daten

KI ist eine datenhungrige Anwendung, und Daten sind der Schlüssel zur Erzielung zuverlässiger Ergebnisse. Sie sind der Treibstoff der KI und ermöglichen es Modellen, zu lernen, sich anzupassen und intelligente Entscheidungen zu treffen. Die Qualität und Relevanz dieser Daten sind von größter Bedeutung, da fehlerhafte oder unzureichende Daten zwangsläufig zu fehlerhaften KI-Ergebnissen führen.

Die HPE AI Essentials Software umfasst ein integriertes Data Lakehouse Gateway, das autorisierten Benutzern unabhängig von Format oder Standort sicheren und transparenten Zugriff auf Ihre Unternehmensdaten bietet. Ein globaler Namespace verbindet operative Datenbanken über S3- und NFS-Endpunkte nahezu in Echtzeit direkt mit dem Data Lakehouse. Durch diese Funktion entfällt die Notwendigkeit, Daten vor der Verarbeitung zu verschieben, und Benutzer erhalten effektiv eine konsistente, einheitliche Ansicht der Daten aus unterschiedlichen Quellen.

Dieser einheitliche Zugriff wird durch feinkörnige Zugriffskontrollen geregelt, die für Datensicherheit sorgen und gleichzeitig eine nahtlose Zusammenarbeit ermöglichen. Darüber hinaus können Kunden ihre bevorzugten Analyse-Engines nutzen – unabhängig davon, ob es sich um Microsoft SQL, Apache Spark, BI-Tools oder KI-/ML-Tools handelt –, um gemeinsam an Daten zu arbeiten, die in verschiedenen Architekturen gespeichert sind.

Multi-Tenancy für Unternehmen

Die HPE AI Essentials Software verbessert ihre kollaborativen Fähigkeiten durch die Einführung einer robusten Mandantenfähigkeit für Unternehmen. Diese Weiterentwicklung ermöglicht mehreren Benutzern die nahtlose Zusammenarbeit an Projekten in einer

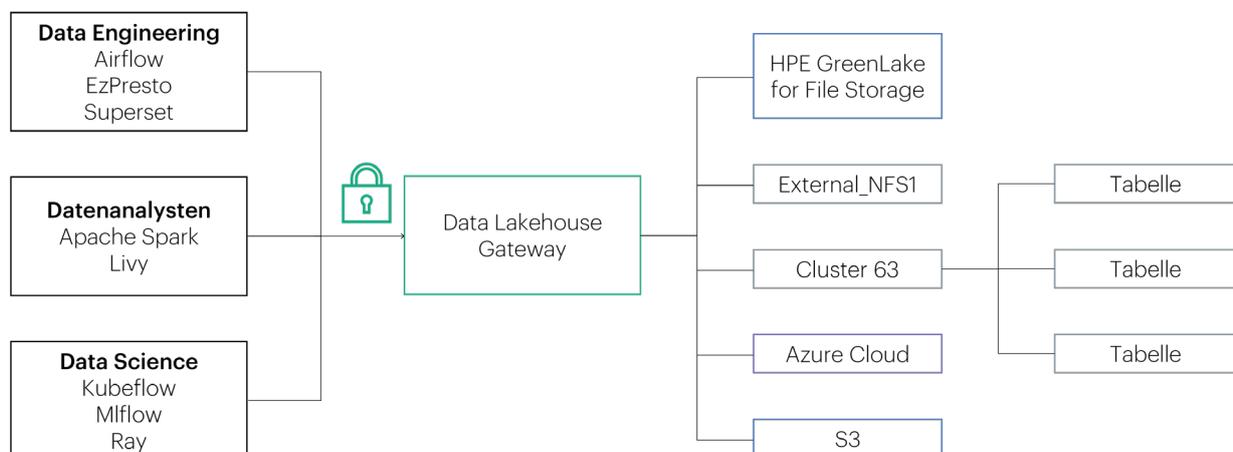


Abbildung 2. Data Lakehouse Gateway ermöglicht einheitlichen, geregelten Zugriff auf Daten aus verschiedenen Quellen

sicheren, gemeinsam genutzten Umgebung. Administratoren erhalten eine granulare Kontrolle über Rollen und Berechtigungen und können so festlegen, wer auf was zugreifen kann. Außerdem können sie den Zugriff auf Datenquellen basierend auf einzelnen Projekten präzise festlegen. Darüber hinaus können mit der Möglichkeit zur Festlegung von GPU-Kontingenten Ressourcen effizient zwischen verschiedenen Teams und Projekten zugewiesen und verwaltet werden, was optimale Leistung und eine faire Nutzung für die Benutzer ermöglicht. Durch dieses Multi-Tenancy-Modell können Teams ihre KI-Initiativen mit größeren Effizienz-, Sicherheits- und Kollaborationsfähigkeiten beschleunigen.

Schaffen Sie schnell KI-Wert

Neben nahtlosem Datenzugriff und Multi-Tenancy steht Data Engineers und Data Scientists ein vorintegrierter Katalog von KI- und Datentools zur Verfügung. Dadurch wird die Komplexität der Infrastruktur verringert und es können Innovationen entwickelt und Lösungen schnell bereitgestellt werden. Über die vorintegrierten Angebote hinaus erhalten Datenexperten volle Flexibilität. Sie können Modelle und Tools aus externen Repositories wie Hugging Face sowie NVIDIA Blueprints nahtlos direkt in den Katalog importieren und integrieren.

Tools und Frameworks Framework importieren

NVIDIA AI Enterprise **Data Engineering** Analyse Data Science

AirFlow
Version 2.10.0 | Bereit

Workflow-Engine für die geplante Batch-Aufnahme von Daten aus einer Vielzahl externer Datenquellen

Endpoint <https://airflow.ingress.hou-pcai.hpepci.net>

Chart Version 1.4.4

[Öffnen](#)

EzPresto
Version Prestodb: 0.287-1.6.0 | Bereit

Verteilte Abfrage-Engine für analytische Abfragen von Daten jeder Größe

Endpoint <https://ezpresto.ingress.hou-pcai.hpepci.net>

Chart Version 1.6.0

[Öffnen](#)

Superset
Version 4.0.1 | Bereit

Apache Superset ist eine moderne, Business Intelligence-Webanwendung für Unternehmen

Endpoint <https://superset.ingress.hou-pcai.hpepci.net>

Chart Version 0.11.217

[Öffnen](#)

Abbildung 3. Data Engineers profitieren vom sofortigen Zugriff auf Airflow für die Orchestrierung, EzPresto für Abfragen und Superset für die Visualisierung von Datenpipelines

Verabschieden Sie sich von den Problemen der komplexen Datentransformation. Der Modellkatalog vereinfacht alles mit vorab validierten Hardware- und Software-Stacks, die eine nahtlose Interoperabilität ermöglichen, sodass Integrationsprobleme der Vergangenheit angehören. Data Engineers können direkt auf die Daten im Data Lakehouse zugreifen und so eine mühsame Vorabnahme oder Datenverschiebung umgehen. Dank dieser neu gewonnenen Freiheit können sie ihre Energie auf die Entwicklung zuverlässiger Pipelines konzentrieren, ohne von der Komplexität der Infrastruktur und von Datensilos belastet zu werden.

Tools und Frameworks Framework importieren

NVIDIA AI Enterprise Data Engineering **Analyse** **Data Science**

Importiert **cs-test**
Version 0.2.11 | Bereit

cs-test

Endpoint <https://hpe-customer-support.prod.discover.hpepci.com>

Chart Version 0.2.11

[Öffnen](#)

Kubeflow
Version 1.9.0 | Bereit

Notebooks, Modelltraining, Pipelines und Model Serving-Toolset für maschinelles Lernen

Endpoint <https://kubeflow.prod.discover.hpepci.com>

Chart Version 1.2.20

[Öffnen](#)

Ray
Version 2.35.0 | Bereit

Einheitliches Framework zum Skalieren von KI- und Python-Anwendungen.

Endpoint <https://ray-dashboard.prod.discover.hpepci.com>

Chart Version 0.5.4

[Öffnen](#)

Importiert **LF**
Neueste Version | Unbekannt

Test

Endpoint <https://testendpoint.prod.discover.hpepci.com>

Chart Version 0.1.3

MLflow
Version 2.16.0 | Bereit

MLflow ist eine Plattform zur Optimierung der Entwicklung von maschinellem Lernen, einschließlich der Verfolgung von Experimenten und des ML-Lebenszyklus

Endpoint <https://mlflow.prod.discover.hpepci.com>

Chart Version 0.7.37

HPE MLIS
Version 1.2.0 | Bereit

Die HPE Machine Learning Inference Software ist eine benutzerfreundliche Lösung, die die Bereitstellung vereinfacht und steuert...

Endpoint <https://mlis.prod.discover.hpepci.com>

Chart Version 1.2.0

Abbildung 4. Direkt nutzbarer Modellkatalog für Data Scientists. Elemente mit einem gelben Band wurden in die HPE AI Essentials Software importiert

Eine vereinfachte Infrastruktur gibt Data Scientists außerdem die Freiheit, zu experimentieren. Dank vorintegrierter Tools können sie sich voll und ganz auf die Entwicklung fortschrittlicher Modelle, aufschlussreiche Experimente und die Schaffung greifbarer KI-Werte konzentrieren. Anpassbare Jupyter-Notebooks und JupyterLab-Umgebungen bieten die exakte Kontrolle, die erforderlich ist, um Anwendungen genau nach ihren Spezifikationen zu erstellen, zu entwickeln, auszuführen und bereitzustellen.

Reichhaltiges ISV-Ökosystem

Die HPE AI Essentials Software bietet durch ihr reichhaltiges Ökosystem validierter ISV-Partner einen außergewöhnlichen Mehrwert. Dieses umfangreiche Netzwerk ergänzt die Plattform, indem es eine vielfältige Palette vorab validierter Lösungen und Anwendungen bereitstellt. Kunden können Dutzende vorab validierter Anwendungsfälle schnell importieren und bereitstellen, wodurch die Time-to-Value verkürzt und ein ROI bei deutlich geringerem Risiko rascher erzielt wird. Dieses reichhaltige Ökosystem, kombiniert mit Funktionen wie zuverlässiger Multi-Tenancy, ermöglicht Unternehmen den Zugriff auf die spezialisierten Tools und die kollaborative Umgebung, die sie benötigen, um komplexe KI-Herausforderungen zu bewältigen und selbstbewusst neue Möglichkeiten zu erschließen.

Wirklich unternehmensreife KI-Plattform

Kann HPE Private Cloud AI Ihre KI-Implementierungen beschleunigen? Hier sind einige wichtige Fragen, die Sie untersuchen sollten.

1. Welche konkreten Schritte im KI-Bereitstellungsprozess führen zu den größten Verzögerungen oder Ressourcenabflüssen für Ihr Team?
2. Konnten Sie den Zeit- und Ressourcenaufwand für die Integration und Wartung Ihrer KI-Infrastruktur und Ihres Software-Stacks quantifizieren?

¹ ["The economic benefits of HPE Private Cloud AI with NVIDIA AI Computing by HPE"](#), Enterprise Strategy Group, March 2025.

² ["On-premises AI approaches: the advantages of a turnkey solution, HPE Private Cloud AI"](#), Principled Technologies, April 2025.

[HPE.com besuchen](#)

Jetzt chatten

© Copyright 2025 Hewlett Packard Enterprise Development LP. Die hier enthaltenen Informationen können jederzeit ohne vorherige Ankündigung geändert werden. Neben der gesetzlichen Gewährleistung gilt für Produkte und Services von Hewlett Packard Enterprise (HPE) ausschließlich die Herstellergarantie, die in den Garantieerklärungen für die jeweiligen Produkte und Services explizit genannt wird. Aus dem vorliegenden Dokument sind keine weiterreichenden Garantieansprüche abzuleiten. Hewlett Packard Enterprise haftet nicht für hierin enthaltene technische oder redaktionelle Fehler oder Auslassungen.

Microsoft ist eine eingetragene Marke oder Marke der Microsoft Corporation in den USA und/oder anderen Ländern. NVIDIA ist eine Marke und/oder eingetragene Marke der NVIDIA Corporation in den USA und anderen Ländern. Alle Marken Dritter sind Eigentum der jeweiligen Rechteinhaber.

a00141386DEE

HEWLETT PACKARD ENTERPRISE

[hpe.com](#)

3. Können Sie etwaige Herausforderungen oder Verzögerungen beschreiben, mit denen Ihr Team in der Vergangenheit beim Versuch, KI-Anwendungen einzuführen, aufgrund von Einschränkungen der zugrunde liegenden Infrastruktur konfrontiert war?
4. Wie wirkt sich die Möglichkeit, die zugrunde liegenden Komponenten Ihrer KI anzuzeigen und zu steuern, auf die Fähigkeit Ihres Teams aus, Ihre Modelle zu debuggen, zu optimieren oder zu skalieren?

Beschleunigen Sie Ihre KI-Bereitstellung und steigern Sie den ROI

Haben Sie genug von der Komplexität und den Verzögerungen, die mit der Zusammenstellung von Hard- und Software und der Orchestrierung für KI verbunden sind? HPE Private Cloud AI durchbricht diese Komplexität. Branchenanalysten bestätigen, dass Sie mit unserer gemeinsam mit NVIDIA entwickelten Lösung im Vergleich zu Selbstbau-Initiativen sieben bis zwölf Monate Bereitstellungszeit einsparen können. Dabei geht es nicht nur um Geschwindigkeit. Auch Ihre Gesamtbetriebskosten werden erheblich gesenkt, und zwar um schätzungsweise 1,72 bis 3,57 Millionen US-Dollar, je nach Größe Ihres Systems.¹

Wenn Effizienz und schnelle Wertschöpfung für Ihre Private Cloud entscheidend sind, werden Sie HPE Private Cloud AI schätzen. Der schlüsselfertige Ansatz wurde hinsichtlich des Werts sowohl im Vergleich zu Referenzarchitekturen als auch zu selbst erstellten Alternativen als der einfachste und schnellste eingestuft.²

Die Lösung ist einsatzbereit, Sie können also mit der Integration aufhören und mit der Innovation beginnen.

Weitere Informationen

[HPE.com/Private-Cloud-AI](https://www.hpe.com/Private-Cloud-AI)

