

7 Schritte zur Erschließung des Werts von Daten für KI

Leitfaden für effektives Datenmanagement



Schaffen eines Wettbewerbsvorteils mit Daten

In der sich schnell entwickelnden Unternehmenslandschaft von heute sind Daten eine wichtige Ressource – der Eckpfeiler von Innovation und Wettbewerbsvorteilen. Führungskräfte, die ihre Geschäftsentscheidungen durch die intelligente Nutzung solider, genauer und umfassender Datensätze beschleunigen, heben sich von der Masse ab. Nirgends wird dies deutlicher als bei der Einführung von [generativer KI \(GenAI\)](#).

GenAI verspricht beispiellose Fortschritte und Effizienzen, und perfekte Daten sind sicherlich keine Voraussetzung für die kreativen, iterativen und flexiblen Ergebnisse, die sie bieten kann. Aber irgendwann müssen Sie Ihr „Datenhaus“ bewerten,

um sicherzustellen, dass alles in Ordnung ist. Mit anderen Worten, Sie benötigen mehr als nur **Zugriff** auf Daten. Sie benötigen **robuste Datenmanagementverfahren**. Dies wird mit zunehmender Skalierung noch notwendiger.

Ein starkes Datenmanagement ist notwendig, aber nicht einfach.

Die Bereitstellung von [KI-Workloads](#) ist das Ergebnis eines komplexen und anspruchsvollen Prozesses, bei dem Data Scientists und andere Personen geeignete Datensätze identifizieren und deren Sauberkeit und Vollständigkeit sicherstellen. Dieses Unterfangen geht über technische Herausforderungen hinaus und erfordert unternehmensweite Prozesse und robuste Datenmanagement-Frameworks.

64 %

der Unternehmen glauben, dass künstliche Intelligenz dazu beitragen wird, ihre Gesamtproduktivität zu steigern.¹

70 %

der Top-Performer hatten Schwierigkeiten bei der Integration von Daten in KI-Modelle²

83–92 %

aller KI-Projekte scheitern³

¹ Forbes-Berater: [So nutzen Unternehmen künstliche Intelligenz im Jahr 2024](#)

² McKinsey-Bericht: [Der Stand der KI Anfang 2024](#)

³ Fortune: [„Want your company's AI project to succeed? Don't hand it to the data scientists, says this CEO.“](#)



Die Entwicklung des Datenmanagements

Um Sie bei der Navigation durch diese komplexe Landschaft zu unterstützen, haben wir einen umfassenden Überblick über die Entwicklung des Datenmanagements erstellt, der in sieben wesentliche Schritte unterteilt ist. Diese Schritte sind das Ergebnis umfangreicher Workshops und Beratungen, die von den erfahrenen Data Scientists von Dell mit einer Vielzahl von Unternehmen durchgeführt wurden. Sie stellen gemeinsame Herausforderungen und erfolgreiche Strategien für die Erstellung skalierbarer und effektiver [KI-Modelle](#) vor. In den kommenden Monaten werden wir uns eingehender mit jedem Schritt befassen, damit Sie die besonderen Hindernisse in Ihrem Unternehmen meistern und Ihre KI-Journey erfolgreich abschließen können.

Dieses E-Book bietet einen klaren, umsetzbaren Rahmen für das Verständnis und die Umsetzung effektiver Datenmanagementpraktiken. Wenn Ihr Unternehmen diese Prinzipien beherzigt, kann es vom Experimentieren mit GenAI zur Skalierung der Datentransformation von einer Rohressource in einen strategischen Vermögenswert übergehen und so den Weg für Innovation und nachhaltige Wettbewerbsvorteile in der KI-gesteuerten Zukunft ebnen.

Laden Sie diese [Infografik](#) herunter, um eine einfache Referenz zu diesen 7 Schritten zu erhalten.



SCHRITT 1

Geschäftsanforderungen identifizieren



SCHRITT 2

Relevante Datenermittlung beschleunigen



SCHRITT 3

Datenexploration und Datenzugriff vereinfachen



SCHRITT 4

Analysen, ML-Experimente und Modellierung optimieren



SCHRITT 5

Produktion von Daten und Analysen skalieren



STEP 6

Datenmanagement und Governance automatisieren



STEP 7

Geschäftsergebnisse bewerten



SCHRITT 1

Geschäftsanforderungen identifizieren



Dieser Schritt gibt die Richtung für Ihr Datenmanagement vor und richtet die durchzuführenden Maßnahmen an den strategischen Geschäftsanforderungen aus. Ohne definierte Ziele und Metriken ist es unwahrscheinlich, dass Sie einen nennenswerten Wert erzielen.

- Beginnen Sie damit, das operative Ziel und den damit verbundenen Nutzen zu verstehen.
- Einigen Sie sich im gesamten Unternehmen auf die gewünschten Ergebnisse und die Art und Weise der Erfolgsmessung.
- Erstellen Sie eine klare Vision des zu schaffenden Werts.

Wenn Sie wissen, **was** Sie erreichen wollen und **wie** Sie es messen können, stellen Sie sicher, dass alle nachfolgenden Datenmanagementmaßnahmen sinnvoll und auf klare, erreichbare Ziele ausgerichtet sind.

⁴ Technologie von Seagate: [Daten überdenken: Machen Sie mehr aus Ihren Daten – vom Edge bis zur Cloud.](#)

SCHRITT 2

Relevante Datenermittlung beschleunigen



Beschleunigen Sie die Ermittlung relevanter Daten mit einem klaren Fahrplan. Nicht alle verfügbaren Daten sind notwendig. Data Scientists müssen schnell Daten identifizieren, die für die Lösung des Problems relevant sind. Stellen Sie eine klare Verbindung zwischen Daten und ihrem Wert her, indem Sie Quelldatensätze katalogisieren und Metadaten erstellen. Dieser fokussierte Ansatz stellt sicher, dass Ihre Datenbemühungen effizient und effektiv sind.

Durch die schnelle Ermittlung relevanter Daten sparen Sie **Zeit** und **Ressourcen**, sodass sich Ihre Data Scientists auf Daten mit hoher Wirkung konzentrieren können, die Ihre strategischen Ziele direkt unterstützen.

68 %

der Daten, die Unternehmen zur Verfügung stehen, bleiben ungenutzt⁴



SCHRITT 3

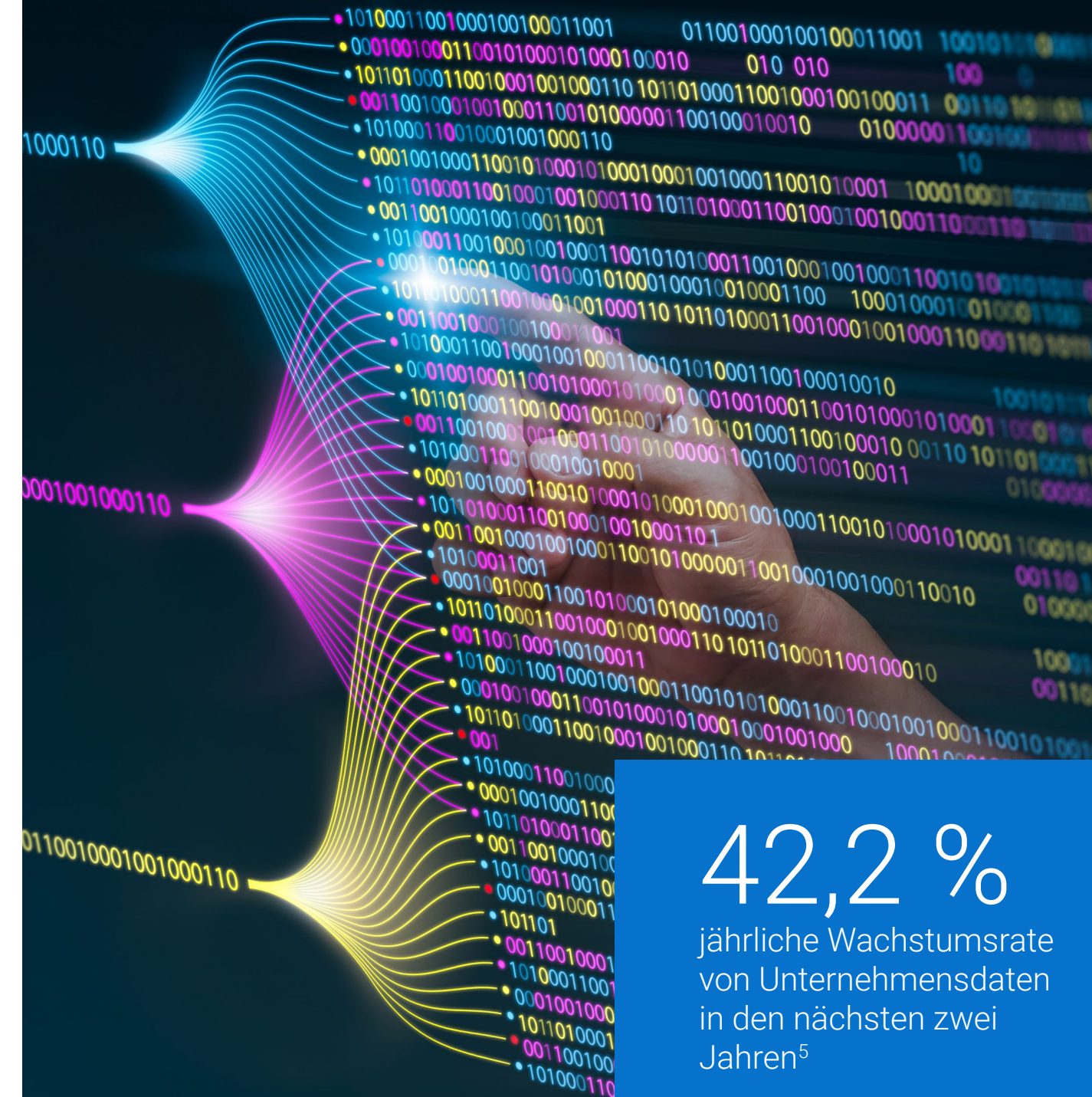
Datenexploration und Datenzugriff vereinfachen



Wenn Sie Daten zur Analyse an zentralisierte Speicherorte verschieben, verringern Sie den Wert von Echtzeit-Anwendungsfällen. Vermeiden Sie diese Ineffizienz, indem Sie Data Scientists den Zugriff auf Daten dort ermöglichen, wo sie abgelegt sind. Standardisieren Sie Prozesse und implementieren Sie die automatische Erkennung, um strukturierte und unstrukturierte Daten an leicht zugänglichen Speicherorten anzuordnen und so die Erkennung und Nutzung in Echtzeit zu erleichtern. Dieser Ansatz spart nicht nur Zeit, sondern erhöht auch die Agilität Ihrer Datenvorgänge und ermöglicht **schnellere Einblicke** und eine **reaktionsschnellere Entscheidungsfindung**.

Angesichts des explosionsartigen Datenwachstums ist ein effektives Datenmanagement erforderlich, um [KI-Workloads](#) erfolgreich zu skalieren. Tools wie [Dell Data Lakehouse](#) bieten einen vollständigen Lösungsstack für die Erkennung, Abfrage und Verarbeitung aller Unternehmensdaten, unabhängig von Standort oder Datenquelle, mit einer einzigen föderierten Abfrage-Engine.

5 Technologie von Seagate: [Daten überdenken: Machen Sie mehr aus Ihren Daten – vom Edge bis zur Cloud](#).





SCHRITT 4

Analysen, ML-Experimente und Modellierung optimieren



Regen Sie zu ständigen Experimenten und Modellierungen an, um die Variablen zu identifizieren, die Probleme lösen können. Erwägen Sie die Erstellung synthetischer Daten, die besonders bei Herausforderungen in Bezug auf Datenqualität und Datenschutz hilfreich sind und den Prozess beschleunigen können, insbesondere wenn Sie Ihre KI-Journey beginnen. Die Nutzung vorab trainierter grundlegender Modelle, die nur erweitert und optimiert werden müssen, kann ein guter Ausgangspunkt sein.

Sobald Sie Ihr Projekt erweitert haben, konzentrieren Sie sich auf den einfachen Zugriff auf Daten. Auf diese Weise kann Ihr Team Analysen durch iterative Tests optimieren. Eine Plattform, die mehrere Iterationen und Algorithmen unterstützt, kann wichtige Datenvariablen aufdecken, die Effektivität der generativen KI steigern und den Wert von Daten erschließen. Dieser iterative Prozess des Testens, Lernens und Verfeinerns stellt sicher, dass Ihre **Modelle robust** und Ihre **Erkenntnisse umsetzbar** sind, was kontinuierliche Verbesserungen und Innovationen vorantreibt.

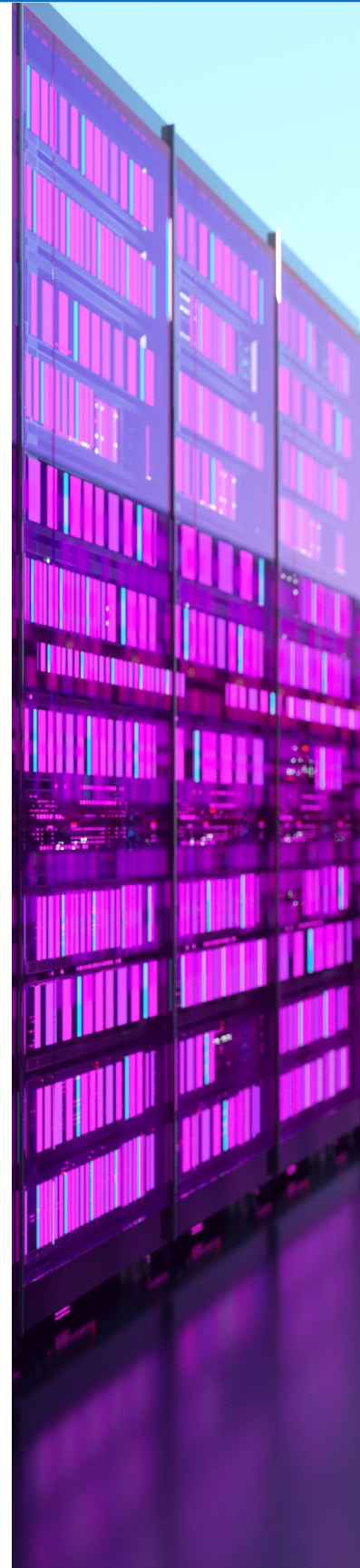


SCHRITT 5

Produktion von Daten und Analysen skalieren



Wechseln Sie von einem Data-Science-**Projekt** zu einem zuverlässigen, wiederholbaren Data-Science-**Produkt**. Diese Produkte können unabhängig betrieben und regelmäßig auf Verbesserungen überprüft werden. Ein effizienter Datenerfassungsprozess beschleunigt die Zeit bis zur Gewinnung von Informationen und ermöglicht Datenpipelines mit integrierter Intelligenz, um Daten in Echtzeit zu analysieren und zu nutzen und so vordefinierte Ziele zu erreichen. Durch die Skalierung von Datenprodukten stellen Sie Konsistenz und Zuverlässigkeit sicher und verwandeln **Ad-hoc-Projekte** in **nachhaltige, wertschöpfende Vermögenswerte**.



STEP 6

Datenmanagement und Governance automatisieren



Wenn Produkte wiederholbar werden, sollten Sie die Datenmanagement- und Governance-Prozesse Ihres Unternehmens automatisieren. Auf diese Weise kann das System sich selbst überwachen und Anomalien melden, bevor sie zum Problem werden. Setzen Sie auf ganzheitliche Observability, um sicherzustellen, dass Workflows während des gesamten Datenmanagementprozesses integrierten Standards, Policies und Sicherheitsmaßnahmen für die Data Governance entsprechen. Die Automatisierung **reduziert manuellen Aufwand** und **verbessert die Genauigkeit und Compliance**, wodurch ein robustes Framework für das Management von Daten in großem Maßstab entsteht.



STEP 7

Geschäftsergebnisse bewerten



Werten Sie Daten aus Geschäftsergebnissen aus und lassen Sie diese in den Prozess einfließen. So können Sie eine Kultur des kontinuierlichen Lernens und Anpassens fördern. Diese fortlaufende Auswertung erfasst tatsächliche Ergebnisse im Vergleich zu vorhergesagten Ergebnissen und ermöglicht damit eine Verbesserung und Optimierung dieser. Die Nutzung von Funktionen mit minimalem menschlichem Eingriff gewährleistet eine nachhaltige Leistung und einen nachhaltigen Wert. Durch die kontinuierliche Analyse der Ergebnisse schaffen Sie eine Feedbackschleife, die zu **kontinuierliche Verbesserungen** führt und **Ihre Dateninitiativen** an sich ändernden Geschäftszielen ausrichtet.



Umsetzung von Daten mit Dell AI Factory with NVIDIA.

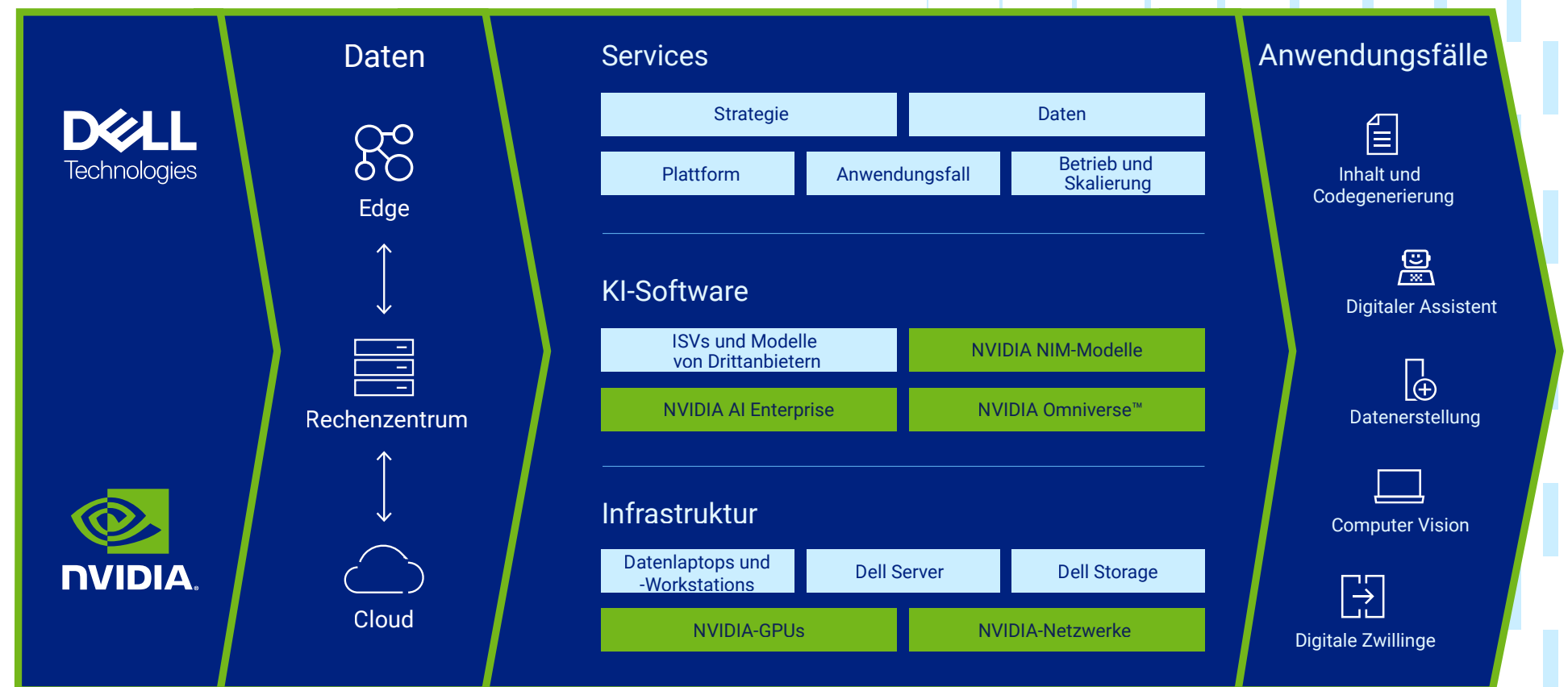
Ihr aktuelles Rechenzentrum und IT-Betriebsmodell sind wahrscheinlich nicht dafür ausgelegt, die Geschwindigkeit und den Umfang von GenAI zu nutzen. Was sie brauchen, ist eine neue Art von Rechenzentrum, das speziell für die spezifischen Anforderungen von [KI](#) entwickelt wurde.

Sie brauchen die AI Factory.

So wie physische Fabriken die industrielle Revolution angetrieben haben, werden „AI Factories“ die KI-Revolution vorantreiben. Doch statt physischer Güter produzieren AI Factories Intelligenz. Sie sind vielseitig und flexibel und können überall dort eingesetzt werden, wo sich Ihre Daten befinden – in Clouds, in Rechenzentren, auf Workstations und an Edge-Standorten.

2,4 x jährliche Steigerung der Ausgaben für das Training umfangreicher ML-Modelle – hauptsächlich für Hardware wie Grafikprozessoren und Server⁶

⁶ Epoch AI: [Wie viel kostet das Trainieren von Frontier AI-Modellen?](#)





Das Verständnis Ihrer Anwendungsfälle und die Anwendung der richtigen Datenmanagementlösungen sind die Grundlage für die Optimierung Ihrer KI-Bemühungen.

Dell AI Factory with NVIDIA vereint die erforderlichen Bausteine, mit denen Sie Ihre KI-Innovationen beschleunigen können:

- Fokus auf Daten
- KI-optimierte Infrastruktur
- KI-Software und -Modelle
- Offene Ökosysteme
- Expertenservices
- und Best Practices

Es handelt sich um eine Technologie, die sich an Ihren geschäftlichen Anforderungen orientiert, um schnell die richtige Lösung mit besseren Gesamtbetriebskosten zu liefern. Dell und NVIDIA haben mehr als 340.000 Stunden darauf aufgewendet, diese Lösungen zu validieren und Referenzdesigns und Benchmarking bereitzustellen.

Gemeinsam bieten NVIDIA und Dell eine Full-Stack-Lösung, die Ihre KI-Einführung beschleunigt: die **Dell AI Factory with NVIDIA**.

NVIDIA AI Enterprise bietet Ihnen folgende Vorteile:



Zugriff, Management, Bereitstellung und Support

GenAI-Modelle, -Anwendungen, -Lösungen und -Accelerators in einer Reihe von Infrastrukturlösungen.



Nutzung von NVIDIA NIMS

für eine robuste Datenpipeline, um alle Vorteile dieser Datenpipeline voll auszuschöpfen



Einfacher Übergang

von einem Modell zum nächsten, wobei die Integrität Ihrer Datenpipelines gewahrt bleibt, was Investitionsschutz bietet und es Ihnen ermöglicht, immer die besten verfügbaren Modelle zu nutzen



Nachhaltiger Erfolg durch kontinuierliches Datenmanagement

Das Datenmanagement ist keine einmalige Aufgabe. Es handelt sich um einen fortlaufenden Prozess, bei dem eine Reihe von Praktiken und Tools zum Einsatz kommen, mit denen Sie fortlaufend Wert aus Ihren Daten schöpfen können. Und dies ist für jede erfolgreiche GenAI-Initiative von entscheidender Bedeutung. Dell bietet in Zusammenarbeit mit NVIDIA das weltweit umfassendste KI-Lösungsportfolio.



End-To-End-Lösungen

Vom Desktop über das Rechenzentrum bis zur Cloud – unsere durchgängigen Lösungen, unterstützt von NVIDIA KI, bieten Sicherheit, Zugänglichkeit und Skalierbarkeit.



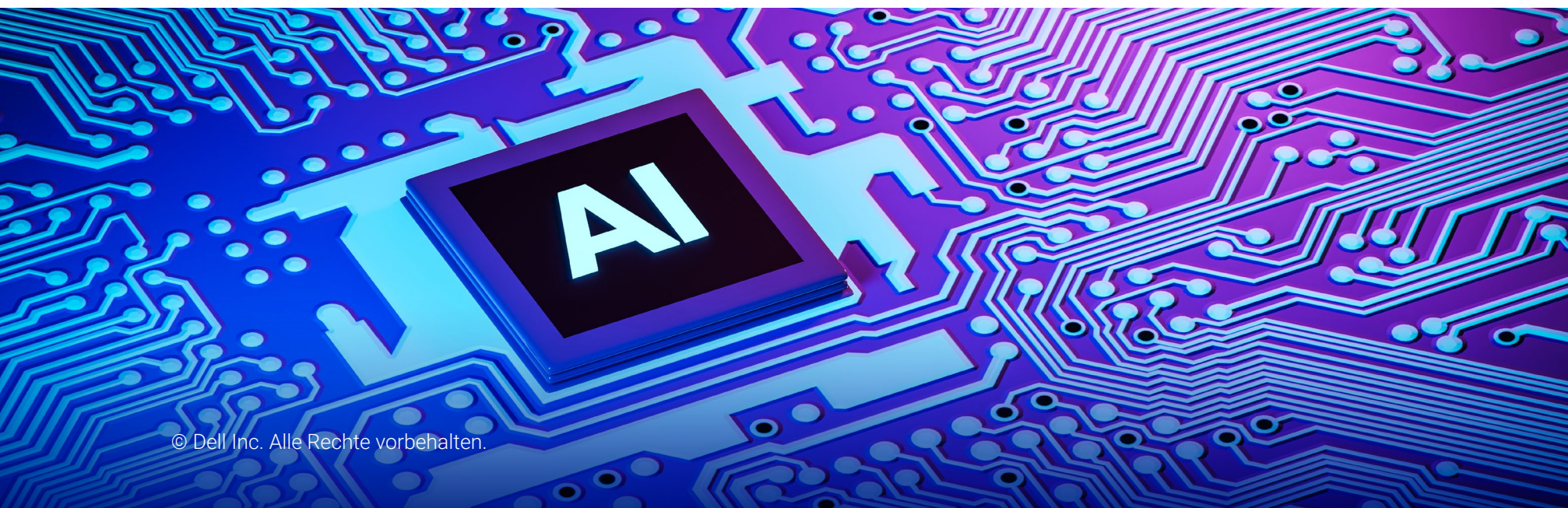
Voller Einsatz für das AI

Die Dell AI Factory with NVIDIA verkörpert das Engagement von Dell, KI zu nutzen und zu implementieren. Sie unterstützt Kunden bei der Beschleunigung ihrer KI-Initiativen, um ihre Daten zu maximieren und ihre Ziele zu erreichen.



Vereinfachte Datennavigation

Mit Dell und NVIDIA können Sie die Entwicklung des Datenmanagements mühelos durchlaufen. Und im Zusammenhang mit KI-gesteuerten Daten-Workloads ist der Weg genauso wichtig wie das Ziel.





Datenmanagement ist keine Aufgabe, die nur einmal durchgeführt werden muss.

Das Datenmanagement ist ein fortlaufender Prozess, bei dem eine Reihe von Praktiken und Tools zum Einsatz kommen, die Ihnen dabei helfen, kontinuierlich Werte aus Ihren Daten zu schöpfen. Und dies ist für jede erfolgreiche GenAI-Initiative von entscheidender Bedeutung.

Es gibt keine Pauschallösung.

Sie verfügen über eine einzigartige Datengrundlage – Ihr eigenes spezifisches Ökosystem aus Hardware und Software sowie Datenstandort und -typen und Ihre eigenen, individuellen Anwendungsfälle. Die in diesem E-Book beschriebenen Schritte dienen als übergreifender Leitfaden, der Sie dabei unterstützen soll, die spezifischen Anforderungen Ihres Unternehmens zu erfüllen.

Und Sie sind nicht auf sich allein gestellt.

[Dell Services](#) bietet umfassendes Fachwissen in jeder Phase des Lebenszyklus für eine kürzere Time-to-Value, und zwar abgestimmt auf Ihr Unternehmen. Von der Ausrichtung einer erfolgreichen Strategie und der Validierung von Daten bis hin zur schnellen Implementierung Ihrer Datenplattform und Gewährleistung sicherer, optimierter Abläufe stehen Ihnen unsere vertrauenswürdigen ExpertInnen zur Seite.

Ganz gleich, an welchem Punkt Ihrer Reise Sie sich befinden, wir helfen Ihnen dabei, Unternehmensdaten effektiv zu nutzen, um KI-Projekte voranzutreiben und die Time-to-Value zu beschleunigen.

[BESUCHEN SIE DELL.COM](#)



Glossar

Generative KI

Generative KI bezieht sich auf einen Teilbereich der künstlichen Intelligenz, der sich auf die Erstellung neuer Inhalte oder Daten konzentriert. Diese Art von KI verwendet Algorithmen und Modelle wie Generative Adversarial Networks (GANs) und Transformer, um Text, Bilder, Musik und andere Medientypen zu erzeugen. Generative KI ist einzigartig in ihrer Fähigkeit, originelle Ausgaben zu erzeugen, die den Stil und Inhalt ihrer Trainingsdaten nachahmen, was sie für kreative Aufgaben, die Erstellung von Inhalten und die Simulation komplexer Szenarien sehr wertvoll macht.

KI

Künstliche Intelligenz (KI) ist ein weit gefasstes Gebiet der Informatik, das sich auf die Entwicklung von Systemen konzentriert, die Aufgaben ausführen können, die normalerweise menschliche Intelligenz erfordern. Zu diesen Aufgaben gehören Lernen, logisches Denken, Problemlösung, Wahrnehmung, Sprachverständnis und Entscheidungsfindung. KI umfasst verschiedene Techniken und Ansätze, wie maschinelles Lernen, Verarbeitung natürlicher Sprache und Robotik. Ziel ist es, intelligente Agents zu entwickeln, die sich an neue Situationen anpassen, ihre Leistung im Laufe der Zeit verbessern und menschliche Aktivitäten über eine Vielzahl von Anwendungen hinweg unterstützen oder automatisieren können.

KI-Workloads

KI-Workloads beziehen sich auf die spezifischen Aufgaben oder Prozesse, die von KI-Systemen verarbeitet werden. Diese Workloads können je nach Anwendung sehr unterschiedlich sein und Datenverarbeitung, Modelltraining, Inferenz und Analyse umfassen. KI-Workloads zeichnen sich durch hohe Rechenanforderungen aus, die für das Training von Modellen, die Verarbeitung großer Datenvolumen und die Entscheidungsfindung in Echtzeit erforderlich sind. Das Management von KI-Workloads umfasst die Optimierung von Hardware, Software und Algorithmen, um diese anspruchsvollen Aufgaben effizient zu bewältigen und sicherzustellen, dass KI-Systeme effektiv und in großem Umfang arbeiten.

Referenzen und weiterführende Literatur

[Infografik](#): Die Entwicklung des Datenmanagements

[Dell Professional Services for Generative AI](#)

[Datenverwaltungslösungen von Dell](#)

[Dell AI Factory with NVIDIA](#)

[Die 5 wichtigsten Sicherheitsüberlegungen für GenAI](#)