

Generative KI ist hier: Sind Sie bereit dafür?

Best Practices für den Einsatz von
KI für Unternehmensdaten

Inhaltsverzeichnis

Schaffung der Grundlage für Erfolg	1
Die entscheidende Rolle Ihrer Daten für generative KI	2
Erfolgreiche generative KI durch betriebliche Bereitschaft	3
Überlegungen zur betrieblichen Bereitschaft	4
Die Nutzung der richtigen Daten erfordert Zusammenarbeit	5
Halten Sie generative KI-Modelle nah an Ihren Daten	6
Generative KI und Datenmanagement	7
Die richtige Dimensionierung Ihrer generativen KI-Modelle	8
KI-Bereitstellungsmodelle: Bewerten des Kosten-Nutzen-Verhältnisses	9
Die richtige Infrastruktur für Ihre Modelle	10
KI für Ihre Teams	11
Die letzte Meile	12
Neuanfänge	13
Erste Schritte mit generativer KI	14
Dell Technologies kann Ihren Weg beschleunigen	15



Schaffung der Grundlage für Erfolg

Generative KI ist in aller Munde

Auch wenn sie noch in den Kinderschuhen steckt, wird generative KI bereits in großem Umfang von Unternehmen und Privatpersonen genutzt. In einer Umfrage von Dell gaben 91 % der Unternehmens- und IT-Verantwortlichen an, dass sie generative KI in irgendeiner Form bereits in ihrem Alltag einsetzen, während 71 % angaben, dass sie sie auch schon für die Arbeit verwendet haben.¹

KI für den Privatgebrauch (wie z. B. ChatGPT und DALL-E) hat sich bereits auf vielfältige Weise auf die Gesellschaft ausgewirkt, da sie die menschliche Interaktion mit Technologie revolutioniert und menschenähnliche Ergebnisse bei Gesprächstexten, Bildern, Audio, Video und Softwarecode erzielt.

Der Ausgangspunkt

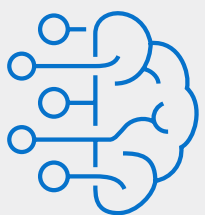
Wird die generative KI verantwortungsvoll eingesetzt?		Wie werden alle ins Boot geholt?
Wie lassen sich die eigenen Daten am besten nutzen?	Wie lassen sich Risiken am besten minimieren?	

Risiken und Chancen

Generative KI wird sich gleichermaßen auf Unternehmen auswirken und sie auf eine Weise transformieren, von der selbst Visionäre derzeit nur eine vage Vorstellung haben. Im Zuge einer möglichst schnellen Einführung kann es für Führungskräfte jedoch verlockend sein, zu schnell zu agieren und dabei wichtige Implikationen im Zusammenhang mit Daten, Compliance, Governance und anderen Risiken zu übergehen.

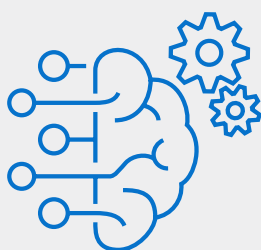
Unternehmen müssen auf ihrem Weg zur generativen KI andere Aspekte als Privatpersonen berücksichtigen. Sie können das Vertrauen ihrer KundInnen und den hohen Wert ihrer Daten nicht riskieren, nur um als erstes über die Ziellinie zu laufen.

Generative KI ist nicht perfekt, aber für Ihr Unternehmen *muss* sie es sein. In diesem E-Book erfahren Sie, wie Sie Ihre Initiativen im Bereich generative KI sowohl mit Vorsicht als auch mit Zuversicht umsetzen können.



KI

Künstliche Intelligenz (KI) wendet fortschrittliche Analysen und logikbasierte Techniken an, um Ereignisse zu interpretieren und die Entscheidungsfindung sowie Aktionen zu unterstützen und zu automatisieren.²



Generative KI

Generative KI bezeichnet eine Reihe von Technologien und Techniken, die enorme Datenmengen nutzen, einschließlich großer Sprachmodelle, um neue Inhalte zu erzeugen (z. B. Text, Video, Bilder, Audio und Code). Sie verwendet Eingaben wie z. B. Prompts in natürlicher Sprache oder andere nicht auf Code basierende und nicht herkömmliche Eingaben.³

1. Interne Untersuchung von Dell
2. [IT-Glossar von Gartner: Künstliche Intelligenz](#)
3. [Forrester: Ignoring Generative AI Will Be A Costly Mistake for Enterprises](#), Forrester, 14. Februar 2023.

Die entscheidende Rolle Ihrer Daten für generative KI

Daten als Alleinstellungsmerkmal

Obwohl generative KI bereits ihr erhebliches Potenzial für die Transformation von Unternehmen und ganzen Branchen gezeigt hat, fangen Unternehmens- und IT-Verantwortliche gerade erst damit an, die Vielzahl der Möglichkeiten für ihre Unternehmen in Betracht zu ziehen.

Eines ist jedoch sicher: Hochwertige Daten sind das entscheidende Element für Ihre generativen KI-Projekte. Je mehr Sie Ihre eigenen Daten für Ihre Projekte verwenden, desto höher ist der Nutzen für Ihr Unternehmen. Das Potenzial Ihrer eigenen Daten ist quasi endlos.

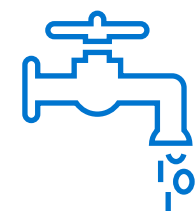
Am Anfang Ihrer generativen KI-Initiative ist es wichtig, eine umfassende Strategie zu entwickeln, um die Einzigartigkeit Ihrer Daten voll auszuschöpfen und für spezialisierte Projekte zu nutzen, die für Ihr Unternehmen einen Mehrwert schaffen. Die Auswahl der richtigen Projekte erfordert eine sorgfältige Bewertung der Anforderungen, Risiken, Eingaben und Ergebnisse für jeden Anwendungsfall. Je ausgereifter Ihre generativen KI-Funktionen mit der Zeit werden, desto mehr Optionen stehen Ihnen für die Feinabstimmung Ihrer Modelle zur Verfügung, um damit im gesamten Unternehmen noch mehr Innovationen zu ermöglichen. Wenn Sie sich möglichst frühzeitig den Wert Ihrer Daten bewusst machen, können Sie Ihr Unternehmen besser auf die Zukunft der Arbeit vorbereiten.

Datenrisiken bei generativer KI

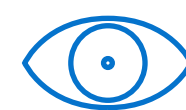
Daten und Risiken gehen Hand in Hand. Daten sind zwar die treibende Kraft für Ihre generativen KI-Projekte, Sie müssen sich jedoch auch über die potenziellen Risiken bewusst sein, die das Hosten generativer KI-Modelle in Public Clouds mit sich bringt. Dazu gehören:



**Verlust von
geistigem Eigentum**



Datenleck



**Datenschutz-
probleme**



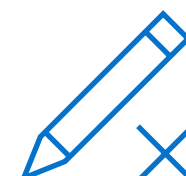
**Compliance-
Verstöße**



**Glaubwürdigkeits- und
Integritätsverlust**



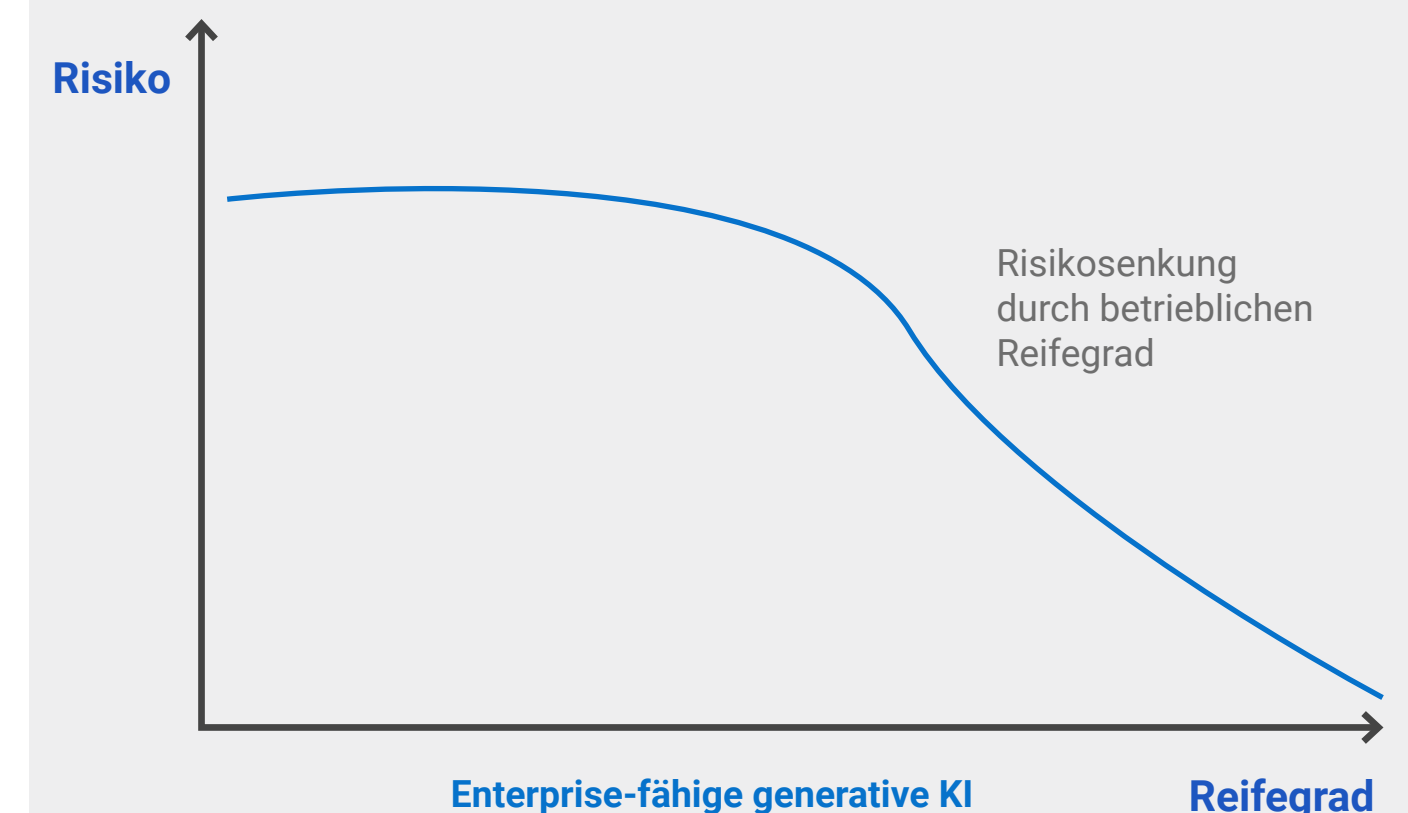
Einseitigkeit



**Verletzung von
geistigem Eigentum**

Risikomanagement und Nutzensteigerung

Am Anfang Ihres Wegs müssen Sie in Technologien und Schulungen investieren, um Risiken zu reduzieren und Ihrem Unternehmen mehr Kontrolle sowie eine höhere Wertschöpfung zu ermöglichen.



Erfolgreiche generative KI durch betriebliche Bereitschaft

Jedes Unternehmen hat individuelle Anforderungen an generative KI

Generative KI-Technologien können Unternehmen auf unzählige Arten von Nutzen sein. Auch wenn bereits Hunderte von Anwendungsfällen in allen möglichen vertikalen Märkten zutage gefördert wurden, können Sie mit den passenden Anwendungsfällen für Ihr Unternehmen, basierend auf Ihren eigenen Daten und Ihren individuellen Anforderungen, schnellere Erfolge erzielen.

Der Anfang



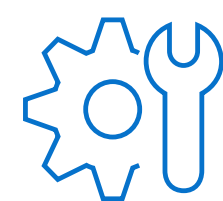
MitarbeiterInnen und Teams

Bereiten Sie Ihr Unternehmen auf den Einsatz und die Chancen generativer KI vor. IT-Abteilungen blicken hierfür nach innen, während geschäftliche Entscheidungsträger nach außen schauen.



Prozesse und Richtlinien

Definieren und kommunizieren Sie, wie Ihr Unternehmen KI nutzen soll. Es ist von entscheidender Bedeutung, dass MitarbeiterInnen sich mit KI als Teil der geschäftlichen Tätigkeit beschäftigen.



Technologie

Stellen Sie sicheren Zugriff auf generative KI in Ihrem gesamten Unternehmen bereit. Führen Sie mit allen Stakeholdern des Unternehmens offene Gespräche darüber, wie die IT diese Funktionen bereitstellen wird.

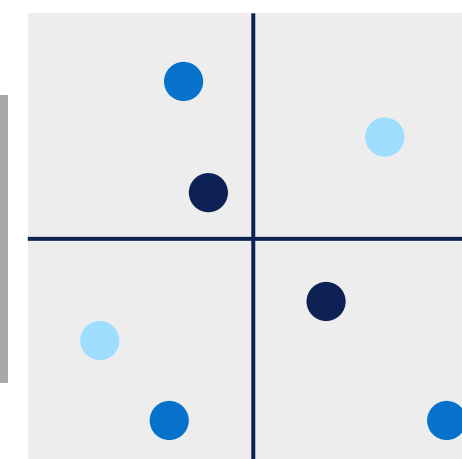
Der Status Quo und wo die Reise hingehen soll

Die jeweiligen Anwendungsfälle für generative KI bilden die Grundlage für die Planung und Bereitstellung der technischen Infrastruktur. Die folgenden Punkte helfen Ihnen dabei, Ihr Investitionsniveau zu bestimmen:

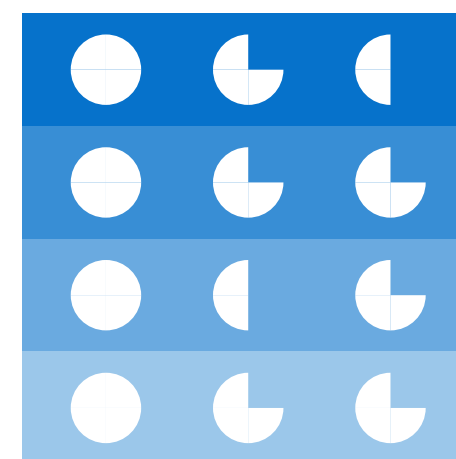
Ermitteln Sie den aktuellen Status Quo Ihrer Umgebung, um Topologiediagramme zu entwickeln und Informationen zu wichtigen technischen Systemen zu dokumentieren. Mit diesen Erkenntnissen können Sie Ihre strategische Vision und die Leitprinzipien für Ihre zukünftigen generativen KI-Projekte festlegen.

Setzen Sie sich mit den geplanten Investitionen, langfristigen Strategieentscheidungen sowie den betrieblichen und finanziellen Informationen für die Zielumgebung auseinander. Erstellen Sie einen Blueprint Ihres zukünftigen Wunschzustands und holen Sie Feedback von allen Stakeholdern ein, bevor Sie Ihre Vision final festlegen.

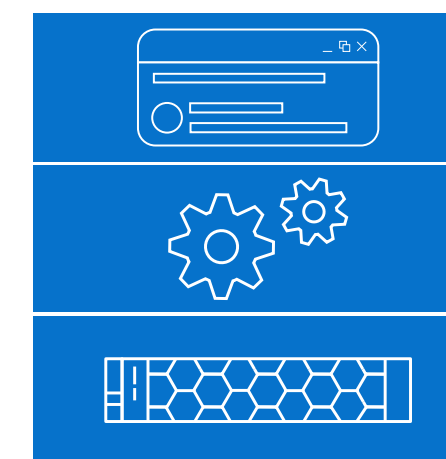
Priorisierte Anwendungsfälle



Zielfunktionen



Zukünftige IT-Architektur



Dank der explosiven Zunahme KI-gesteuerter Automatisierung müssen SpezialistInnen nun weniger Zeit mit der Analyse von Bildern und der Quantifizierung von Daten für einfachere, sich wiederholende Aufgaben verbringen. Dadurch haben sie nicht nur mehr Zeit für komplexe Probleme, sie können auch schneller genauere Diagnosen liefern.⁴

—Nikolas Stathonikos

Principal Investigator, AI
Development and Implementation,
UMC Utrecht

Überlegungen zur betrieblichen Bereitschaft

Wenn Sie für Ihre Branche eine hochspezialisierte Fähigkeit aufbauen wollen, müssen Sie wahrscheinlich Ihr eigenes Modell trainieren. Die technischen Anforderungen dafür sind jedoch hoch. Im Gegensatz dazu bieten private Instanzen von vorab trainierten großen Sprachmodellen (Large Language Models, LLMs) mit generativer KI – wie LLaMA 2 oder Falcon – Vorteile, was die Geschwindigkeit und Bereitstellung betrifft, erfordern jedoch möglicherweise eine Feinabstimmung, wenn nicht alle Ihre Anforderungen erfüllt werden. RAG (Retrieval-Augmented

Generation) ermöglicht Ihnen die Einbeziehung von unternehmensspezifischem Wissen und ist relativ einfach bereitzustellen und zu verwenden. Unternehmen können auch SaaS mit generativen KI-Funktionen erwerben.
Unabhängig davon, welchen Weg Sie einschlagen, müssen Sie bei der Entscheidung, ob Sie Ihre generative KI-Lösung selbst erstellen oder kaufen wollen, die gewünschten Geschäftsergebnisse, die Eignung Ihrer Unternehmensdaten und Ihre potenziellen Anwendungsfälle berücksichtigen.

 Herstellung  Kaufen		
Branchenspezifisch	Funktion	Produktivität
▪ Medikamentenentwicklung ▪ Materialdesign ▪ Optimierte Ernteerträge	▪ Marketing ▪ Vertrieb ▪ Beschaffung	▪ Salesforce ▪ Microsoft 365 ▪ Box

Daten-Governance
Generative KI-Projekte erfordern die strenge Einhaltung von Daten- und Datenschutzbestimmungen und unterliegen Einschränkungen hinsichtlich der Datenhoheit. Auch wenn es äußerst nützlich ist, Ihre eigenen Daten für generative KI-Projekte zu verwenden, muss der Umgang mit den Daten und die Verarbeitung sehr sorgfältig erfolgen.

Vorbereiten Ihrer Daten für KI
Ihre Unternehmensdaten lassen sich am besten für generative KI verwenden, wenn sie bereinigt und konsistent sind und zentral gespeichert werden. Achten Sie darauf, Ihre Daten zu überwachen und regelmäßig zu erweitern, um Zuverlässigkeit, Konsistenz und Integrität zu gewährleisten.

Proof-of-Concept und Minimum Viable Product
Demonstrieren Sie die Machbarkeit generativer KI-Initiativen in Ihrem Unternehmen mit Proof-of-Concept-Projekten. Mit kleinen, leicht umzusetzenden Projekten lässt sich die Aufmerksamkeit von MitarbeiterInnen und Führungskräften gewinnen und die Wahrnehmung generativer KI und ihres Nutzens im gesamten Unternehmen steigern.

KI als langfristige Investition
Auch wenn es am besten ist, klein anzufangen, sollten Sie sich darüber im Klaren sein, dass KI bleiben und zu einem wichtigen Rädchen im Unternehmensgetriebe werden wird. Wir befinden uns noch ganz am Anfang der KI-Revolution, aber Unternehmensbudgets sollten erhebliche zukünftige Investitionen einplanen.

Die Nutzung der richtigen Daten erfordert Zusammenarbeit

Zuordnen von Anwendungsfällen zu Ihren Daten

Die besten Anwendungsfälle für generative KI nutzen Ihre umfangreichen Daten und beinhalten zur Unterstützung die richtige Mischung aus Kompetenzen, Budgets und Ressourcen. Es ist sehr wichtig, dass die Geschäfts- und IT-Teams zusammenarbeiten, um die Prioritäten festzulegen. Ein Punkt ist jedoch sicher: Geschäftliche Stakeholder werden wahrscheinlich schneller agieren wollen als die IT. Das bedeutet, dass Sie die richtige Balance zwischen Geschwindigkeit und Risiko finden müssen.

Bei der Einschätzung Ihrer internen Bereitschaft müssen Sie sich darüber bewusst sein, dass Ihre Daten sich am besten für generative KI eignen, wenn sie bereinigt und konsistent sind und zentral gespeichert werden. Achten Sie darauf, Ihre Daten zu überwachen und regelmäßig zu erweitern, um Zuverlässigkeit, Konsistenz und Integrität zu gewährleisten. Darüber hinaus erfordern generative KI-Projekte die strenge Einhaltung von Daten- und Datenschutzbestimmungen und unterliegen Einschränkungen hinsichtlich der Datenhoheit.



Viele Stakeholder interessieren sich für generative KI

Die Zusammenarbeit zwischen den Geschäfts- und IT-Teams ist entscheidend, um Prioritäten für Anwendungsfälle festzulegen und einen Konsens bezüglich der Bereitschaft zu erzielen.

Wichtige Rollen und Verantwortlichkeiten:

Geschäftsführer und Führungskräfte: Sie identifizieren die Geschäftschancen und -prioritäten, für die generative KI eingesetzt werden soll.

IT-Verantwortliche: Sie stellen sicher, dass die richtige technische Infrastruktur mit der nötigen Zuverlässigkeit, Skalierbarkeit und Sicherheit für aktuelle und zukünftige Projekte vorhanden ist. Sie werden bei Datenentscheidungen miteinbezogen.

Data Scientists: Sie entwickeln und stellen die grundlegenden generativen KI-Modelle bereit.

FachexpertInnen: Sie liefern Fachwissen und Daten aus einem spezifischen Bereich für das Training und die Validierung generativer KI-Modelle.

Halten Sie generative KI-Modelle nah an Ihren Daten

Warum Nähe zu den Daten wichtig ist

Um bei der Workload-Platzierung die besten Ergebnisse zu erzielen, sollten Sie sie in der Umgebung einsetzen, die für Ihre geschäftlichen und technischen Anforderungen am sinnvollsten ist. Denken Sie jedoch daran, dass bei generativer KI viele Daten, für die es bisher keine starken Anwendungsfälle gibt, in Zukunft interessant werden können und Sie möglicherweise dazu animieren, die Platzierung zu überdenken.

Nachfolgend finden Sie einige Merkmale für die Ermittlung der am besten geeigneten Umgebungen. Ihre eigenen Datenimplementierungen und -präferenzen mögen sich vielleicht unterscheiden, die Konzepte, die auch im Diagramm enthalten sind, sollten Ihnen jedoch beim Einstieg helfen.

Überlegungen zur Platzierung generativer KI

Sicherer Zugriff: Der einfachste Ansatz bedeutet oft, die vollständige Kontrolle über Ihre eigene Infrastruktur, Daten und Services zu behalten.

Allgemeine Anwendungsfälle: Die Public Cloud bietet Vorteile für allgemeine Anwendungsfälle, für spezifische Anwendungsfälle bieten private Umgebungen jedoch bessere Anpassungsmöglichkeiten.

Genauigkeit und Anpassung: Die Private Cloud ermöglicht mehr Anpassung für den Zugriff auf aktuelle Daten und die Erstellung von Leitplanken für eine höhere Genauigkeit und bessere Modellperformance.

Kürzere Time-to-Value: Die Public Cloud ermöglicht einen schnellen Zugang zu generativer KI für Test- und Experimentierzwecke. Die Private Cloud bietet zahlreiche Optionen, um Modelle mit mehr Kontrolle an Ihre Anwendungsfälle anzupassen.

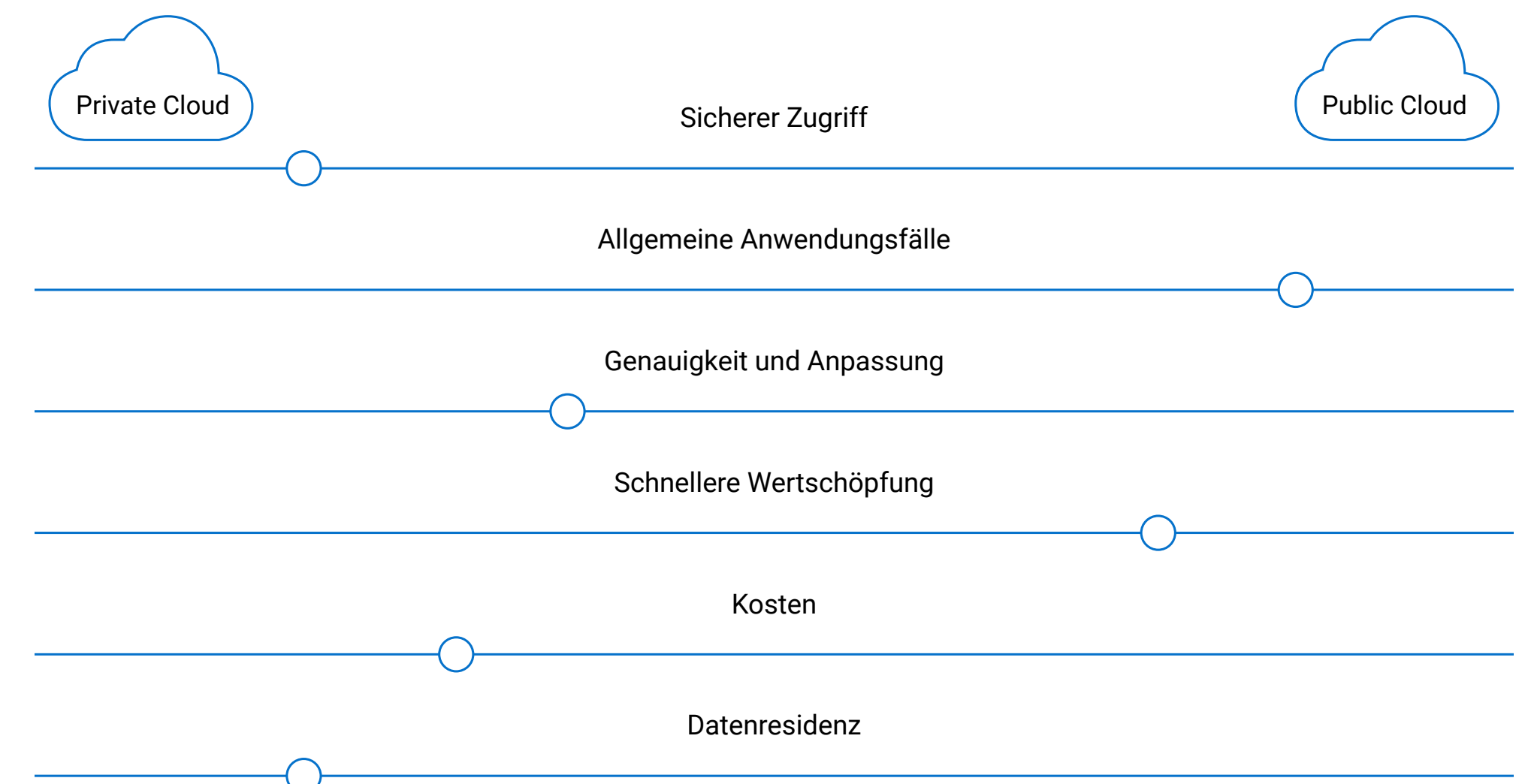
Kosten: Die Private Cloud bietet mehr Transparenz und die Möglichkeit, Modelle richtig zu dimensionieren, sodass Sie mehr Kontrolle über die Kosten haben.

Datenresidenz: KI funktioniert am besten, wenn sie mit Ihren Daten kombiniert wird, wo auch immer sie sich befinden: am Edge, in der Cloud oder vor Ort.

5 Gründe für den Einsatz von KI für Ihre Daten

Die Wichtigkeit von Daten kann im KI-Zeitalter nicht oft genug betont werden und die Ausschöpfung des vollen Potenzials der generativen KI hängt von Ihrer Fähigkeit ab, sie zusammen mit Ihren eigenen proprietären Daten zu nutzen. Hier sind fünf Gründe für den Einsatz von KI:⁵

1. Sie haben die Kontrolle über die Sicherheit und den Zugriff auf Ihre Daten.
2. Sie können zusätzliche Leitplanken schaffen und das Risiko für Imageschäden reduzieren.
3. Sie profitieren von Ihren eigenen Echtzeitdaten.
4. Sie erzielen Kosteneinsparungen.
5. Sie steigern die Energieeffizienz.



Generative KI und Datenmanagement

Was ist Ihre Datenmanagementstrategie?

Die meisten Unternehmen verfolgen bei ihrer generativen KI-Strategie einen zweistufigen Ansatz. Sie experimentieren mit taktischen Bereitstellungen, um zu lernen und nicht abgehängt zu werden, und entwickeln gleichzeitig eine langfristige Strategie, die flexibel genug für die unzähligen Anwendungsfälle ist, die sich im Laufe der Zeit herauskristallisieren werden. Dieser Ansatz erfordert daher auch eine zweistufige Datenmanagementstrategie.

Langfristig: Unternehmen müssen über eine angemessene Datenmanagementstrategie für generative KI verfügen, die wichtige Aspekte wie Data Governance, Datenqualität, Datenintegrität und Datensicherheit berücksichtigt. Dies ist nicht nur eine technologische, sondern auch eine organisatorische Angelegenheit, da viele dieser Aspekte eine enge Zusammenarbeit mit den Stakeholdern des Unternehmens erfordern.

Unternehmen benötigen außerdem ein Repository für ihre generativen KI-Daten, wie z. B. einen Data Lake oder ein Data Lakehouse, das die notwendigen Kontrollmöglichkeiten sowie die Flexibilität bietet, Datenvolumen für generative KI-Anwendungsfälle und die zugehörigen Modelle nach Bedarf zu kombinieren.

Es ist wichtig, diese Datenmanagementfunktionen stufenweise aufzubauen, um eine gleichzeitige Entwicklung von Datenmanagement- und generativen KI-Fähigkeiten zu ermöglichen und schrittweise die Funktionen des grundlegenden Daten-Repositorys für generative KI zu erweitern.

Dell Technologies Services für die Datenvorbereitung

Kurzfristig: Unternehmen sind äußerst bestrebt, generative KI-Anwendungsfälle ein- und auszuführen, um sobald wie möglich vom Nutzen, den generative KI-Modelle für ihre Daten bieten, zu profitieren. Für diesen Fast-Track-Ansatz ist eine Datenvorbereitung erforderlich, die Folgendes umfasst:

- Anonymisieren von Daten zum Schutz vor der Offenlegung personenbezogener Daten
- Kennzeichnen der Daten
- Normalisieren der Daten über Datenquellen hinweg

Darüber hinaus müssen auch Datenpipelines erstellt werden, um die Daten in ein Modell zu integrieren. Dell bietet dedizierte Data Preparation Services für generative KI, um Unternehmen durch alle entscheidenden Schritte zu leiten und die Time-to-Value für Anwendungsfälle mit hoher Priorität zu beschleunigen.



Der Weg zur Maximierung des Nutzens dieser LLMs liegt darin, Ihre eigenen Daten basierend auf dem jeweiligen Anwendungsfall zu kombinieren und mit diesen Daten zu trainieren. All dies fängt mit einer starken Datenmanagementstrategie an, um die Qualität Ihrer Daten zu gewährleisten, und geht dann in die Planung des komplexen Prozesses der Anpassung von Modellen über, da für die Anwendungsfälle der meisten Unternehmen bereichsspezifische Informationen aus ihrem jeweiligen vertikalen Segment in ein bereits vorhandenes großes Sprachmodell einbezogen werden müssen.

—Jonathan Seckler,
Dell Technologies

Die richtige Dimensionierung Ihrer generativen KI-Modelle

Generative KI-Lösungen in unterschiedlichen Größen

Die verschiedenen Arten generativer KI erfordern entsprechende Einstiegspunkte und Investitionen, um den Erfolg zu gewährleisten.

LLMs, die mit riesigen Textmengen trainiert wurden, sind wie Enzyklopädien. Sie eignen sich hervorragend für allgemeine Anwendungsfälle, aber nicht für spezifische Fragen bezüglich Ihrer Unternehmensdaten.

Ihre Daten sind extrem aussagekräftig und einzigartig. Je mehr und besser Sie sie nutzen, desto mehr können Sie sich auf bereichs- und unternehmensspezifische Anwendungsfälle mit einem gezielten Branchennutzen für bestimmte Aufgaben oder Funktionen konzentrieren, für die nur Sie allein die Daten haben.

Ihre Daten verbessern die Effizienz und den Nutzen generativer KI



Richten Sie Ihr Modell für eine höhere Effizienz an Ihrem Anwendungsfall aus

Die genaue Bestimmung Ihrer Anwendungsfälle erleichtert die richtige Dimensionierung Ihres Modells und vermeidet die Verschwendung von Ressourcen für ein Modell, das größer ist, als Sie benötigen. Kleinere Modelle sind kostenbewusster und energieeffizienter. Die richtige Dimensionierung Ihres LLM für die Verwendung auf einer Workstation, einem Laptop oder am Edge ist ebenfalls hilfreich.

Laut Forrester wird generative KI den Energieverbrauch in Zukunft auf verschiedene Art und Weise weiter reduzieren⁴:

Messung und Reporting: KI kann Automatisierung ermöglichen, um zeitnahe Empfehlungen für Wege zu Netto-Null-Emissionen bereitzustellen.

Governance und Visualisierung: Generative KI ist ein idealer Anwendungsfall für die Erstellung einer einzigen Informationsquelle, die weniger Ressourcen für das Management erfordert.

Umfrage: KI kann dazu beitragen, revolutionäre Ansätze zur Entwicklung weniger energieintensiver Technologien und Prozesse zu beschleunigen.

KI-Bereitstellungsmodelle: Bewerten des Kosten-Nutzen-Verhältnisses

Die ersten drei nachfolgend genannten Bereitstellungsmodellarten werden aktuell von 90 % der Unternehmen implementiert. Für welches KI-Modell Sie sich entscheiden, hängt vom Grad der Data-Science-Bereitschaft Ihres Unternehmens, den Bereitstellungsmustern und den damit verbundenen Implikationen ab. Die Modellaugmentierung ist ein guter Ausgangspunkt und die meisten Unternehmen gehen von dieser Phase in die Feinabstimmung Ihrer Modelle über.

Einfache Inferenzierung mit vorab trainierten Modellen

Beim sogenannten „Prompt Engineering“ wird einem vorab trainiertem Modell eine Frage gestellt und ein Ergebnis erhalten. Die Verwendung von ChatGPT ist ein Beispiel für einfache Inferenzierung.

Modellaugmentierung

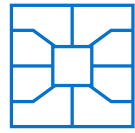
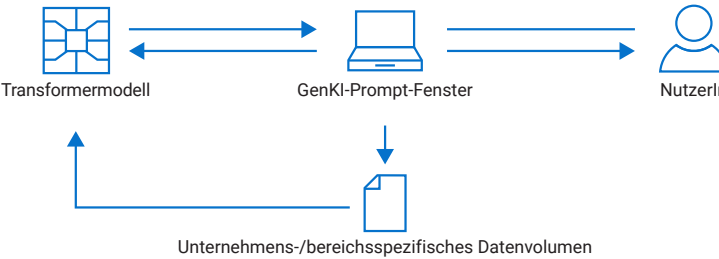
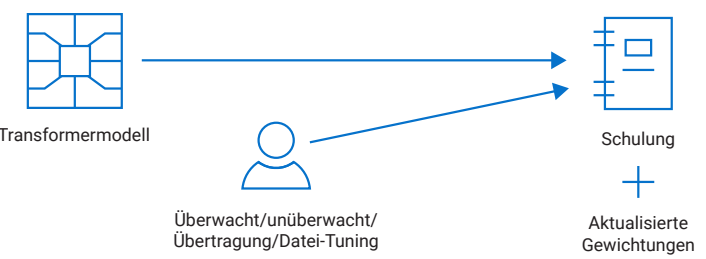
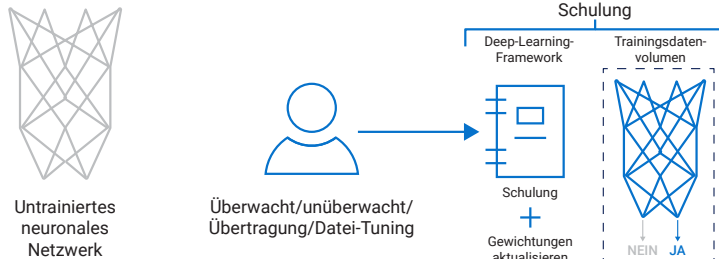
Hierbei handelt es sich um Inferenzierung in Kombination mit dem Zugriff auf Ihre Daten. RAG ist hier ein gängiger Anwendungsfall. Es bietet eine einfache Möglichkeit, ein generatives KI-Modell mit Ihren Daten intelligenter zu gestalten.

Feinabstimmung von Modellen

Dies beinhaltet die Änderung der Gewichtung eines Modells und die Zuführung Ihrer Daten. Bei der Skalierung von Modellen bietet die Feinabstimmung Unternehmen bessere Ergebnisse, erfordert jedoch auch mehr Aufwand.

Modelltraining

Für das Modelltraining wird ein äußerst spezifisches Modell entwickelt und mit einem Datenvolumen trainiert. Dies erfordert in der Regel die meiste Arbeit und die meisten Ressourcen und wird häufig nur für die Lösung sehr komplexer Probleme verwendet.

	Vorab trainiertes Modell	Modellaugmentierung (RAG, PAL)	Feinabstimmung von Modellen	Modelltraining
Modelldefinition	Anfänglich oder vorab trainiertes Modell  Transformermodell			
Aufwand				
Kosten				
Wert und Differenzierung				
Datenintegration				
Infrastruktur	Client – Server	Client – Server	GPU-optimiert	Große GPU-Bereitstellung
Kompetenzen	IT Ops	Entwickler	Data Scientist(s)	Data Scientist(s)



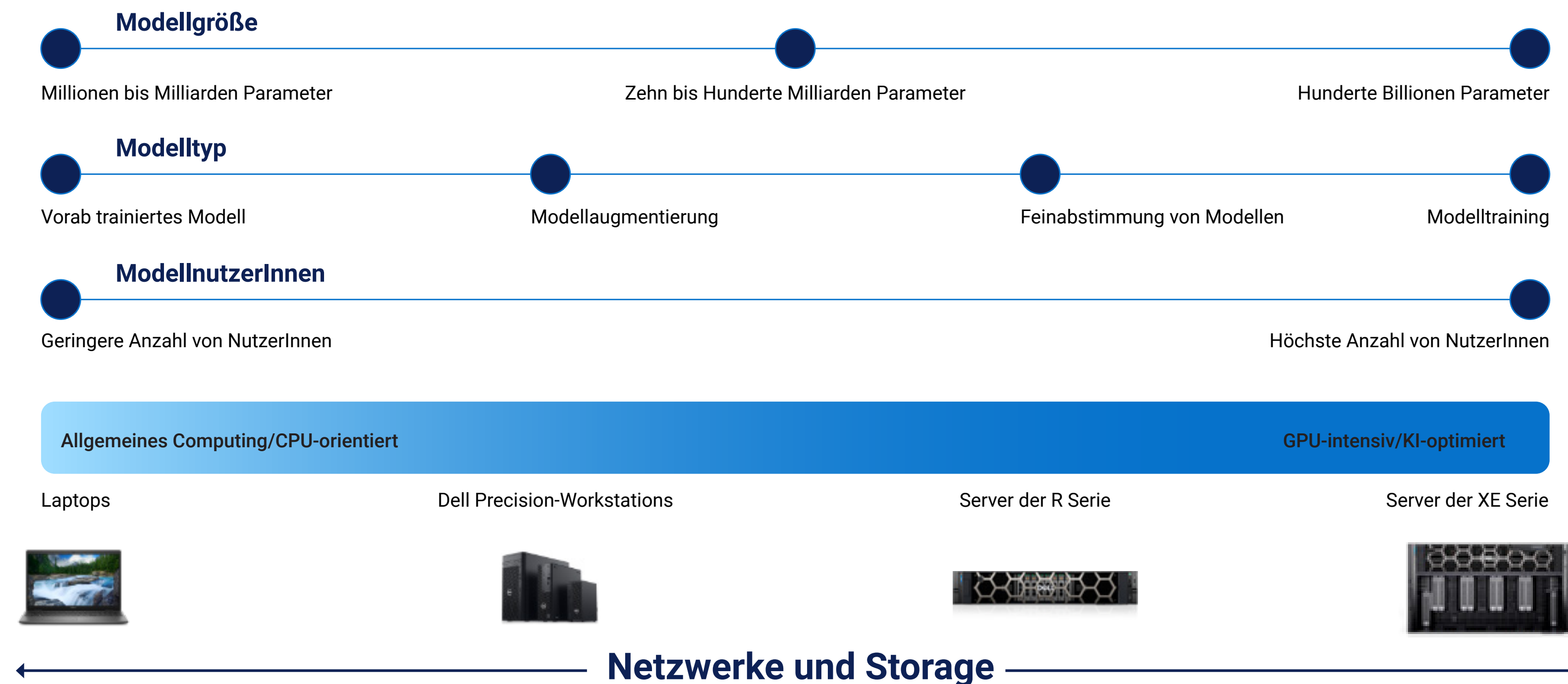
Die richtige Infrastruktur für Ihre Modelle

Die Hardwareanforderungen steigen zusammen mit der Rechenlast

Einfache Inferenzierungs-Anwendungsfälle können mit minimalem Zeitaufwand und vorhandener Hardware umgesetzt werden. Für die Modellaugmentierung wird bereits eine rechenintensivere Infrastruktur benötigt, während die Feinabstimmung von Modellen und das Modelltraining eine stetig steigende Rechenleistung für die Ausführung von Workloads erfordert. Für Ihre rechenintensivsten

Anwendungsfälle müssen Sie wahrscheinlich einen Data Scientist und hochleistungsfähige KI-optimierte Server einsetzen.

Wenn Ihre generativen KI-Funktionen und -Anwendungsfälle ausgereifter werden, müssen Sie Ihre Infrastruktur möglicherweise richtig dimensionieren, um bessere Ergebnisse zu erzielen.



KI für Ihre Teams

Generative KI als überzeugendes Nutzererlebnis

Aus Kunden- und Geschäftsperspektive ist der wichtigste Schritt bei generativer KI die sogenannte „letzte Meile“ des gesamten Prozesses. Sie haben Ihre generativen KI-Instanzen bereits erstellt und Ihre Modelle richtig dimensioniert. Jetzt müssen Sie die generative KI Ihren MitarbeiterInnen zur Verfügung stellen, also höchstwahrscheinlich sowohl technischen und als auch nicht technischen NutzerInnen. Es ist also an der Zeit, Ihre MitarbeiterInnen dabei zu unterstützen, die Chancen zu verstehen, die generative KI bietet.

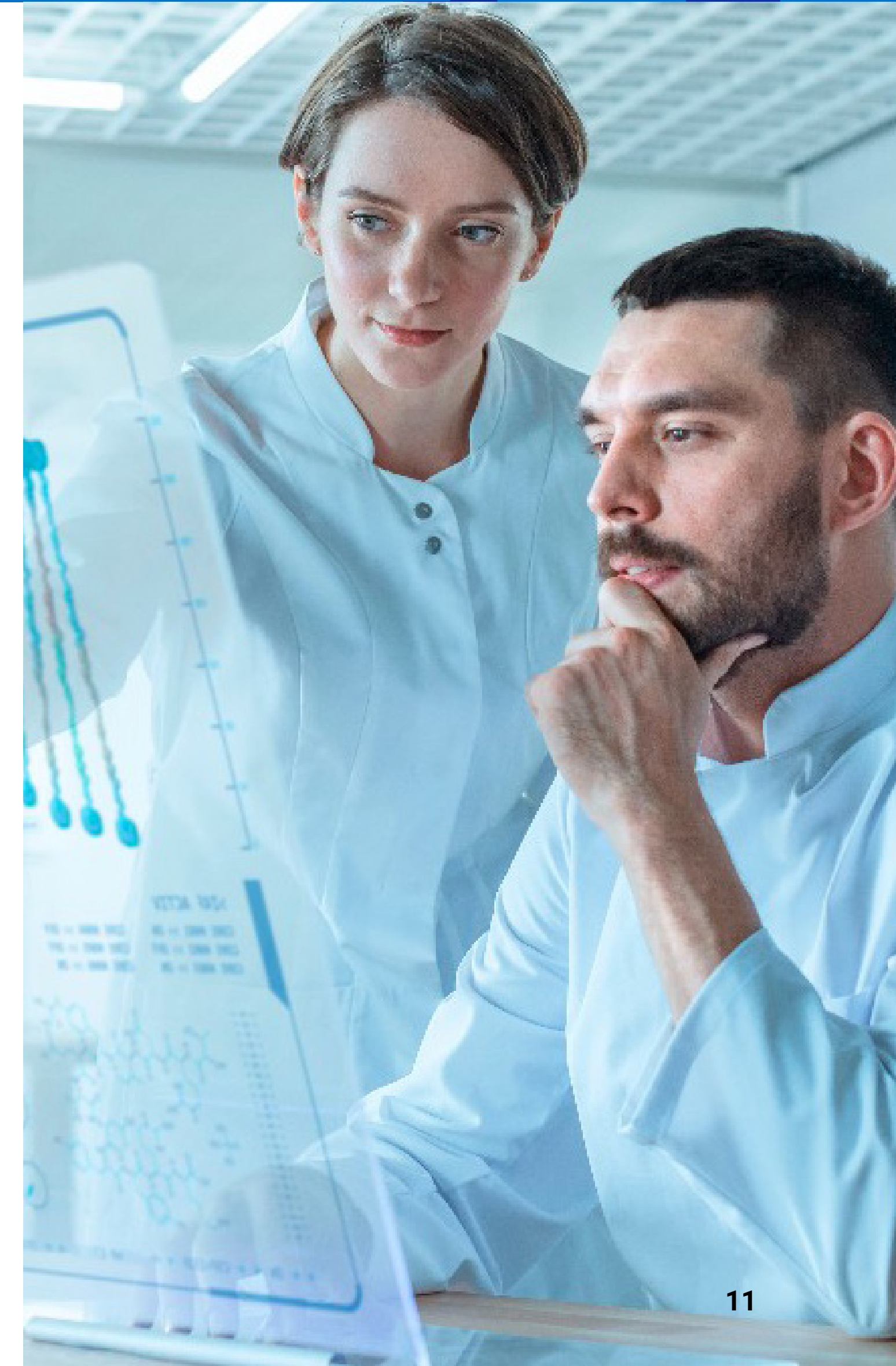
Aufbau eines Frameworks, damit Teams Anwendungsfälle identifizieren können

Unterstützen Sie Ihre Teams mit Tools und Ressourcen wie z. B. Vorlagen, die genutzt werden können, um kreative und intelligente Lösungen für die jeweiligen Aufgabenbereiche zu erstellen, und anschließend unternehmensweit geteilt werden. Dies trägt dazu bei, generative KI zu einem willkommenen und transformativen Prozess in Ihrem Unternehmen zu machen.

Prompt Engineering und Prompt-Bibliotheken

Prompt Engineering kann für manche Ihrer MitarbeiterInnen möglicherweise zu einem faszinierenden neuen Karriereweg werden. Die Auswahl der richtigen Begriffe, Formulierungen oder Symbole, anhand derer KI-Modelle relevante und qualitativ hochwertige Texte erzeugen, ist eine Schlüsselkompetenz für die Entwicklung wirkungsvoller KI-basierter Anwendungen.

Das Schulen Ihre MitarbeiterInnen in Prompt-Engineering-Verfahren kann die Produktivität im gesamten Unternehmen steigern.



Die letzte Meile

Unternehmensweite Fokussierung auf generative KI

Entwickeln und führen Sie Schulungen durch, um die Begeisterung für die Möglichkeiten generativer KI im gesamten Unternehmen zu entfachen, und ermitteln Sie die KI-Enthusiasten, die diese Schulungen durchführen können.

Zentralisierung von Best Practices

Erstellen Sie eine einzige Informationsquelle für Fachwissen, Anwendungsfälle und Best Practices rund um generative KI, damit Ihre Teams die Governance- und Sicherheitsmaßnahmen besser verstehen und Fehler vermeiden können.

Förderung in allen Unternehmensbereichen

Bieten Sie Schulungen zu generativer KI und Best Practices für alle Unternehmensbereiche und -abteilungen an, um MitarbeiterInnen mit der Technologie vertraut zu machen und Vertrauen in diese zu schaffen. Stellen Sie Teams basierend auf Kompetenzen, Rollen und technischen Fähigkeiten zusammen. Stellen Sie sicher, dass Sie in jeder Gruppe einen „Kapitän“ haben, der die Entwicklung leitet.

Unterstützen Sie NutzerInnen dabei, das Optimum aus generativer KI herauszuholen

Generative KI-Tools funktionieren ähnlich wie Chatbots

NutzerInnen geben Fragen ein, was als Prompt bezeichnet wird, und das Tool gibt Antworten in Form von Text, Code, Bildern usw. aus. Um die besten Ergebnisse zu erzielen, sollten die Prompts an den geschäftlichen Anwendungsfällen ausgerichtet und so spezifisch wie möglich hinsichtlich des gewünschten Ergebnisses sein. Nachfolgend sind drei Faustregeln aufgeführt, die NutzerInnen dabei helfen, effektive Prompts zu erstellen:

Teilen Sie dem KI-Modell mit, welche Rolle es spielen soll

Auf diese Weise kann das KI-Modell besser verstehen, welche Ergebnisse erwartet werden. Beispiele hierfür sind Marketingexperte, Texter oder Social-Media-Experte.

Erklären Sie, was Sie benötigen

Geben Sie klar und deutlich an, nach welcher Ergebnisart Sie suchen. Sie können z. B. Artikel, Blogbeitrag, Social-Media-Beitrag oder Webseite angeben. Oder Sie geben einfach etwas wie „Liste“, „Zusammenfassung“ oder „Aufzählungspunkte“ an.

Geben Sie spezifische und relevante Anweisungen

Dadurch erhält das KI-Modell Eingrenzungen, die es dabei unterstützen, sich auf die richtige Ausgabe zu konzentrieren. Dazu gehören beispielsweise: Zielgruppe, Stil oder Ton, die emuliert werden sollen, Schlüsselwörter, die verwendet oder vermieden werden sollen, und die gewünschte Länge.



Neuanfänge

Generative KI: Eine neue Ära der Produktivität und Innovation

Die Welt befindet sich derzeit am Anfang einer epochalen Transformation. Die generative KI ist dabei, sich als Wegweiser auf dieser transformativen Reise zu etablieren, und zeigt Pfade zu beispielloser Produktivität und Innovation für die sich ständig weiterentwickelnde Zukunft der Arbeit auf.

Diese bahnbrechende Technologie ist nicht nur ein Schritt in der Evolution, sondern ein Sprung in eine neue Ära, in der die Grenzen dessen, was wir bisher für möglich erachtet haben, unaufhörlich verschoben werden.

Die Grenzen der Produktivität überwinden

Generative KI durchbricht die Grenzen der menschlichen Produktivität. Diese Technologie arbeitet rund um die Uhr ohne Unterbrechungen oder Ausfallzeiten. Sie kann bei Workload-Schwankungen mühelos nach oben oder unten skaliert werden, was Unternehmen ein bisher unerreichtes Maß an Flexibilität und Effizienz bietet.⁵ Diese konstante und unerschütterliche Produktivität wird den Begriff Effizienz branchenübergreifend neu definieren. Planen Sie Ihre Strategie noch heute, um Ihre geschäftliche Effizienz zu steigern.

Die Kompetenzlücken der Zukunft schließen

Stellen Sie sich eine Zukunft vor, in der generative KI zum großen ausgleichenden Element in der Belegschaft wird. Die digitale Zukunft entwickelt sich stetig weiter und die Kluft zwischen

vorhandenen Kompetenzen und Branchenanforderungen wird immer größer. Herkömmliche Schulungs- und Entwicklungsinitiativen reichen nicht mehr aus, um diese Lücke zu schließen. Generative KI sorgt mit personalisierten Lernerfahrungen, die auf den individuellen Lernstil und das Tempo der jeweiligen Person zugeschnitten sind, für gleiche Bedingungen und Chancen.

Softwareanbieter nutzen generative KI bereits in ihren Produktivitätssuites und diese Praxis wird sich weiterentwickeln und Innovationen am Arbeitsplatz vorantreiben. Diese Technologie befreit MitarbeiterInnen von Routineaufgaben und macht den Weg für höherwertige Aktivitäten frei, die die geistige Entwicklung und Innovationen fördern. Dies bereichert nicht nur die Arbeitserfahrung, sondern stattet MitarbeiterInnen auch mit neuen Fähigkeiten und Kompetenzen aus, mit denen sie sich auf dem sich ständig weiterentwickelnden Arbeitsmarkt behaupten können.

Schnellere Innovationen

Der vielleicht spannendste Aspekt generativer KI liegt jedoch in der Fähigkeit, Innovationen zu katalysieren. Sie dient als Plattform für Visionäre und ermöglicht es Unternehmen jeder Größe und Struktur, eine Branche zu revolutionieren.

Generatives Design, ein Teilbereich der generativen KI, ermöglicht es Unternehmen, zahlreiche Design-Permutationen schnell zu durchlaufen und Produkte so für Performance, Ästhetik und Kosteneffizienz zu optimieren. Dadurch werden Entwicklungszyklen beschleunigt und Unternehmen haben die Möglichkeit, auf Marktanforderungen und -trends zu reagieren, der Konkurrenz einen Schritt voraus zu bleiben und Innovationen schnell umzusetzen.



Erste Schritte mit generativer KI

Der Weg zum Erfolg beginnt mit den ersten Schritten

RAG (Retrieval-Augmented Generation) ist ein idealer Einstiegsanwendungsfall für viele Unternehmen, da auf zusätzliche Ressourcen wie Ihre eigenen Daten zugegriffen wird, um ein Modell zu erweitern, ohne es neu trainieren zu müssen. RAG wird in LLM-Anwendungen verwendet, um relevante Inhalte im Stile einer Wissensdatenbank abzurufen, Nutzer-Prompts mit diesen bereichsspezifischen Inhalten zu optimieren und dann sowohl die Prompts als auch die Inhalte in das LLM einzuspeisen, um eine umfassendere und nützlichere Antwort zu erzeugen.

Sie können RAG-Anwendungsfälle innerhalb Ihrer eigenen Firewall ausführen, ohne dass Sie Daten im Umfang eines Data Lake benötigen, um einen Proof-of-Concept-RAG-Anwendungsfall zu erstellen und auszuführen. Hierbei handelt es sich um einen häufig genutzten RAG-Anwendungsfall, der Zusammenfassungen in natürlicher Sprache erzeugt, die auf Nutzerabfragen und dem Durchsuchen/Abrufen von Unternehmensdokumenten basieren, und um ein klassisches Beispiel dafür, wie Sie den Wert Ihrer eigenen Daten ausschöpfen können.

Die meisten Unternehmen haben Hunderte von Anwendungsfällen, für die RAG eingesetzt werden kann. Die nachfolgenden einfachen Schritte helfen Ihnen bei der Einrichtung lokaler RAG-Anwendungsfälle:

- ☐ Wählen Sie die Hardware für die Testversion aus.
- ☐ Ermitteln Sie die Anwendungsfälle, die Sie lösen möchten, um die richtigen Daten zu identifizieren.
- ☐ Laden Sie Open-Source-Software herunter, die unstrukturierte Daten in strukturierte Daten verwandelt.
- ☐ Erstellen Sie die Vektordatenbank, in der alle Daten enthalten sind, die das LLM verwenden soll.
- ☐ Wählen Sie ein LLM und seine Parametergröße aus. Davon hängt die Genauigkeit und die erforderliche Hardware ab.
- ☐ Und jetzt viel Spaß beim Prompting!

[Lesen Sie unseren Blog](#), um mehr über die ersten Schritte zu erfahren und wie Sie auf unseren Quellcode für RAG zugreifen können.



Die Zusammenarbeit mit Dell Technologies, einem führenden Technologieunternehmen, wird unsere Ergebnisse weiter beschleunigen, indem wir unser Know-how in KI-Cloud-Infrastruktur mit der globalen Reichweite und den Ressourcen von Dell kombinieren. Gemeinsam sind wir bereit, die KI-Landschaft zu revolutionieren und Unternehmen weltweit in die Lage zu versetzen, die gesamte Leistungsfähigkeit von KI für eine bessere und nachhaltigere Zukunft zu nutzen.

— Dave King,
Mitbegründer und CRO
Denvr Dataworks

Dell Technologies kann Ihren Weg beschleunigen

Erfahren Sie mehr darüber, wie Sie KI zu Ihren Daten bringen und generative KI-Projekte aller Formen und Größen erfolgreich erstellen und implementieren. Wir haben verschiedene Ressourcen zusammengestellt, um Ihnen den Einstieg zu erleichtern.

Einführungsservices und -lösungen

Dell Technologies unterstützt Sie mit einem ganzheitlichen End-to-End-Ansatz vom Schreibtisch bis zum Rechenzentrum dabei, Barrieren aus dem Weg zu räumen und KI-Strategien unternehmensweit einzuführen.

Sie profitieren von einem umfassenden Serviceportfolio für generative KI, mit dem Ihr Unternehmen und Ihre IT-Verantwortlichen Strategien entwerfen können, um den Wert von Datensilos und Data Lakes freizulegen, Kompetenzlücken zu schließen und die

richtige Infrastruktur für Ihre generative KI-Zukunft zu ermitteln.

Wir bieten zudem für jede/n MitarbeiterIn und jeden Anwendungsfall eine zentrale Anlaufstelle für den Kundensupport. Auf globaler Ebene und mit globalen Partnern können Ihnen die Services und Lösungen von Dell für generative KI dabei helfen, sich auf neue Chancen und Möglichkeiten zu konzentrieren und die spürbare Wertschöpfung Ihrer generativen KI-Projekten zu beschleunigen.

Weitere Ressourcen zu generativer KI

- [Bringen Sie KI zu Ihren Daten mit Dell Lösungen für generative KI](#)
- [KI-Technologien unterstützt von Dell Precision-Workstations](#)



Starten Sie Ihre generative KI-Strategie im Rahmen eines kostenlosen 1/2-tägigen Strategiegesprächs, um Ihre wichtigsten Geschäftschancen, Herausforderungen und Prioritäten zu ermitteln.

Kontaktieren Sie uns, um Ihren Workshop zu planen >