

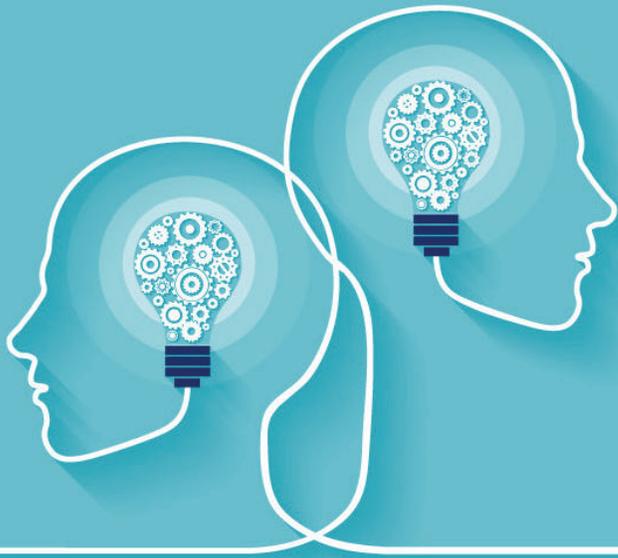
Beschleunigen Sie den Weg zur KI

**mit einer hyperkonvergenten Daten-
und Analyseplattform**

Gesponsert von IBM und Intel

Ritu Jyoti

Februar 2020



EINFÜHRUNG UND ÜBERBLICK

Bis 2024 werden Unternehmen mit künstlicher Intelligenz (KI) durch proaktive, hyperschnelle betriebliche Veränderungen und Reaktionen von Kunden, Wettbewerbern, Regulierungsbehörden und Partnern 50 % schneller reagieren als ihre Konkurrenten. Diese Initiativen zur digitalen Transformation (DX) werden durch Künstliche Intelligenz (KI) unterstützt, die zeitnahe kritische Einblicke, reichhaltigere und immersive Benutzererfahrungen und verbesserte Geschäftsergebnisse ermöglichen.

IDC-Prognosen zufolge werden die Ausgaben für KI bis 2023 weltweit 97,9 Mrd. US-Dollar erreichen, und zwar vorwiegend durch den Einsatz im Bankwesen, Einzelhandel und in der Fertigung. Allerdings ist die Einführung von KI bisher schleppend verlaufen. Die wichtigsten Anwendungen sind derzeit automatisierte Kundenservice-Agenten, Verkaufsprozessempfehlungen, IT- und allgemeine Automatisierungen. Wir gehen jedoch davon aus, dass die Automatisierung der Personalabteilung, digitale Assistenten für Wissensarbeiter in Unternehmen, regulatorische Informationen und fortschrittliche digitale Simulationen, in den kommenden fünf Jahren, die am schnellsten wachsenden Anwendungsfälle sein werden.

KI wird ein echtes Unterscheidungsmerkmal sein, mit Diensten, die vom Edge zu Core bis hin zur Cloud reichen, sowie Hybrid- und Multi-Cloud-Implementierungen als neue Norm. Organisationen, die die KI meistern, werden einen großen Schritt voraus sein und diejenigen, die sie nicht beherrschen, werden zurückfallen. Die effektive Anwendung von maschinellem Lernen (ML) zum geschäftlichen Nutzen erfordert:

- 1. ML-Training:** Machine Learning (ML)-Training umfasst die für den Aufbau von ML-Modellen erforderlichen Schritte und kann Modellgenerierung, -aufbau und -anpassung beinhalten.
- 2. ML-Inferenz:** ML-Schlussfolgerung (d. h. Vorhersage, Scoring oder Model Serve) generiert die Erkenntnisse, die in einen geschäftlichen Anwendungsfall integriert werden müssen, wodurch eine ML-Geschäftsanwendung entsteht, die letztendlich Kundennutzen generiert.

ML-Training und ML-Schlussfolgerung existieren nicht isoliert. Es gibt immer einen Kreislauf, der sie verbindet. Modelle, die durch das ML-Training generiert wurden, müssen zur ML-Inferenz geschickt werden, und sofort oder irgendwann müssen die Erfahrungen aus den Live-Daten genutzt werden, um das Modell in der nächsten Runde des ML-Trainings weiter zu optimieren. Die ML-Schlussfolgerung wird in den geschäftlichen Anwendungsfall integriert, wodurch eine Anwendung entsteht, die Kundennutzen generiert. In der jüngeren Geschichte konnte sich dieser Zyklus über Monate oder sogar Jahre erstrecken, und als solcher konnte man fast vergessen, dass er überhaupt existierte. Mit den Fortschritten bei den Schulungsalgorithmen, der leistungsstarken Hardware und den skalierbaren Analyse-Engines konnten die Laufzeiten für jede Phase jedoch erheblich reduziert werden. Eine KI-optimierte, hyperkonvergierte Daten- und Analyseplattform kann den Weg zur KI beschleunigen und eine schnellere Realisierung des Geschäftswerts ermöglichen.

SITUATIONS- ÜBERBLICK

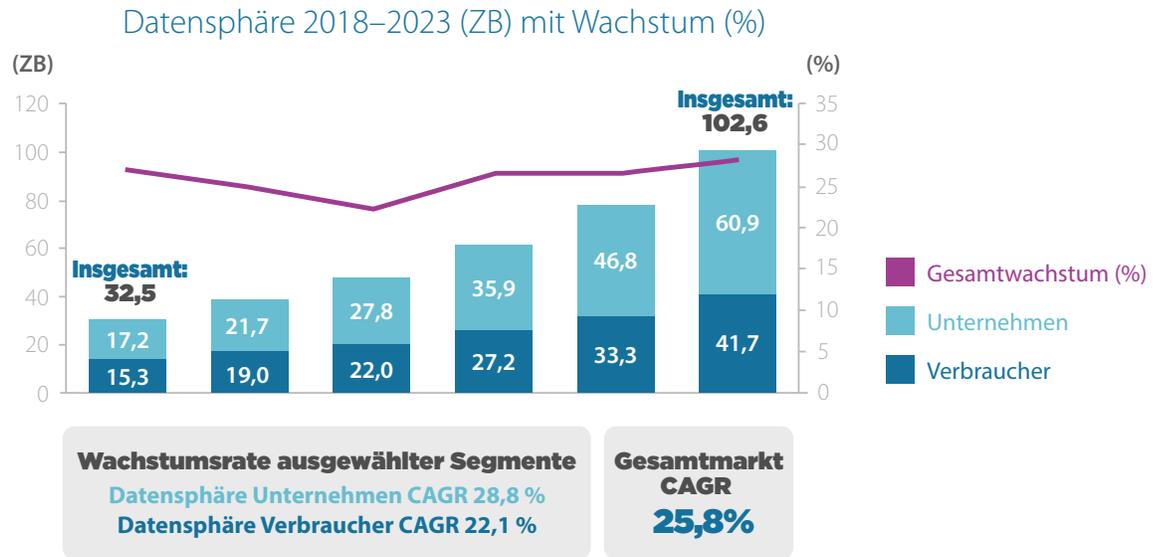


Die globale Datensphäre wächst explosionsartig, ist jedoch nur sehr wenig nützlich für KI

Eine erfolgreiche digitale Transformation setzt voraus, dass Daten in konkrete Erkenntnisse umgewandelt werden. Diese zunehmende Abhängigkeit von datengestützten Insights trägt mit zu einer neuen Phase des Datenzeitalters bei. IDC-Prognosen zufolge soll die globale Datensphäre – also alle weltweit generierten und konsumierten Daten – bis 2023 auf 102,6 ZB anwachsen (siehe Abbildung 1). Dieses enorme Datenvolumen wird potenziell einzigartige Benutzererfahrungen und unzählige neue Geschäftschancen eröffnen.



Abbildung 1. Weltweite globale Datensphäre (102,6 ZB bis 2023)



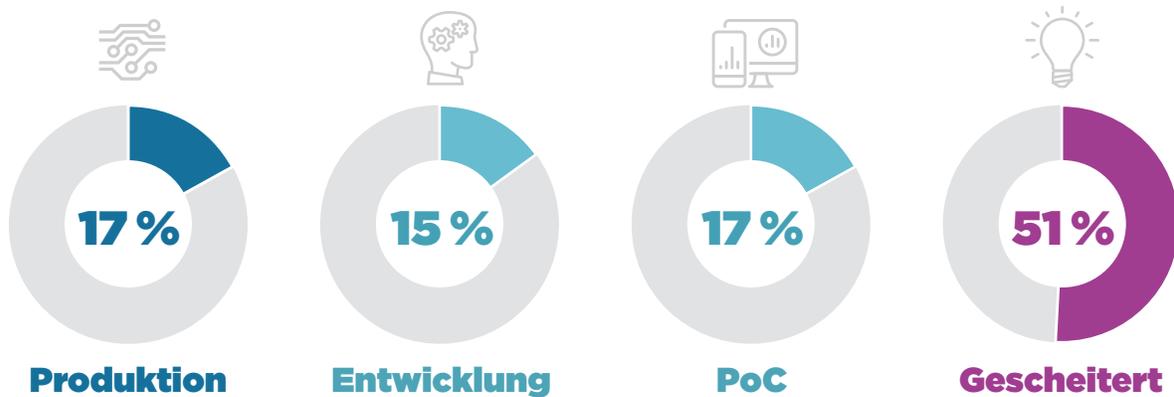
Quelle: IDC, Worldwide Global DataSphere Forecast, 2019–2023:
 Consumer Dependence on the Enterprise Widening, #US44615319

Das explosionsartige Wachstum der generierten Daten ist jedoch nicht das einzige bestimmende Merkmal des neuen Datenzeitalters. Wichtig ist auch, dass nur 27 % der generierten Daten „nutzbar sind, wenn sie getaggt werden“. Dabei werden nur 44 % Daten, die „nutzbar sind, wenn sie getaggt werden“, auch tatsächlich getaggt und nur 21 % der „getaggteten“ Daten werden analysiert. Nach diesem Schritt werden dann nur 15 % der „analysierten“ Daten in KI-Systeme eingespeist. Kalkuliert man diese kaskadierenden Prozentsätze, so ergibt sich ein problematisches Ergebnis: Weniger als 1 % der globalen Datensphäre wird derzeit von KI-Systemen genutzt. Der Rest sind ruhende Daten oder „Dark Data“, die derzeit nicht für Insights oder Entscheidungen herangezogen werden. Da ein so großer Mangel an nutzbaren Daten besteht, generieren Unternehmen synthetische Daten (also Daten-Repositories, die per Programmierung generiert werden), um die für das Training von ML-Modellen erforderlichen Repositories zu erstellen.

Trends und Herausforderungen bei der KI-Übernahme

Künstliche Intelligenz ist im Wettbewerb ein echtes Differenzierungsmerkmal. Mit ihr verbessert sich die geschäftliche Agilität, und die Zeit bis zur Markteinführung neuer Produkte und Dienstleistungen verkürzt sich. Dem im Mai 2019 erstellten *Global Artificial Intelligence (AI) Survey* von IDC zufolge sind nur 17 % aller KI-Initiativen in der Produktion. Weitere 15 % sind in der Entwicklung und 17 % in der Proof-of-Concept-Phase (siehe Abbildung 2). Im Kontrast dazu sind mit 51 % über die Hälfte gescheitert.

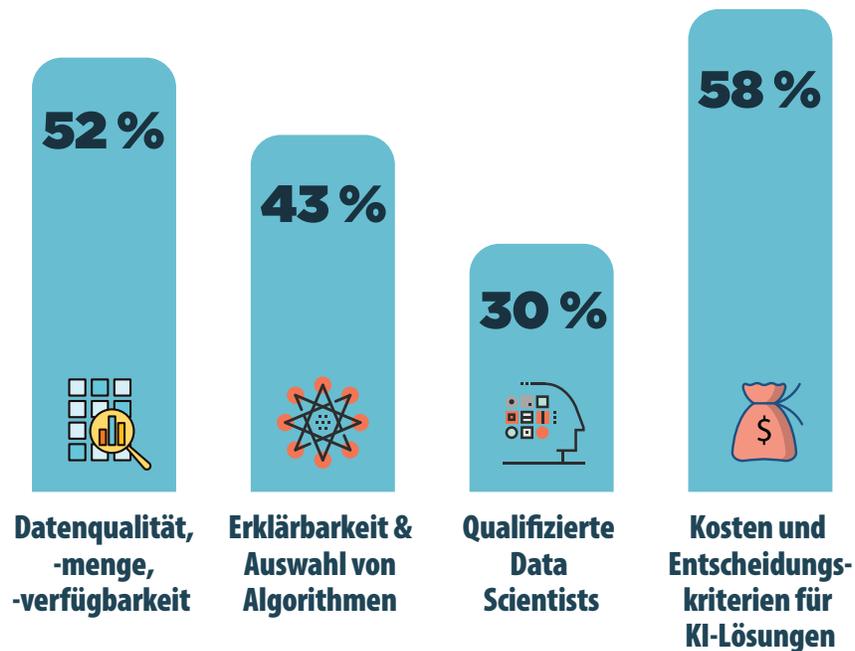
Abbildung 2. KI-Bereitstellungen – derzeitiger Stand



Quelle: Global Artificial Intelligence (AI) Survey IDC, Mai 2019

Hindernisse bei Kosten und Entscheidungskriterien für die KI-Lösung, Datenqualität und -verfügbarkeit, Schwierigkeiten mit den Algorithmen und Mangel an datenwissenschaftlichen Kenntnissen sind die Schlüsselfaktoren, die Unternehmen bei der Implementierung von KI ausbremsen (siehe Abbildung 3).

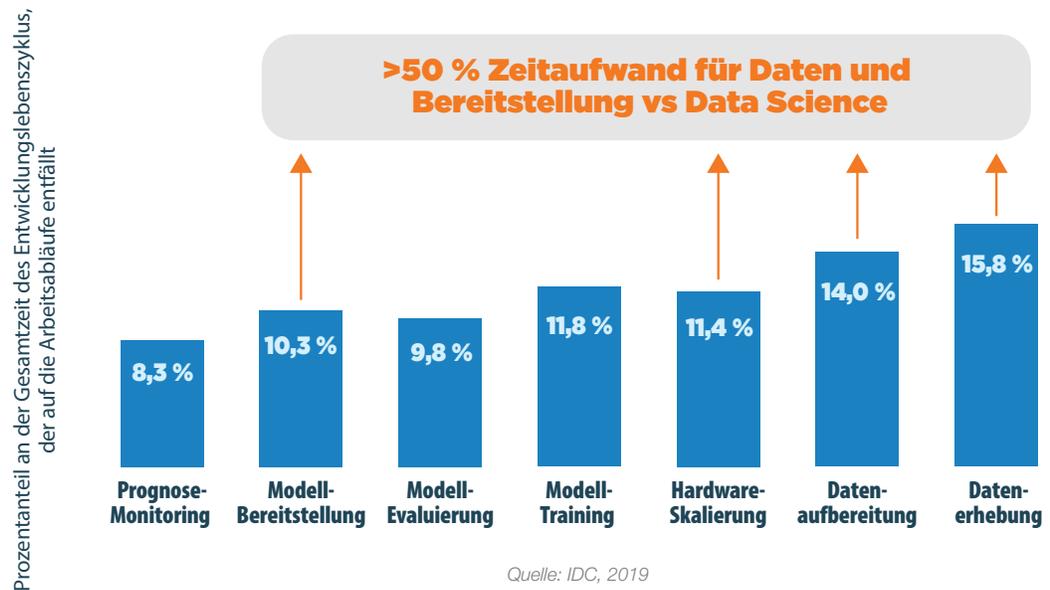
Abbildung 3. Wichtigste Hindernisse für KI-Bereitstellungen



Quelle: Global Artificial Intelligence (AI) Survey, IDC Mai 2019

Einer anderen IDC-Studie zufolge wenden Unternehmen über 50 % ihrer Zeit für Datenaufbereitung und -bereitstellung anstelle des eigentlichen Data Science auf (siehe Abbildung 4). Die Skalierung von Infrastruktur und Leistung behindert den geschäftlich nutzbringenden Einsatz.

Abbildung 4. Daten- und Bereitstellungsmaßnahmen sind zeitaufwendig



Neue Herausforderung im Wettbewerb: MLOps

Um die Herausforderungen bei der KI-Bereitstellung zu bewältigen, werden Unternehmen innovativ: mit der neu entstehenden Praxis der MLOps (zusammengesetzt aus „maschinellern Lernen“ und IT-„Operations“). Schwerpunkt der MLOps sind die Zusammenarbeit und Kommunikation zwischen Datenwissenschaftlern und Betriebsspezialisten. Dies gilt für den gesamten Lebenszyklus von ML/Deep Learning (DL) vom Experiment bis zur Produktion, u. a. mit:



Datenintegration und -katalogisierung



AutoML



Support bei Daten- und Konzeptdrift



Überwachung von Compliance und Sicherheitsvorkehrungen in der Produktion



Unterstützung bei der
Eliminierung von Verzerrungen
(Bias) und Transparenz



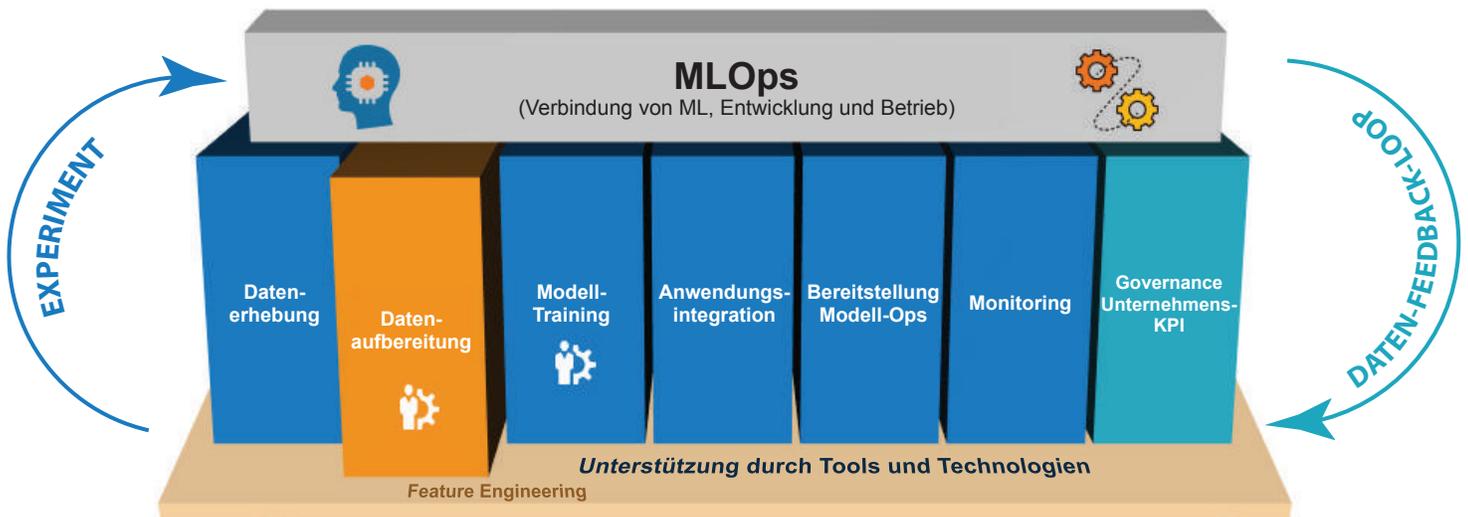
Operationalisierung von Modellen



Sicherheit und Zugriffssteuerung

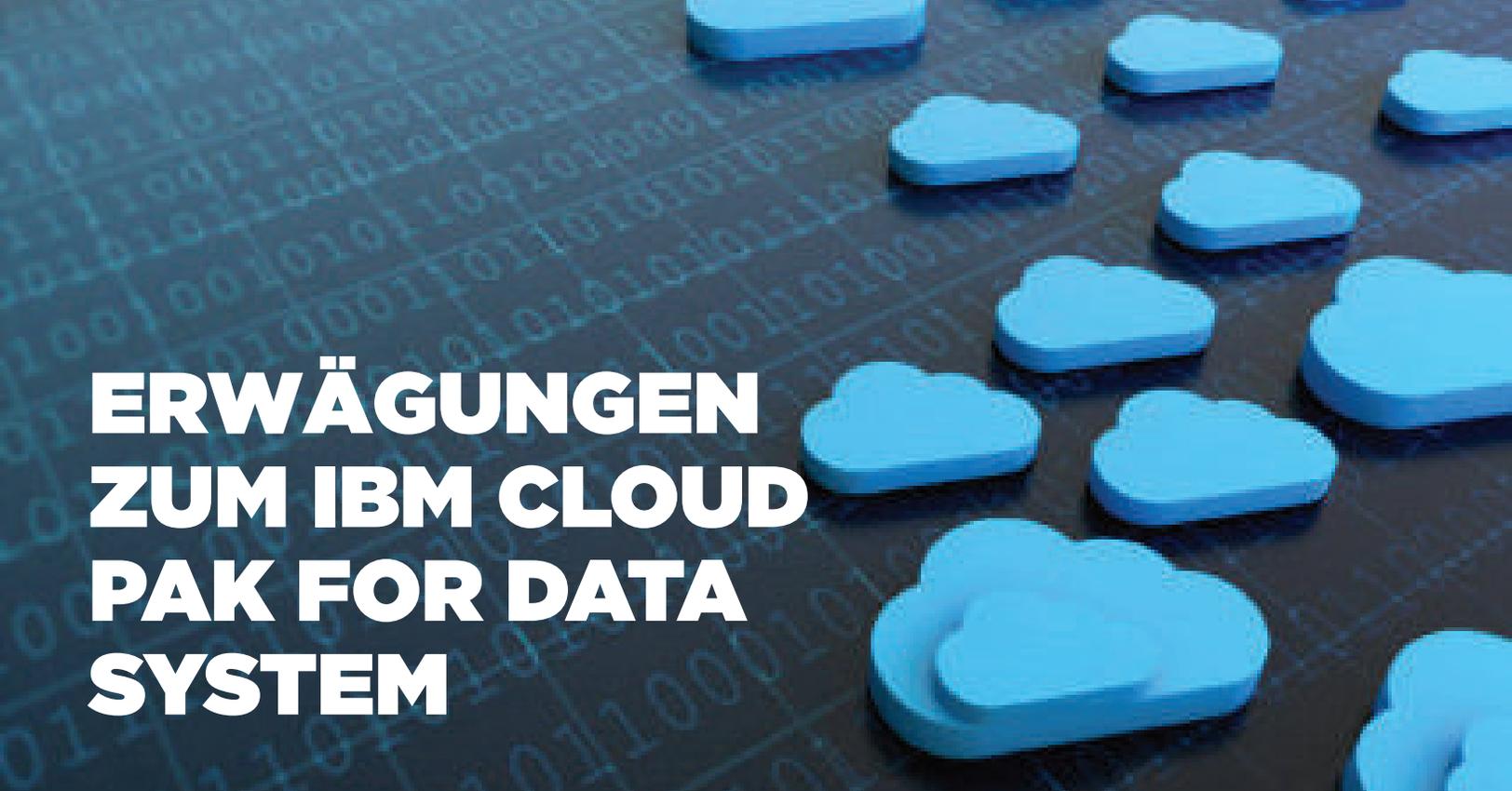
Mit Unterstützung durch Tools und Technologien bietet MLOps Containerunterstützung für Hybrid-, Multicloud- und Edge-Bereitstellungen. So lässt sich ein von „End-to-End“ optimierter, elastischer Stack schaffen: Mit ihm kann mit kleinen Kapazitäten und Leistungen gestartet werden, die dann skaliert werden können – bei gleichzeitig konsistenter und vorhersehbarer Leistung. Effektive MLOps erlauben so eine Kostenkontrolle, vereinfachen das Management und verkürzen die Zeit bis zur Wertschöpfung (siehe Abbildung 5).

Abbildung 5. MLOps über den ML/DL-Lebenszyklus



Praxis der Zusammenarbeit zwischen Data Scientists, Business Analysten, Datenarchitekten und Betriebsspezialisten zur Unterstützung des Managements des ML/DL-Lebenszyklus in der Produktion

Quelle: IDC, 2019

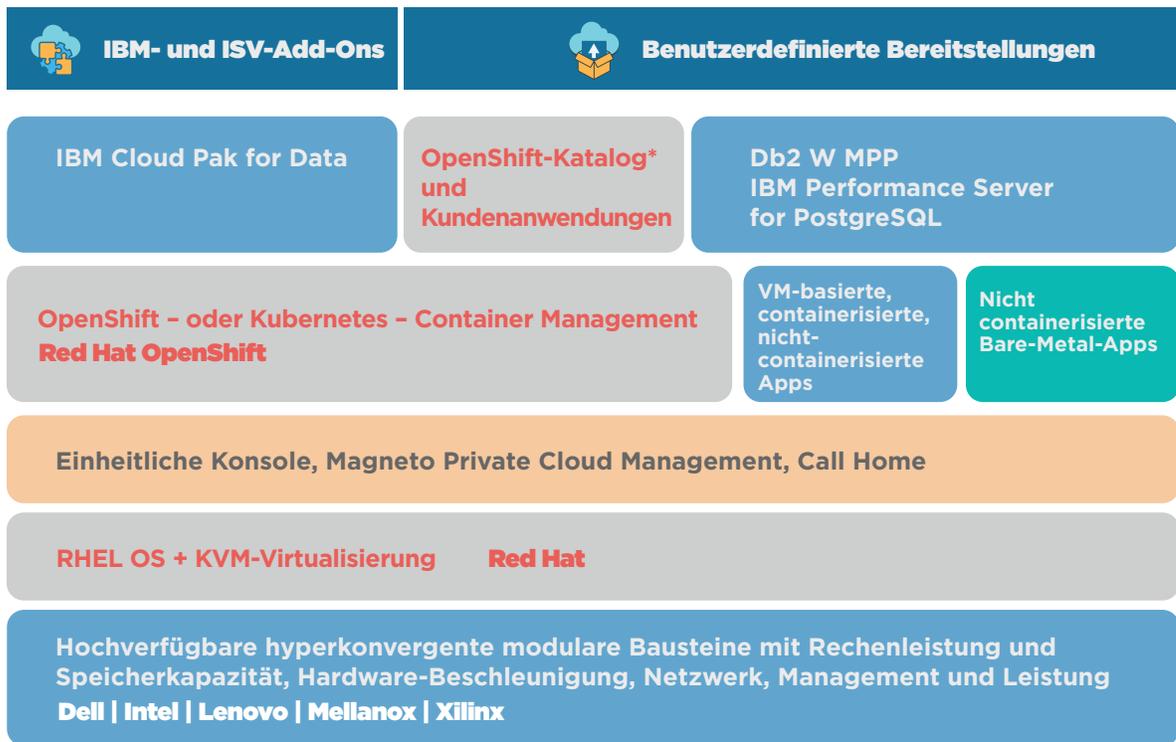


ERWÄGUNGEN ZUM IBM CLOUD PAK FOR DATA SYSTEM

Das IBM Cloud Pak for Data System wurde für die MLOps-Ziele des vereinfachten KI-Lebenszyklus-Managements und der verbesserten Zusammenarbeit konzipiert. Als hyperkonvergente, cloud-native Daten- und KI-Plattform stellt IBM Cloud Pak for Data System eine vorkonfigurierte und sichere Umgebung mit Governance zur Erhebung, Organisation und Analyse von Daten bereit (siehe Abbildung 6). Mit dieser Plattform können Organisationen eine breite Palette von Daten- und KI-Diensten nutzen und in Anwendungen integrieren: So lassen sich die Zeit bis zur Wertschöpfung, bis zur Ermittlung von Einsichten und bis zur Markteinführung verkürzen.

IBM Cloud Pak for Data System baut auf der Red Hat OpenShift Container Platform auf und kombiniert Speicherkapazität, Rechenleistung, Netzwerk und Software in Plug & Play-Knoten. Diese hyperkonvergente Architektur macht das Software- und Hardware-Management einfacher, und mit ihr lässt sich eine private Cloud schon innerhalb weniger Stunden bereitstellen. Ein flexibles nutzungsbasiertes Zahlungsmodell unterstützt die Kostenkontrolle, wobei sich gleichzeitig sowohl Rechenleistung als auch Speicherkapazität skalieren lassen.

Abbildung 6. Architektur des IBM Cloud Pak for Data System



* Kundenseitige Container-Plattform-Lizenz erforderlich

Auf Basis von  Red Hat

Das IBM Cloud Pak for Data System ermöglicht Organisationen die Erhebung, Organisation und Analyse von Daten für KI-gestützte Anwendungen. Das IBM Cloud Pak for Data System ist eine Multicloud-Daten- und KI-Plattform: Sie bietet eine Informationsarchitektur für die KI mit Flexibilität, Sicherheit und Kontrolle bei gleichzeitigen Vorteilen der Cloud, ohne dass Daten bewegt werden müssen. Watson Studio ist in der Konfiguration inbegriffen. Watson Machine Learning und Watson Machine Learning Accelerator sind als Erweiterungsdienste verfügbar.

Das datenwissenschaftliche „End-to-End“-Toolkit von IBM ist hilfreich für Datenwissenschaftler auf jeder Kenntnisstufe:

- Datenaufbereitung
- Aufbau von ML/DL-KI-Modellen
- Training von KI-Modellen durch interaktives oder Batch-Paradigma
- Bereitstellung von Modellen und Management ihres Lebenszyklus

- GPU-Beschleunigung für das Modell-Training und Nutzung von GPUs und Field Programmable Gate Arrays (FPGAs) für Inferenzaufgaben
- Skalierung auf unternehmensweite Bereitstellungen von Datenwissenschaftlern und ML/DL-Modellen

Wichtigste Leistungsmerkmale des IBM Cloud Pak for Data System

Das IBM Cloud Pak for Data System bietet einen modularen Ansatz für Rechenleistung, Netzwerk und Speicherkapazität auf Standardhardware. Zu den wichtigsten Leistungsmerkmalen gehören:

- Red Hat OpenShift-Unterstützung, die für IBM Cloud Pak for Data Services zertifiziert ist
- Open-Source-Governance-Kapazitäten für das Risikomanagement und eine schnellere Umsetzung von Open-Source-basierten KI-Projekten
- Automatisierte Entwicklung von KI-Modellen mit Unterstützung durch AutoAI
- Integriertes Data Science und maschinelles Lernen
- Hochleistungsanalytik (auf Basis des IBM Performance Server for PostgreSQL, der zu 100 % mit Netezza kompatibel ist) in einer „Cloud in a Box“-Konfiguration für die Nutzung der Vorteile hyperkonvergenter Modularität
- Visuelle Anwendungserstellung, visuelles Debugging nahezu in Echtzeit und Support für Red Hat AMQ-Streams
- Neue branchenspezifische Beschleuniger
- Neue Angebote, so genannte Pakete, die IBM Cloud Pak for Data-Berechtigungen enthalten, die zur Ausführung des Dienstes erforderlich sind
- Neues Lizenzmodell für IBM Cloud Pak for Data, das es Unternehmen ermöglicht, Lizenzen für IBM-Software in einem Cloud-zentrierten Modell zu erwerben, das mit einem Red Hat-Abonnementmodell abgestimmt ist, um ein einheitliches Käuferlebnis für beide Produktportfolios zu gewährleisten
- Erweiterbare Drittanbieter-Dienste wie ‚Figure Eight‘ zur Unterstützung der Annotation von Trainingsdaten sowie maschineller Lerninitiativen

Intel für schnellere KI-Ergebnisse

Intel liefert ein innovatives und flexibles Portfolio von Prozessoren und Beschleunigern für die gesamte Datenpipeline. Viele der am weitesten verbreiteten KI- und Analyse-Frameworks wurden für die skalierbaren Intel Xeon-Prozessoren optimiert, um die Leistung deutlich zu steigern. Diese sind als optimierte Softwareerweiterungsdienste mit direktem Zugriff von IBM Cloud Pak for Data System-Software erhältlich.

Intel Xeon Prozessoren, Beschleuniger und Workload-optimierte Software-Add-Ons können dazu beitragen, KI-Ergebnisse zu beschleunigen und die Zeit bis zum Handeln zu verkürzen. Diese Technologien können dazu beitragen, die Entwicklung und Erkenntnisse mit optimierten Frameworks zu beschleunigen und mit einer modernisierten, leistungsstarken und sicheren Plattform, die die Daten schützt, Daten schneller zu sammeln und zu organisieren.

Intel Data Center SSDs mit NVMe-Schnittstelle sind Bestandteil der IBM Cloud Pak for Data Lösung: Sie bieten den Durchsatz und die konsistent niedrige Latenz, die für die Unterstützung von Anwendungen wie Analytik gebraucht werden. Basis der Software von IBM Cloud Pak for Data ist die Red Hat OpenShift Container Platform. Sie profitiert von der engen Zusammenarbeit zwischen Intel und Red Hat mit Investitionen in Co-Engineering und Optimierungen anhand von Intels Bibliotheken und Tools. So lässt sich sicherstellen, dass die Software durch Hardware verbesserte Technologien ausschöpfen kann.

Anpassungsfähig im Hinblick auf neue Geschäftsanforderungen

Das IBM Cloud Pak for Data System wurde als „Plug & Play“ konzipiert und bietet zusätzlich „Plug & Grow“-Optionen, so dass Bare-Metal-Rechen- und Speicherknoten aufgenommen, erkannt und bereitgestellt werden können, damit auch sich neu entwickelnde Geschäftsanforderungen bewältigt werden können. Mit der integrierten Datenvirtualisierung wird der Zugriff auf neue Datenquellen über eine einheitliche Konsole möglich: Sie macht die Bewegung von Daten überflüssig und bietet eine nahtlose Benutzererfahrung über Daten- und KI-Funktionen hinweg.

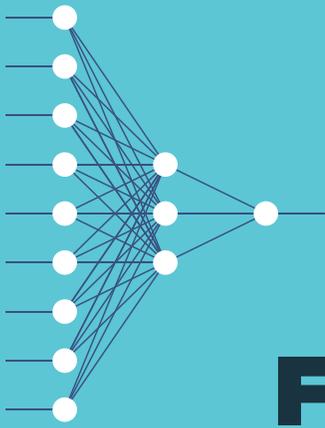


HERAUSFORDERUNGEN UND CHANCEN

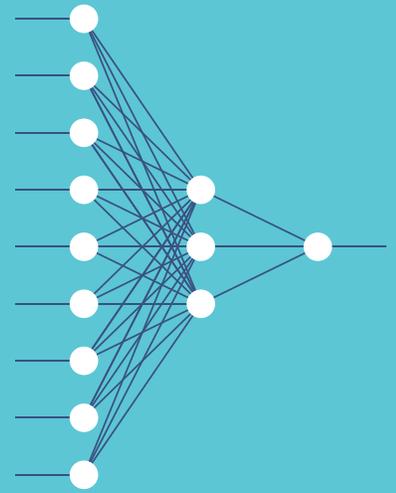
Als Haupthindernisse für die Einführung der KI werden unrealistische geschäftliche Erwartungen an die KI und mangelnde Kenntnisse in Data Science angeführt. IDC empfiehlt IBM Partnerschaften mit einer großen Bandbreite von Systemintegratoren, u. a. Anbietern von Unternehmensleistungen mit KI-Schwerpunkt der jüngsten Generation: Ziel sollte die Vereinfachung von Aufbau, Betrieb und Bereitstellung von „End-to-End“-KI für Unternehmen weltweit sein.

Derzeitige KI-Anwendungen werden Einfluss auf jeden Aspekt unseres Lebens haben – u. a. Verkehr, Finanzen, Einzelhandel, Gesundheitswesen, intelligente Fertigung, Bildung und Dienstleistungsindustrie. KI-Technologien werden bei digital vernetzten Autos, intelligenter Fertigung und medizinischer Bilderkennung an vorderster Front stehen. Für Inferenz am Edge, bei reduzierter Stellfläche und unter häufig unwirtschaftlichen Bedingungen kann IBM Angebote in strapazierfähiger Ausführung erwägen, die Funktionen der Betriebs- und Kundentechnologie einbeziehen. Die inhärente Effizienz der HCI-Architektur mit ihrer Zusammenstellung von Rechen-/ Speicher-/Netzwerktechnik in jedem Knoten kann IBM die Möglichkeit eröffnen, nach unten zu skalieren: So ließen sich Anforderungen der Inferenz am Edge und von Umgebungen mit begrenztem Platz erfüllen. Um sicherzustellen, dass das Angebot leistungsstark und leistungsfähig ist, um unterschiedliche Workload-Anforderungen zu erfüllen, kann IBM versuchen, heterogene Rechenleistung zu vereinen - einschließlich diskreter Beschleuniger wie FPGAs, GPUs, ASICs und ASSPs.

IDC schätzt das cloud-zentrierte Software-Lizenzierungsmodell von IBM auf Abonnementbasis mit flexiblem nutzungs-basiertem Kapazitätsmodell als äußerst gelungen ein. Mit der Reifung der KI-Technologien und dem Aufbau von Vertrauen in die KI in Organisationen durch ergebnisorientierte Preisgestaltung, bei der für Ergebnisse und nicht für Technologie gezahlt wird, kann IBM die Möglichkeit ergebnisorientierter Lizenzmodelle in Erwägung ziehen.



FAZIT



Alle Unternehmen streben nach erhöhter Agilität und rascherer Markteinführung: Da kann es verlockend sein, sich kopfüber in KI-Initiativen zu stürzen. KI entfaltet ihren Nutzen jedoch nicht wie durch Zauberei. Der Nutzen von KI ist vielmehr das Ergebnis strategischer Planung und des ständigen Einsatzes für Datenmanagement.

Vergessen Sie nicht: Daten sind für die KI entscheidend und jeder einzelne Schritt im KI-Lebenszyklus ist von kritischer Bedeutung. Bei KI und ML sind ohne Daten keine Fortschritte möglich: Daher müssen das richtige Verständnis und Management des Lebenszyklus dieser Daten gewährleistet sein. Bei der richtigen Verwaltung der Daten kann KI die Kapazitäten und Chancen einer Organisation transformieren.

Wie können Unternehmen jedoch das Datenmanagement verbessern und KI-Initiativen umsetzen, wenn die erforderlichen Data Scientists so dünn gesät sind? Um dem derzeitigen Mangel an Datenwissenschaftlern abzuhelpfen, sollten sich Unternehmen mit KI-gestützter Self-Service-Datenaufbereitung und automatisierter/unterstützter Funktionsentwicklung auseinandersetzen, mit denen bestehende Teams schon heute das Potenzial der Datenwissenschaft nutzen können. Diese Anwendungen können entweder Daten selbst erfassen oder nahtlos in bestehende Datensysteme integriert arbeiten. Die richtigen Daten sind für die „Feature Generation“ entscheidend – ein Teil des „Feature Engineering“, d. h. Erstellen und Auswählen der Features oder Attribute, die in einem Vorhersagemodell verwendet werden sollen. Darüber hinaus sind Praxis und Disziplin der Zusammenarbeit über Dateningenieure, Datenarchitekten, Datenwissenschaftler, KI-App-Entwickler und MLOps-Spezialisten hinweg unbedingt zu berücksichtigen.

KI wird zu einem entscheidenden Merkmal für die Differenzierung im Wettbewerb – von Edge über Core bis zur Cloud. Durch die geschickte Nutzung der Leistung einer KI-optimierten und skalierbaren, integrierten Lösung lassen sich die Arbeitsabläufe Ihrer Teams verschlanken, der einfache Zugriff auf Daten kann unterstützt werden und konsistente Leistung kann in Verbindung mit einer auf Mikrodienste ausgerichteten Entwicklung sicher und kontrolliert skaliert und gewährleistet werden. Mit einer offenen, interoperablen und standardisierten „End-to-End“-Lösung, welche die unkomplizierte Portabilität von KI-Initiativen möglich macht, erschließt Ihr Unternehmen sich die Möglichkeiten für eine wirkungsvolle Nutzung des wahren Potenzials der KI.

Hinweise des Sponsors

Um mehr über IBM Cloud Pak for Data System, einschließlich Produktdetails und Vorteile zu erfahren, besuchen Sie <https://www.ibm.com/products/cloud-pak-for-data/system>

Ein 7-tägiger kostenloser Test ist verfügbar unter <https://www.ibm.com/account/reg/us-en/signup?formid=urx-34120>

Über IDC

International Data Corporation (IDC) ist der weltweit führende Anbieter von Marktinformationen, Beratungsdienstleistungen und Veranstaltungen auf dem Gebiet der Informationstechnologie und der Telekommunikation sowie der Verbrauchertechnologiemärkte. IDC unterstützt IT-Profis, Business-Entscheider und Investoren bei fundierten Entscheidungen über Geschäftsstrategien und den Einkauf von Technologie. Mehr als 1100 IDC-Analysten in mehr als 110 Ländern bieten globale, regionale und lokale Expertise zu Chancen und Trends in Technologie und Wirtschaft. Seit 50 Jahren bietet IDC strategische Insights, um unseren Kunden zu helfen, ihre wichtigsten geschäftlichen Ziele zu erreichen. IDC ist ein Tochterunternehmen von IDG, einem weltweit führenden Medien-, Research- und Veranstaltungs-Technologieunternehmen.



Internationaler Hauptsitz

5 Speen Street

Framingham, MA 01701

USA

+1 508 872 8200

Twitter: @IDC

idc-community.com

www.idc.com

Urheberrechtshinweis

Externe Veröffentlichung von IDC-Informationen und -Daten: Die Veröffentlichung aller IDC-Informationen, die im Rahmen von Werbemaßnahmen, Pressemitteilungen oder Werbematerial zum Einsatz kommen sollen, muss vorab schriftlich vom entsprechenden IDC Vice President oder Country Manager genehmigt werden. Derartige Anforderungen sind unter Beilage eines Entwurfs des geplanten Dokuments an uns zu richten. IDC behält sich das Recht vor, die externe Nutzung ohne Angabe von Gründen zu versagen.

Copyright 2020 IDC. Jede Wiedergabe ohne vorherige schriftliche Genehmigung ist strengstens untersagt.

