



# Der Weg zu KI und geschäftsfähigen Daten beginnt mit der Informationsarchitektur

Lernen Sie eine zuverlässige Basislösung für die Analyse mit zentralen Funktionen für Governance und Katalogisierung kennen

# Einführung

## Inhaltsverzeichnis

- Einführung
- Die Bausteine einer zuverlässigen Basislösung für die Analyse
- Maschinelles Lernen beschleunigt die Governance
- Eine einzige Basis für viele Einsatzgebiete
- Schlussfolgerung

## Kernaspekte

- Compliance kann Unternehmen ermutigen, kontinuierliche und nutzbringende Datengovernance-Strategien zu implementieren
- Maschinelles Lernen (ML) automatisiert Governance- und Integrationsinitiativen in großem Umfang, wobei die Schwierigkeiten umfangreicher Datenvolumen und die Beschränkungen menschlicher Leistungsfähigkeit überwunden werden
- Datengovernance ist sowohl in On-Premises- als auch in Multicloud-Umgebungen wirksam.

In Unternehmen aller Art multiplizieren sich Daten schnell hinsichtlich Menge und Vielfalt. In Multicloud-Umgebungen erhöhen eine Reihe von Datenquellen exponentiell den Strom eingehender Informationen, vom Internet der Dinge (IoT) und sozialen Medien über mobile Geräte bis hin zu Virtual-Reality-Implementierungen und optischer Verfolgung. Während Unternehmen bereitwillig in künstliche Intelligenz (KI) investieren, haben die meisten die sorgfältige Prüfung vernachlässigt, ihre Daten besser kennenzulernen oder die Datenqualität sicherzustellen, damit man von KI-Lösungen profitieren kann. In vielen Unternehmen sind die Daten nicht zugänglich, unzuverlässig oder verstoßen gegen Datenschutzvorschriften.

Weltweite Verordnungen wie die Datenschutz-Grundverordnung (DSGVO), der California Consumer Privacy Act (CCPA) und das allgemeine brasilianische Datenschutzgesetz (Lei Geral de Proteção de Dados, LGPD) konzentrieren sich auf personenbezogene Kunden- und Mitarbeiterdaten. Diese Art von Vorschriften bieten Unternehmen die Möglichkeit, neue datengestützte Geschäftsmodelle zu transformieren und zu entwickeln, insbesondere durch die hohen Strafen, die sich aus Verstößen ergeben, die Einfluss auf die Produktivität oder den Markenwert nehmen können. Um Datenschutzverpflichtungen zu erfüllen und personenbezogene Daten zu schützen, müssen Unternehmen zuerst ihre verschiedenen Datentypen identifizieren und klassifizieren. Unternehmen, die Schwierigkeiten damit haben, Kundendaten zu erfassen oder ordnungsgemäß zu verwenden, können akute Probleme bekommen. Um dieser Herausforderung gerecht zu werden, implementieren Unternehmen [geregelt Informationsarchitekturen](#), die Verordnungen einhalten und gleichzeitig weiterhin datengesteuerte betriebliche Leistung und Innovation unterstützen.

# Die Bausteine einer zuverlässigen Basislösung für die Analyse

Die Annäherung an Datenschutzverordnungen als Verpflichtung und als Gelegenheit, Dateninfrastrukturen zu modernisieren, bietet einen erheblichen Vorteil. Auf diese Weise werden Unternehmen ermutigt, Strategien für die Datengovernance zu implementieren, die neue Geschäftsmodelle hervorbringen und zu datengesteuerten Erkenntnissen führen. Initiativen für einheitliche Governance und Integration gelten sowohl für strukturierte als auch für unstrukturierte Daten in Public und Private Clouds. Die Implementierung einheitlicher Governance und Integration ist an sich schon wichtig, ihr Nutzen wirkt sich jedoch auch auf andere Unternehmensbereiche aus, insbesondere auf die Regulierung von KI-Modellen für Data Scientist.

Wenn ein Unternehmen Datengovernance einsetzt, um seinen Daten vertrauen zu können, wissen Benutzer, dass die Daten aus einer qualitativ hochwertigen Quelle stammen. Sie wissen, wie die Daten im Unternehmen verwendet werden und wie dadurch Analyseprojekte verbessert werden. Für Analyseinitiativen sind zuverlässige Daten notwendig, um effektives Arbeiten zu ermöglichen, unabhängig davon, wie innovativ die Tools auch sein mögen. Die Vorteile zuverlässiger, geschäftsfähiger Daten scheinen unbegrenzt. Analyse kann neue Produktdesigns und Marketingprogramme empfehlen und Vertriebs-, Lieferketten- und Kundendienstinitiativen verbessern. Analyse kann sogar betriebliche Ineffizienzen ans Licht bringen, die, sobald sie behoben werden, die betriebliche Agilität verbessern und den Reinertrag steigern können.

Die Implementierung von Daten- und KI-Governance in Ihrem Unternehmen setzt sich aus folgenden Bausteinen zusammen.

## Datenerkennung und -qualität

Unternehmen sind sich der großen Datenvolumen, die in ihrem Unternehmen gespeichert sind, möglicherweise nicht bewusst. Der erste Schritt bei der Datengovernance ist die Bestandsaufnahme betrieblicher Daten. Konzentrieren Sie sich zuerst auf Datasets im Rahmen eines bestimmten Projekts und gehen Sie dann zu anderen Geschäftsbeispielen über, um eine breitere betriebliche Abdeckung zu erreichen. Redundante, veraltete oder belanglose Daten sind nicht nur teuer bezüglich Speicherung und Management, sondern **stören zudem die Entscheidungsfindung und den Betrieb**. Sie können auch die Compliance erschweren und Analysearbeiten behindern. Daten müssen bestimmte Qualitätskennzahlen erreichen und beibehalten, damit eine nachgeordnete Nutzung erfolgreich möglich ist.

## Katalogisierung

Sobald Daten identifiziert wurden und ein entsprechendes Profil erstellt wurde, werden sie unter Verwendung von Metadaten tags katalogisiert, um Datentypen, Verwendung, Eigentum, Datenabstammung etc. zu ermitteln. Da Unternehmen in bestimmten Branchen gemeinsame Anforderungen haben, können vordefinierte Branchenmodelle den Katalogisierungsprozess beschleunigen, indem direkt verfügbare Geschäftsbedingungen und -strukturen verwendet werden. Durch Fortschritte beim maschinellen Lernen können Geschäftsbedingungen automatisch zugeordnet werden, um einen **Unternehmenskatalog** in wenigen Stunden zu erstellen. Dank der Basis für die Katalogisierung durch einheitliche Governance und Integration können Unternehmen ihre KI-Modelle, Notizbücher und andere Datenquellen regulieren und damit eine zentrale Bibliothek für betriebliches Wissen erstellen. Eine solche Basis ist eine Ressource für viele Datenbenutzer im Unternehmen, einschließlich Data-Engineers, Data-Stewards und Line-of-Business-Benutzern (LOB) wie Analysten, Data-Scientists und Marketingexperten.

## Übertragung, Transformation und Synchronisation von Daten

Daten aus verschiedenen Quellen können problemlos integriert, transformiert und bei Bedarf mit anderen Systemen physisch oder virtuell gemeinsam genutzt werden. Dieser Prozess führt strukturierte und unstrukturierte Daten zusammen und ermöglicht die Integration in offene Technologien wie Apache Atlas und Hadoop. Durch die Erstellung eines automatisierten Datenflusses und Synchronisation wird sichergestellt, dass die neuesten Daten in Data-Lakes, Data-Warehouses, Datamarts und Point-of-Impact-Lösungen verfügbar sind. Da die Datenmengen zunehmen, unterstützt die Replikation umfangreiche Volumen mit niedriger Latenz. Unternehmen können Virtualisierung verwenden, ohne Daten ausgehend von ihren Anforderungen verschieben zu müssen.

Bis zum Jahr 2019 wird die Analyseausgabe von Fachanwendern mit Self-Service-Funktionalität die von professionellen Data Scientist übertreffen.

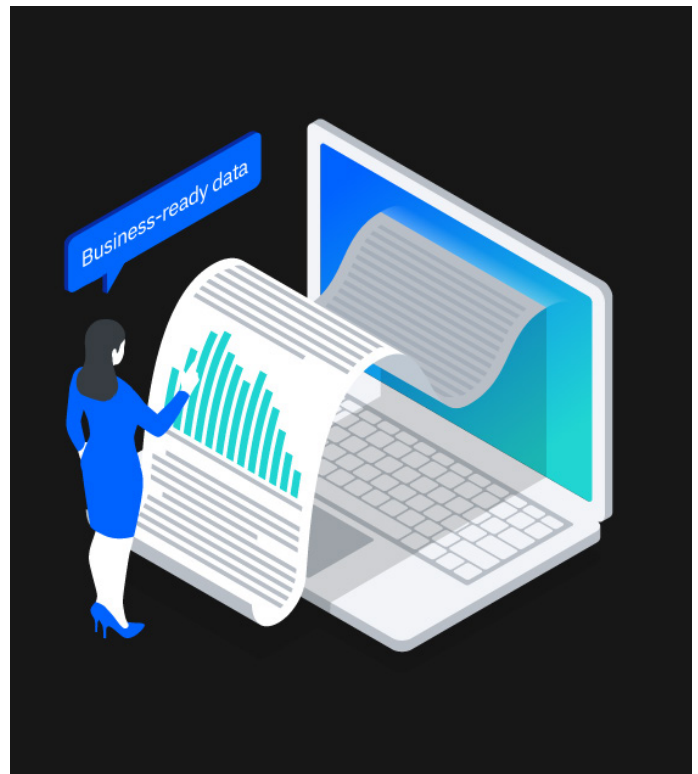
## Master Data Management

Laut Gartner wird bis zum Jahr 2019 die Analyseausgabe von Fachanwendern mit Self-Service-Funktionalität die von professionellen Data Scientist übertreffen. Für Unternehmen ist es entscheidend, sich auf eine umfassende, zuverlässige und einheitliche Ansicht kritischer Entitäten wie Kunden, Produkten und Konten verlassen zu können. **Moderne Master Data Management-Implementierungen (MDM)** bieten analytische grafikbasierte Exploration, eine hoch präzise ‚Matching Engine‘, ein Data-First-Konzept bei der Auswahl passender Algorithmen und durch maschinelles Lernen unterstützte Stewardship-Prozesse. Darüber hinaus verfügen MDM-Lösungen über agilen Self-Service-Zugriff, Governance-Tools und benutzerfreundliche Dashboardfunktionalität.

## Datenschutz

Unternehmen müssen ihre strategischen und sensiblen Informationsassets proaktiv schützen. Das Management des Datenlebenszyklus reicht von der Erstellung bis zur Löschung unter Verwendung von Verfahren wie Schriftgutverwaltung, Aufbewahrungspflicht und Archivierungsspeicherung. Wenn kognitives Lernen auf die Dokumente und Protokolle eines Unternehmens angewendet wird, können Risiken jetzt automatisch, basierend auf dem Unternehmenskontext identifiziert werden.

Daten, die sowohl für Betriebsabläufe als auch für Compliance reguliert werden, sind geschäftsfähig und können direkt für Entscheidungsfindung, Verbesserung oder Innovation verwendet werden. Da die Datenmengen zunehmen, unterstützt die Replikation umfangreiche Volumen mit niedriger Wartezeit. Unternehmen können Virtualisierung verwenden, ohne Daten ausgehend von ihren Anforderungen verschieben zu müssen.



# Maschinelles Lernen beschleunigt die Governance

Maschinelles Lernen erweitert jetzt die menschliche Intelligenz und gleicht die erheblichen Beschränkungen menschlicher Leistungsfähigkeit dank aktueller technologischer Fortschritte aus. Es automatisiert Governance- und Integrationsinitiativen in großem Umfang, wobei die Schwierigkeiten umfangreicher Datenvolumen überwunden werden, was zu einer leistungsfähigen Datengovernance im gesamten Unternehmen führt. Wenn ein Unternehmen beispielsweise über 20.000 Datenbegriffe verfügt, benötigt ein Team von sechs Personen normalerweise

sechs Monate für die manuelle Klassifizierung der Begriffe, um Analysen zuverlässig und vertrauenswürdig zu unterstützen. Mit maschinellem Lernen kann dieser Prozess in wenigen Tagen oder sogar Stunden abgeschlossen werden, je nach Menge der Datenbeständen. Durch diese Beschleunigung wird der Governance-Prozess insgesamt kostengünstiger. Durch maschinelles Lernen können das Management von Compliance-Verpflichtungen vereinfacht und der Weg zu produktiven Analyseinitiativen geebnet werden.

## Manuelle Klassifizierung

Sechs Monate bis zum Abschluss



## Durch maschinelles Lernen gestützte Klassifizierung

Tagen oder Stunden bis zum Abschluss



# Eine einzige Basis für viele Zwecke

Wenn es eine regulierte Basis gibt, kann sie für verschiedene Geschäftsbereiche und im ganzen Unternehmen verwendet werden, wie in den folgenden allgemeinen Beispielen beschrieben.

## Regulierter Data-Lake

Von vielen Unternehmen in den verschiedensten Branchen wurden Projekte im Bereich Big Data-Vorhersageanalyse durchgeführt. Ein wichtiger erster Schritt ist die Speicherung umfangreicher Mengen strukturierter und unstrukturierter Daten in einem Data-Lake.

Unternehmen haben für ihre Data-Lakes Hadoop- oder Handcodierungslösungen verwendet. Das Fehlen einer Governance-Struktur gepaart mit dem Fehlen einer Strategie für das Management von Datenstandards, Geschäftsbedingungen, Herkunft, Nutzung und Qualität kann zu einem Datensumpf führen. Dies geschieht, wenn Benutzer ihre Daten weder kennen noch diesen vertrauen. Unternehmen sehen, dass für ihren Data-Lake eine Strategie und Governance notwendig sind, damit erfolgreiches Arbeiten möglich ist.

In einer **regulierten Implementierung** werden Daten in einem Data-Lake Geschäftsbedingungen zugeordnet, die jeder Datenbenutzer leicht versteht und die im gesamten Unternehmen einheitlich sind. Wenn diese Daten für Benutzer verfügbar sind, kann dies die Wertschöpfung für praktisch jedes Self-Service-Data-Science-, Datenexplorations- oder KI-Projekt beschleunigen, was Agilität bietet. Dieser Zugriff bildet die erforderliche Grundlage zur Unterstützung von Multicloud-Umgebungen, On-Premises-Architekturen und den verschiedensten Datenquellen.

## Anwendungsmodernisierung

Unternehmen investieren in großem Umfang in die Modernisierung von Anwendungen, um die Effizienz zu steigern, Kosten zu reduzieren und einen Wettbewerbsvorteil zu erzielen. Je nach Unternehmen können sich Ideen für die Anwendungsmodernisierung auf verschiedene Art äußern. Zu den wichtigsten Überlegungen zählen unter anderem Testdatenmanagement, Datenvirtualisierung und Konnektivität. Unternehmen

verwenden jetzt agile Methodik zur Durchführung von Test- und Entwicklungsarbeiten gleichzeitig mit **virtuellem Datenzugriff**. Sie verlassen sich auch auf die Verbindung von Geschäftsanwendungen und Daten über flexible Integrationsfunktionen.

## 360-Grad-Ansicht von Kundendaten

Es ist wichtig, dass Mitarbeiter über zuverlässige, aktuelle und genaue **zentral angezeigte Stammdaten** in Bezug auf Kunden, Produkte oder sonstige Entitäten verfügen. Fehlerhafte oder veraltete Informationen können Kundeninteraktionen beeinträchtigen und Vertrauen untergraben, zu Kontofluktuationen führen oder die Kosten für die Lieferkette steigern. Daten, die über Prozesse basierend auf einheitlicher Governance und Integration erhoben werden, können dazu beitragen, über Kundeninteraktionen Vertrauen, Markentreue und Eigenkapital zu verbessern und die Agilität der Lieferkette zu steigern.

## Optimierung des Data-Warehouse des Unternehmens

Mit dem Data Warehouse des Unternehmens stellt die Optimierung der Architektur ein Upgrade und eine umfassende Veränderung dar, wie Daten abgerufen, gespeichert, vorbereitet, reguliert und analysiert werden. Eines der **effektivsten Optimierungskonzepte** ist die Auslagerung von ETL-Jobs (Extract, Transform, Load), Daten, die nicht mehr verwendet werden, und Daten, die in Explorationsmodellen benötigt werden. Dieser Prozess reduziert nicht nur Kosten, sondern bietet auch die Möglichkeit, dass die Daten mit anderen Datentypen in Umgebungen wie regulierten Data-Lakes kombiniert werden, um eine dynamische Datenexploration zu ermöglichen.

## Einhaltung von Vorschriften

Eine zuverlässige Basislösung für die Analyse ermöglicht und **beschleunigt die Einhaltung** gesetzlicher Vorschriften. Der Weg zum Schutz personenbezogener Daten beginnt mit der Definition über welche personenbezogenen Daten das Unternehmen verfügt. Der maßgebliche Datenkatalog enthält die geltenden Regeln für Datenqualität, -aufbereitung und -analyse sowie Compliance-Richtlinien.

### Fazit

Bei der digitalen Transformation von Unternehmen werden Führungskräften die Vorteile von Governance im Rahmen ihrer Daten- und KI-Modelle bewusst, sei es in On-Premises- oder in Multicloud-Umgebungen. Durch Konzentration auf zentrale Governance-Verfahren bereiten Unternehmen ihre Daten und KI nicht nur auf die analytische Verarbeitung und Erkenntnisgewinnung vor, sondern auch auf die Einhaltung der entsprechenden Verordnungen. Bei umfangreichen Datenvolumen tragen Verfahren für maschinelles Lernen und KI dazu bei, menschliche Leistungsfähigkeit und Intelligenz im Rahmen von Aufgaben wie Datenzuordnung, Katalogisierung, Abgleich großer Datenvolumen und Beibehaltung der Datenqualität zu erweitern.

Führungskräfte mit Visionen wissen, dass sich der Zeitaufwand für die Erstellung einer soliden Basislösung für einheitliche Governance und Integration heute und in naher Zukunft auszahlen wird. Sie realisieren, dass ihre Unternehmen wichtige Vorteile erzielen, wenn sie Governance als Enabler geschäftlicher Optimierung, Innovation und Compliance für Daten- und KI-Initiativen nutzen. Es ist entscheidend, Lösungen zu verwenden, die Datenoperationsmanagement von der Erstellung bis zur Nutzung umfassen. Für die Optimierung dieser Operationen sind Synergieeffekte, Skalierung und gemeinsame Nutzung erforderlich.

**IBM Deutschland GmbH**  
IBM-Allee 1  
71139 Ehningen  
[ibm.com/de](http://ibm.com/de)

**IBM Österreich**  
Obere Donaustraße 95  
1020 Wien  
[ibm.com/at](http://ibm.com/at)

**IBM Schweiz**  
Vulkanstrasse 106  
8010 Zürich  
[ibm.com/ch](http://ibm.com/ch)

Die IBM Homepage finden Sie unter:  
[ibm.com](http://ibm.com)

IBM, das IBM Logo und [ibm.com](http://ibm.com) sind eingetragene Marken der IBM Corporation in den USA und/oder anderen Ländern. Weitere Produkt- und Servicennamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der IBM Marken finden Sie auf der Webseite „Copyright and trademark information“ unter: [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml).

Dieses Dokument ist zum Datum seiner Erstveröffentlichung aktuell und kann jederzeit von IBM geändert werden. Nicht alle Angebote sind in jedem Land verfügbar, in dem IBM tätig ist.

Alle hier beschriebenen Kundenbeispiele dienen lediglich zur Veranschaulichung. Die tatsächlichen Ergebnisse beim Leistungsverhalten sind abhängig von der jeweiligen Konfiguration und den Betriebsbedingungen.

Vertragsbedingungen und Preise erhalten Sie bei den IBM Geschäftsstellen und/oder den IBM Business Partnern. Die Produktinformationen geben den derzeitigen Stand wieder. Für IBM Produkte gelten die Gewährleistungen, die in den Vereinbarungen vorgesehen sind, unter denen sie erworben werden.

Der Kunde ist für die Einhaltung der geltenden Gesetze und Verordnungen selbst verantwortlich. IBM erteilt keine Rechtsberatung und gibt keine Garantie bzw. Gewährleistung bezüglich der Konformität von IBM Produkten oder Services mit den geltenden Gesetzen und gesetzlichen Bestimmungen.

© Copyright IBM Corporation 2019

32017532-DEDE-01

### Weitere Informationen

IBM Lösungen für einheitliche Governance und Integration unterstützen Sie dabei, eine zuverlässige Basislösung für die Analyse zu erstellen, um KI im richtigen Maß zu unterstützen.

**Besuchen Sie die Website unter [ibm.com/unified-governance-integration](http://ibm.com/unified-governance-integration)**

#### Mehr erfahren

Informieren Sie sich über kognitive Datengovernance und erfahren Sie, wie IBM Lösungen mit maschinellem Lernen und KI eine Führungsrolle einnehmen. [Whitepaper lesen.](#)

#### Beratung durch Experten

Sprechen Sie mit Vordenkern, Distinguished Engineers und Experten für einheitliche Governance und Integration, die mit Tausenden von Kunden zusammengearbeitet haben, um überzeugende Daten-, Analyse- und KI-Strategien zu entwickeln. [Termin für 30-minütige Beratung vereinbaren](#)

