

# AUF KI/MACHINE LEARNING BASIERENDE DATENVERWALTUNG FÜR DIE DIGITALE TRANSFORMATION

ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) Bericht,  
Verfasser: John L. Myers

Oktober 2018



IT & DATA MANAGEMENT RESEARCH • INDUSTRY ANALYSIS • CONSULTING

IM AUFTRAG VON:



Informatica™

# ZUSAMMENFASSUNG

Im Zeitalter datenbasierter Unternehmen steigt die Bedeutung von Initiativen zur Förderung der digitalen Transformation von Jahr zu Jahr. Unternehmen nutzen neue Datenquellen und beziehen Daten aus dem Internet of Things, Echtzeit-Anwendungen und mobilen Apps. Zudem versuchen sie, dafür zu sorgen, dass Prozesse und Abläufe papierlos vonstatten gehen. Sie bemühen sich, sich durch innovative Angebote von der Konkurrenz abzusetzen, um neue Umsatzquellen zu nutzen, ihren Kunden bessere Erlebnisse zu bieten und ihre Betriebskosten zu senken, um die Gewinnmargen zu vergrößern. Das geht nur mithilfe neuer Methoden zur Verwaltung ihrer neuen, digital transformierten Datenbestände.

Business-Analysten und Datenwissenschaftler haben in datenbasierten Unternehmen die Aufgabe, innovative Erkenntnisse zu ermitteln. Business-Analysten untersuchen Möglichkeiten zur Gewinnung von Wettbewerbsvorteilen mithilfe von Ereignisdaten. Datenwissenschaftler entscheiden, welche Daten für die Algorithmen für Advanced Analytics und Machine Learning am besten geeignet sind. Doch wenn sie die in ihrer Datenlandschaft gesammelten Informationen nicht verstehen, ist dies reine Zeitverschwendung. Business-Analysten müssen in der Lage sein, Datensätze ordnungsgemäß zu untersuchen und zu pflegen, um das Pricing, die gezielte Kundenansprache und Betriebsabläufe zu verbessern. Datenwissenschaftler müssen in der Lage sein, schnell die richtigen Domäneninformationen für ihre Modelle zu finden, da die Qualität ihrer Arbeit ansonsten sinkt. Doch Zugriff auf Daten zu haben und in der Lage zu sein, diese Daten auch tatsächlich zu finden sind zwei verschiedene Paar Schuhe.

Traditionelle Analytics-Umgebungen, wie Enterprise Data Warehouses und Data Marts, werden der vielfältigen Datennutzung und der Vielzahl von Daten-Initiativen nicht mehr gerecht. Diese traditionellen Analytics-Umgebungen sind für Unternehmen zwar noch wichtig, doch für datenbasierte Unternehmen kaum noch ausreichend. Unternehmen suchen nach flexibleren und agileren Methoden und Best Practices für innovative, datenbasierte Initiativen.

## WICHTIGE FAKTEN

**WER:** Führungskräfte, Architekten und Business-Stakeholder in datenbasierten Unternehmen

**WANN:** Vorhandene Data-Governance- und Datenverwaltungspraktiken werden der datenbasierten Unternehmenskultur und dem aktuellen Geschäftstempo nicht mehr gerecht

**WAS:** Anleitung, wie datenbasierte Unternehmen ihre Abläufe durch Datenverwaltungsplattformen verbessern können, die auf KI und Machine Learning basieren

Daher sind moderne Architekturen und Ansätze für Analytics erforderlich. Datenbasierte Unternehmen müssen Mitarbeitern in sämtlichen Geschäftsbereichen Zugriff auf die richtigen Daten ermöglichen, um fundierte Entscheidungen treffen zu können. An dieser Stelle kommt oft ein Enterprise Data Catalog zum Einsatz, der umfassende und gepflegte Metadaten umfasst, sowohl technische, betriebliche als auch verfahrenstechnische Informationen. Aufgrund des Wachstums und des Umfangs der Datenlandschaft müssen moderne Datenkataloge KI/ML nutzen, um plattformübergreifend skaliert zu werden, damit nicht nur Stichproben von sich ändernden Daten, sondern das gesamte Datenvolumen regelmäßig ermittelt und inventarisiert werden kann.

Mit all diesen Informationen können KI-/ML-Funktionen Daten in verschiedenen Speicherorten katalogisieren und kategorisieren, ermitteln und zuordnen. Zudem können Informationen für jeden Mitarbeiter des Unternehmens zuverlässig gefunden werden. Der sichere Datenzugriff sollte für jeden Mitarbeiter möglich sein – und nicht nur für einige wenige Anwender.

# INHALTSVERZEICHNIS

Datenbasierte Unternehmen treiben die digitale Transformation aktiv voran.....	1
Demokratisierung: Auf KI basierender Enterprise Data Catalog.....	2
Big Data, Data Lakes, Data Science und Machine Learning.....	3
Strategische Data Governance .....	4
Datenschutz und -sicherheit .....	5
Hybrid- und Multi-Cloud-Bereitstellungen.....	6
ÜBER UNSEREN SPONSOR: Informatica.....	7

# DATENBASIERTE UNTERNEHMEN TREIBEN DIE DIGITALE TRANSFORMATION AKTIV VORAN

Unternehmen setzen datenbasierte Strategien erfolgreich um, die präzisere und agilere Betriebsabläufe ermöglichen. Der Einzelhandel ist beispielsweise gerade im Wandel begriffen, da Märkte, Produktangebote und Vertriebskanäle transformiert werden. So verwendet Amazon beispielsweise datenbasierte Strategien, um Kunden und Lieferanten besser zu verstehen und um maßgeschneiderte Angebote zu erstellen. Andere datenbasierte Unternehmen nutzen neue Absatzchancen, indem Lieferanten, die bislang einfach nicht erkannt wurden, in den Bereichen Personentransport und Hotelunterkünfte miteinander in Beziehung gebracht werden. So haben beispielsweise Uber, Lyft und Airbnb neue Geschäftskategorien geschaffen, die auf Agilität und Daten basieren anstatt auf den Ergebnissen eines Ertragsmodells. Dadurch haben sie den Weg für andere Unternehmen geebnet, ebenfalls innovative Schritte zu ergreifen.

Datenbasierte Initiativen beruhen auf der Fähigkeit, detaillierte Transaktionsdaten aus dem gesamten Unternehmen abzurufen. Beispielsweise anhand von Sensorwerten von IoT-Geräten oder Echtzeit-Ereignissen von Unternehmensanwendungen, wie Bestell- und Zahlungsplattformen. Die digitale Transformation äußert sich darin, dass vorhandene Betriebsabläufe papierlos werden. Die Transformation von Ereignissen und Prozessen bietet nicht nur Einblick in moderne Initiativen, sondern ermöglicht auch einen Überblick über den gesamten Betrieb eines Unternehmens.

## 1. STRATEGISCHE NUTZUNG VON DATEN, UM WETTBEWERBSVORTEILE ZU GEWINNEN

### WICHTIGE ASPEKTE FÜR DATENBASIERTE STRATEGIEN:

- Auffindung und Demokratisierung von Daten
- Bereitstellung und Nutzung von Daten mit verschiedenen Latenzen
- Skalierbarkeit, um Datenwachstum und neuen Datentypen gerecht zu werden
- Nutzung von Metadaten und KI/ML zur Automatisierung von Prozessen und Steuerung des Nutzerverhaltens
- Data Governance, um Risiken zu minimieren und den geschäftlichen Mehrwert von Daten zu maximieren
- Richtlinien und Systeme, um sensible Daten zu erkennen und zu schützen

Aufgrund der Verknüpfung dieser neuen Datenquellen und neu formatierten Informationen müssen Unternehmen ihre technischen und betrieblichen Abläufe modernisieren, um ihre Daten effektiv zu verwalten und zu pflegen. Dazu müssen bei Infrastruktur und Praktiken folgende Aspekte berücksichtigt werden:

- Auf KI/ML basierender Enterprise Data Catalog
- Big Data, Data Lakes, Data Science und Machine Learning
- Strategische Data Governance/MDM
- Datenschutz und -sicherheit (personenbezogene Daten und DSGVO)
- Hybrid- und Multi-Cloud-Umgebungen

# DEMOKRATISIERUNG: AUF KI BASIERENDER ENTERPRISE DATA CATALOG

Einer der Vorteile datenbasierter Unternehmen ist es, dass alle Mitarbeiter unternehmensweit Zugriff auf Daten haben, die für die Entscheidungsfindung erforderlich sind. Es gibt jedoch einen Unterschied dazwischen, Datenzugriff zu haben und in der Lage zu sein, die Informationen zu finden, die erforderlich sind, um einen effektiven Geschäftsprozess, ein innovatives Umsatzmodell oder einen modernen Daten-Algorithmus umzusetzen. Unternehmen müssen die unternehmensweite Datennutzung unterstützen. An dieser Stelle kommt oft ein Enterprise Data Catalog zum Einsatz, der umfassende und gepflegte Metadaten umfasst, d. h. sowohl technische, betriebliche als auch verfahrenstechnische Informationen.

Business-Analysten und Datenwissenschaftler haben in datenbasierten Unternehmen die Aufgabe, innovative Erkenntnisse zu ermitteln. Business-Analysten untersuchen Möglichkeiten zur Gewinnung von Wettbewerbsvorteilen mithilfe von Ereignisdaten. Business-Analysten müssen in der Lage sein, Datensätze ordnungsgemäß zu untersuchen und zu validieren, um Produktpreise, die gezielte Kundenansprache und Betriebsabläufe zu verbessern. Doch wenn sie die Inhalte und aktualisierten Informationen in der Datenlandschaft nicht verstehen, nützt das alles gar nichts.

Genauso müssen Datenwissenschaftler die richtigen Daten für ihre Modelle nutzen, da ansonsten die Performance darunter leidet. Datenwissenschaftler bemühen sich, nur die besten Daten für Algorithmen für Advanced Analytics und Machine Learning zu verwenden. Doch wenn sie die in ihrer Datenlandschaft enthaltenen Informationen nicht verstehen, ist dies reine Zeitverschwendung, da sie zunächst einmal nach den richtigen Domäneninformationen suchen müssen.

## 2. LANGFRISTIGE VISION FÜR METADATEN-STRATEGIEN: KI/ML FÜR ERMITTLUNG UND KATALOGISIERUNG VON DATEN

### WICHTIGE ASPEKTE FÜR DEN ENTERPRISE DATA CATALOG:

- Datenzugriff für alle Nutzer erforderlich, und nicht nur für einige wenige Anwender
- Metadaten müssen neben technischen Attributen (Name, Typ, Speicherort) auch semantische Attribute (betriebliche Bedeutung und Qualität) und Betriebsdaten (Ursprung, Aktualität und Zugriff) umfassen
- Bei Katalogen ohne KI/ML sind mehr manuelle Eingriffe und Personal für die Wartung erforderlich

In der Vergangenheit gab es bereits erste Versionen des Enterprise Data Catalog. Datenarchitekten oder Datenverwalter haben diese Umgebungen manuell erstellt, doch dieser Ansatz bietet nicht mehr den Zugriff, den datenbasierte Unternehmen benötigen. Aufgrund des stetigen Wachstums der modernen Datenlandschaft benötigen Datenkataloge KI-/ML-Funktionen, um eine Skalierbarkeit über verschiedene Plattformen hinweg zu unterstützen und Daten regelmäßig zu ermitteln und zu inventarisieren. Mit all diesen Informationen können ML-Algorithmen und KI-Automatisierungsfunktionen Daten in verschiedenen Speicherorten katalogisieren und kategorisieren, ermitteln und zuordnen. Zudem können Informationen für jeden Mitarbeiter des Unternehmens zuverlässig gefunden werden. Dadurch profitieren wirklich alle Mitarbeiter von Datenzugriff und nicht nur einige wenige Anwender.



# BIG DATA, DATA LAKES, DATA SCIENCE UND MACHINE LEARNING

Seit der ersten datenbasierten Nutzung und den ersten Data-Science-Praktiken, wie Machine Learning, haben sich die Anforderungen an Umgebungen geändert, so dass herkömmliche Analytics-Umgebungen einfach nicht mehr ausreichen. Traditionelle Analytics-Umgebungen, wie Enterprise Data Warehouses und Data Marts, bieten Unternehmen auch heute noch zahlreiche Vorteile. Doch Unternehmen bekommen die Einschränkungen dieser Architekturen zu spüren, da die Anforderungen an die unternehmensweite Datennutzung und Komplexität und Geschwindigkeit von Advanced Analytics steigen. Daher benötigen diese Unternehmen flexiblere und agilere Architekturen, um diese Initiativen zu unterstützen.

Um Innovation und Agilität gerecht zu werden, müssen moderne Analytics-Architekturen und Infrastrukturen datenbasierte Strategien unterstützen. Unternehmen benötigen Umgebungen, in denen Daten von verschiedenen Speicherorten abgerufen werden können und die die unterschiedlichen Anforderungen unterstützen, die sich aufgrund von datenbasierten Initiativen ergeben.

Data Lakes in Hadoop wurden den Anforderungen datenbasierter Unternehmen hinsichtlich Erfassung und Nutzung von Daten zunächst gerecht. Sie wurden für Speicherung und Abruf neuer Datensätze verwendet, um die Anforderungen von Data-Science-Initiativen zu erfüllen. Doch aufgrund der Einschränkungen bei der Verwaltung von Metadaten, dem Datenzugriff und der Verarbeitungslatenz konnten diese Umgebungen nur von Business-Analysten und Datenwissenschaftlern mit versiertem technischem Fachwissen genutzt werden.

Da es für Unternehmen jedoch erforderlich war, dass sämtliche Nutzer Analytics der nächsten Generation nutzen können, haben sie angefangen, in ihren Data Lakes umfassende Plattformen zur Datenverwaltung umzusetzen.

**70 %** 7 VON 10 UNTERNEHMEN  
NUTZEN EINE STRATEGIE  
FÜR DATA LAKES

## WICHTIGE ASPEKTE FÜR NEXT-GENERATION ANALYTICS:

- Unternehmen können Ihre Ziele für Big Data Analytics nur erreichen, indem sie die gesamte Daten-Pipeline über ihre Datenlandschaft hinweg optimieren – beispielsweise das hybride Daten-Ökosystem
- Um die Arbeit von Datenwissenschaftlern und Business-Analysten zu unterstützen, müssen Daten-Pipelines für Next-Generation Analytics, wie z. B. Machine Learning, Daten mithilfe von Self-Service-Funktionen vorbereiten, bereinigen, integrieren und anreichern können, anstatt darauf zu warten, dass diese Aufgaben von IT-Fachkräften durchgeführt werden

Diese erweiterten Plattformen boten verschiedene Funktionen:

- Einspeisung verschiedener Datentypen
- Skalierbare Integration und Anreicherung von Daten, um Echtzeit- und Batch-Workloads zu unterstützen
- Bereitstellung von Self-Service-Funktionen für Vorbereitung, Zusammenarbeit und Pflege von Daten sowie Metadatenverwaltung
- Festlegung von Richtlinien und Compliance mit integrierten Tools für Data Governance und Datenkatalogisierung
- Bereitstellung von relevanten Daten im gesamten Unternehmen

Informationen in Anwendungen, Data Warehouses und anderen Datenlösungen wurden verwendet, um Kontext und Mehrwert für neue, ereignisbasierte Datenquellen bereitzustellen. Neue Analytics-Datenbanken mit Speicherstrukturen im Spaltenformat und NoSQL-Frameworks zur Datenspeicherung, wie MongoDB und Cassandra, führten dazu, dass diese Umgebung moderner, aber gleichzeitig auch komplexer wurde.

Das hybride Daten-Ökosystem mit auf KI/ML basierenden Datenkatalogen und Self-Service-Tools zur Datenvorbereitung ermöglicht es Datenexperten, Business-Analysten und nichttechnischen Anwendern im gesamten Unternehmen, auf einen großen Bestand verwalteter Daten zuzugreifen.

# STRATEGISCHE DATA GOVERNANCE

In der Vergangenheit wurde Data Governance oft als „Anhängsel“ zur Datenstrategie angesehen. Datenarchitekten und -verwalter wendeten taktische Data Governance jeweils nur auf eine einzige Datenquelle an. Das geschah oft manuell. Als sich die Datenlandschaft eines Unternehmens noch innerhalb der Grenzen eines On-Premise Rechenzentrums befand und vornehmlich aus traditionellen, relationalen Datenbanken bestand, war gegen diesen taktischen Basisansatz nichts einzuwenden.

Doch heutzutage ist die Datenlandschaft von Unternehmen wesentlich größer und vielfältiger und beinhaltet Plattformen und Implementierungen, die weit über die Funktionen traditioneller Praktiken hinausgehen. Neue Datenquellen, wie IoT-Gerätesensoren, Echtzeit-Unternehmensanwendungen, moderne Cloud-Anwendungen, Fusionen und Übernahmen führen dazu, dass die Anzahl und die Vielfältigkeit von Plattformen in der Datenlandschaft steigen. Veraltete Praktiken reichen einfach nicht mehr aus, damit für die Datenverwaltung zuständige Teams Daten verwalten und pflegen können.

Erschwerend kommt noch hinzu, dass Daten zu einem Unternehmens-Asset geworden sind, so dass die Unternehmensleitung an der Verwaltung beteiligt werden muss. Daten sind nicht länger nur die Summe sämtlicher Betriebsabläufe eines Unternehmens, sondern müssen heutzutage als strategisches Asset verwaltet werden. Von CIOs über CTOs bis zu CDOs (Chief Data Officer) – die Datenbestände eines Unternehmens müssen strategisch analysiert und verwaltet werden. Daten, Metadaten und Stammdaten müssen ebenfalls verwaltet werden,

# 3 von 4

**FAST 3 VON 4 UNTERNEHMEN NUTZEN STRATEGIEN FÜR DATA GOVERNANCE UND METADATENVERWALTUNG**

## WICHTIGE ASPEKTE FÜR STRATEGISCHE DATA GOVERNANCE UND METADATENVERWALTUNG:

- Data Governance und Datenverwaltung, da Daten zu den wertvollsten Assets von Unternehmen gehören, insbesondere in stark regulierten Branchen und auf globalen Märkten
- Effektive Data Governance, die Abläufe nicht verlangsamt
- Unternehmen benötigen umfassende Prozesse für Zusammenarbeit und Data Governance, so dass Richtlinien mit Abläufen verknüpft werden können

damit ihr geschäftlicher Mehrwert für datenbasierte Initiativen optimiert werden kann. Dazu ist es erforderlich, dass datenbasierte Unternehmen Data Governance strategischer angehen als zuvor. Unternehmen müssen einen ganzheitlichen Ansatz verfolgen, um Data Governance und Richtlinien festzulegen. Diese müssen dann mit Tools für die Datenkatalogisierung, Datenqualität und Datensicherheit verknüpft werden, die Data-Governance-Prozesse unterstützen, um Compliance mit diesen Richtlinien festzulegen. Wenn diese Richtlinien nicht eingehalten werden, müssen Data-Governance-Prozesse und -Tools wichtige Stakeholder darüber informieren, damit sie entsprechende Maßnahmen ergreifen können. Dies ist wichtig, um das Risiko von Non-Compliance mit gesetzlichen Vorschriften (z. B. der DSGVO) zu minimieren und den Wert der Daten zu optimieren, damit beispielsweise Kundenerlebnisse verbessert werden können.

# DATENSCHUTZ UND -SICHERHEIT

Mit Implementierung datenbasierter Initiativen und Bereitstellung von Datenzugriff für mehr Nutzer, wie Mitarbeiter, Lieferanten und Partner, werden Datenschutz und -sicherheit für Unternehmen immer wichtiger. Datenbasierte Unternehmen müssen sich der Auswirkungen des Datenschutzes und der allgemeinen Sicherheit von Unternehmensdaten bewusst sein.

Der Datenschutz ist durch die Umsetzung der DSGVO in der EU zu einem der wichtigsten Themen für Unternehmen geworden und auch die Umsetzung von EULAs (Endnutzer-Lizenzvereinbarungen) wird heutzutage genauestens untersucht. EULAs müssen beinhalten, auf welche Weise Kundeninformationen erfasst und innerhalb des Unternehmens verwendet werden. Datenschutzrichtlinien und Lizenzvereinbarungen müssen aufeinander abgestimmt werden, um Compliance sicherzustellen. Das gilt insbesondere für Branchen, in denen Kunden problemlos zu alternativen Anbietern wechseln können.

Ein weiterer, wichtiger Aspekt des Datenschutzes ist die Frage, wie Informationen abgerufen werden und von wem. Wer im Unternehmen sollte Datenzugriff haben? Wer außerhalb des Unternehmens sollte Datenzugriff haben? Externe Partner sollten nur Zugriff auf Informationen haben, die sich auf ihre Zusammenarbeit mit dem Unternehmen oder ihren Bereich der Supply Chain beziehen. Sie dürfen keinen Zugriff auf personenbezogene Kundendaten und Informationen zu den Aktivitäten anderer Partner haben.

# 83

# %

**MEHR ALS 8 VON 10  
UNTERNEHMEN NUTZEN  
STRATEGIEN ZUR  
SICHERSTELLUNG VON  
DATENSCHUTZ UND -  
SICHERHEIT, UM IHRE DATEN  
UMFASSEND ZU SCHÜTZEN**

## WICHTIGE ASPEKTE FÜR DATENSCHUTZ UND -SICHERHEIT:

- Damit Unternehmensdaten für Analytics-Projekte demokratisiert werden können und Branchenvorschriften entsprechen, müssen sie sicher sein und geschützt werden
- Nutzung risikobasierter Data Security Intelligence Tools, um herauszufinden, wo sich die sensiblen Daten eines Unternehmens befinden
- Wenn Datensicherheit mit Datenschutzrichtlinien kombiniert wird, können Risiken eingeschätzt und durch kosteneffektiven Datenschutz verringert werden (z. B. Data Masking und Verschlüsselung)
- Richtlinien müssen klar definiert und strikt eingehalten werden, indem Datennutzung, Datenqualität, Datenschutz und Compliance überprüft werden (DSGVO, HIPAA, SOX, usw.)

Datenschutz und Datensicherheit gehen Hand in Hand. Datenbasierte Unternehmen speichern riesige Mengen vertraulicher Daten, wie Zahlungsinformationen und Lieferantendaten. Daher müssen sie sicherstellen, dass ihre Daten vor Verlust, Diebstahl oder unbefugtem Datenzugriff geschützt werden. Dazu benötigen Unternehmen Einblick in die betrieblichen Aspekte der verwendeten Datenverwaltungsplattformen – in der gesamten Datenlandschaft. Sie müssen wissen, wo und wie sensible Daten gespeichert, verarbeitet und abgerufen werden, um Sicherheitsrichtlinien zu erstellen, die das Wertvollste schützen, was datenbasierte Unternehmen besitzen – Daten. Mithilfe von Tools zur Bewertung des Risikos und der Offenlegung sensibler Daten können Unternehmen ihre Daten auf intelligente Weise schützen.



# HYBRID- UND MULTI-CLOUD-BEREITSTELLUNGEN

Da Änderungen bei datenbasierten Initiativen immer schneller umgesetzt werden, müssen Unternehmen in der Lage sein, ihre Daten flexibel und effizient zu testen, zu validieren, bereitzustellen und zu operationalisieren. Cloudbasierte Ressourcen sind bei der Bereitstellung die erste Wahl und wesentlich besser geeignet als herkömmliche Implementierungspraktiken. Die Cloud ermöglicht eine wesentlich schnellere Bereitstellung. Zudem bietet sie bessere Self-Service-Funktionen für verschiedene Nutzer im gesamten Unternehmen, verringert den Zeitaufwand für die Verwaltung der Infrastruktur und bietet ein hohes Maß an Skalierbarkeit, um datenbasierte Strategien zu unterstützen. Diese Vorteile bieten herkömmliche Rechenzentren On-Premise nicht.

Darüber hinaus bietet die Cloud den enormen Vorteil, dass technische Ressourcen für Speicherung, Verarbeitung und Analyse im Handumdrehen bereitgestellt werden können – damit kann eine herkömmliche Bereitstellung in einem Rechenzentrum nicht Schritt halten. Die Zeiten, in denen IT-Teams erst nach einer sorgfältigen, detaillierten Prognose des Datenwachstums und der Datenverwendung Hardware kaufen und Software installieren konnten, sind endgültig vorbei. Heutzutage können Unternehmen einen flexiblen und agilen Prozess nutzen, um Konfigurationen zu testen, zu bewerten und zu validieren, so dass das Risiko, die falsche Architektur zu erstellen, verringert wird. Diese Art der Bereitstellung umfasst sowohl öffentliche Ressourcen, wie Amazon Web Services (AWS), als auch private Rechenzentren, in denen Unternehmen ihre sensiblen Daten in virtuellen Private Cloud-Implementierungen speichern.

# 76 %

**FAST 8 VON 10 BIG-DATA-PROJEKTEN WERDEN MITHILFE VON CLOUD-STRATEGIEN UMGESETZT, WIE BEISPIELSWEISE DIE BEREITSTELLUNG IN PRIVATE, PUBLIC ODER HYBRID-CLOUDS**

## WICHTIGE ASPEKTE FÜR DATENSCHUTZ UND -SICHERHEIT:

- Die Cloud bietet zahlreiche, überzeugende Vorteile, wie eine höhere Betriebseffizienz, Skalierbarkeit, Einfachheit, Sicherheit und Flexibilität
- In einem Bericht von 2016 heißt es, dass sich 60,1 % der Unternehmen für Big Data Analytics auf Hybrid- und Public Clouds verlassen. Initiativen für Big Data Analytics nahmen dabei einen höheren Stellenwert ein als andere Analytics-Initiativen, die laut Umfrageteilnehmer in Planung waren.<sup>1</sup>
- Sicherstellung, dass Cloud-Bereitstellungen den Vorgaben von Sicherheit, Compliance und Datenschutz gerecht werden (z. B. SOC2, SOC3, ISO, HIPAA, CSA, usw.)

Es ist für datenbasierte Unternehmen unerlässlich, die Unterstützung von Analytics-Initiativen zu vereinfachen. Wenn Administrationsteams sich um die Bereitstellung von Systemen und Plattformen sowie die Wartung nicht mehr kümmern müssen, können sie sich ganz auf eine effektive Skalierung konzentrieren. Da Datenverwaltungsteams sich nicht mehr mit taktischen und anderen, für die Bereitstellung erforderlichen Aktivitäten befassen müssen, können sie sich darauf konzentrieren, neue Datensätze und bessere Informationen in strategischer Weise bereitzustellen. Dank cloudbasierter Bereitstellungen gelingt der Sprung von Test- und Validierungs-Sandboxes zur Produktionsumgebung schnell.

Doch diese Umgebungen halten auch verschiedene Herausforderungen bereit. Durch verschiedene, cloudbasierte Implementierungen werden Daten und Datenbestände über eine wesentlich größere Datenlandschaft verbreitet. Mit herkömmlichen Praktiken lässt sich diese Ausbreitung an Daten kaum verwalten und pflegen. Doch mithilfe einer einheitlichen und modernen Hybrid-Integrationsplattform ist es möglich, diese Ausbreitung außerhalb des Rechenzentrums effektiv zu verwalten.

<sup>1</sup> EMA, „The State of Cloud Analytics“, 2016.

# ÜBER UNSEREN SPONSOR: INFORMATICA

Informatica ist der führende Anbieter im Bereich Enterprise Cloud Data Management und treibt die datengestützte, digitale Transformation weiter voran. Informatica ermöglicht es Unternehmen, das Potenzial ihrer Daten voll auszuschöpfen, um agiler zu werden und neue Wachstumschancen zu nutzen und Innovationen voranzutreiben, um sich dadurch langfristige Wettbewerbsvorteile zu verschaffen. In den letzten 25 Jahren hat Informatica mehr als 9.000 Kunden dabei geholfen, das Potenzial ihrer Daten voll auszuschöpfen. Weiterführende Informationen finden Sie unter [www.informatica.com/de](http://www.informatica.com/de).

## Intelligent Data Platform

Die Informatica Intelligent Data Platform ist eine modulare, auf KI und Machine Learning basierende Lösung zur Datenverwaltung, die auf einer modernen Microservices-Architektur beruht, damit Unternehmen ihre Daten über die hybride Unternehmensumgebung hinweg bestmöglich nutzen können. Die auf KI-basierende Plattform umfasst Daten in On-Premise-, Cloud- und Big-Data-Umgebungen, so dass Sie sich immer darauf verlassen können, dass Daten zuverlässig, sicher, abrufbar, aktuell, relevant und verwertbar sind. So können Unternehmen die datenbasierte, digitale Transformation aktiv vorantreiben.

## CLAIRE™

Die CLAIRE Engine stellt Enterprise Unified Metadata Intelligence zur Verfügung, um die Produktivität der Informatica Intelligent Data Platform zu erhöhen. CLAIRE erfasst umfassende Metadaten aus dem gesamten Unternehmen, darunter auch technische, betriebliche, operationelle und nutzungsbasierte Metadaten, die dann als Trainingsdatensätze verwendet werden können. CLAIRE wendet Machine-Learning-Algorithmen und -Methoden auf diese umfassenden Metadaten an, so dass Unternehmen KI aktiv für die Datenverwaltung nutzen können. Da die Metadaten, die der Plattform von Informatica zur Verfügung stehen, sehr umfassend und vielfältig sind, ist CLAIRE in der Lage, Entwickler, Business User, Datenexperten und -verwalter mit intelligenten Empfehlungen und Automatisierung bei der Datenverwaltung zu unterstützen.



# Informatica™

CLAIRE beschleunigt die Datenverwaltung für die Informatica Intelligent Data Platform erheblich. Im Folgenden werden einige Vorteile aufgeführt, die Unternehmen durch CLAIRE entstehen:

- Datenverwaltung: CLAIRE erkennt Domänen und Strukturen neuer, unstrukturierter Daten auf intelligente Weise.
- Produktivitätssteigerung und Self-Service-Funktionen für Business User: CLAIRE erstellt intelligente Empfehlungen für weitere, relevante Datensätze, die eventuell ebenfalls von Interesse sind – basierend auf der aktuellen Arbeit der Business User.
- Administration und Betrieb: CLAIRE erstellt Vorhersagen von Systemlasten, Cloud Bursting oder Self-Tuning, so dass selbst komplexe Aufgaben der Datenverwaltung problemlos ausgeführt werden können.



## Enterprise Management Associates, Inc.

Enterprise Management Associates (EMA) wurde 1996 gegründet und ist ein führendes Branchenanalyse-Unternehmen, das sich darauf spezialisiert hat, „hinter die Kulissen zu schauen“, um fundierte Erkenntnisse zum gesamten Spektrum an IT- und Datenverwaltungstechnologien zutage zu fördern. Dank praktischer Erfahrung, Einblick in Best Practices der Branche und umfassendes Fachwissen aktueller und geplanter Lösungen können Analysten den Kunden von EMA helfen, ihre Ziele zu erreichen. Auf [www.enterprisemanagement.com](http://www.enterprisemanagement.com) oder [blog.enterprisemanagement.com](http://blog.enterprisemanagement.com) erfahren Sie mehr über die Forschung, Analysen und Beratungsdienste von EMA für Business User, IT-Experten und IT-Anbieter. Sie können EMA auch auf [Twitter](#), [Facebook](#) oder [LinkedIn](#) folgen.

---

Dieser Bericht darf ohne die vorherige, schriftliche Erlaubnis von Enterprise Management Associates, Inc. weder teilweise noch vollständig dupliziert, reproduziert oder in einem Abfragesystem gespeichert oder weitergeleitet werden. Alle in diesem Bericht geäußerten Meinungen und Einschätzungen stellen unser Urteil zu diesem Datum dar und können ohne Vorankündigung geändert werden. Bei den in diesem Bericht erwähnten Produktnamen kann es sich um Marken und/oder eingetragene Warenzeichen der jeweiligen Unternehmen handeln. „EMA“ und „Enterprise Management Associates“ sind eingetragene Warenzeichen von Enterprise Management Associates, Inc. in den USA und anderen Ländern.

©2018 Enterprise Management Associates, Inc. Alle Rechte vorbehalten. EMA™, ENTERPRISE MANAGEMENT ASSOCIATES® und das Mobius-Symbol sind eingetragene Warenzeichen oder handelsübliche Marken von Enterprise Management Associates, Inc.

### Hauptsitz:

1995 North 57th Court, Suite 120

Boulder, CO 80301, USA

Tel.: +1 303.543.9500

Fax: +1 303.543.7687

[www.enterprisemanagement.com](http://www.enterprisemanagement.com)

3773.100318