

Die 5 Mythen über den Data Lake

Sobald Unternehmen den Nutzen und die Relevanz ihrer Daten für neue geschäftliche Chancen und Umsatzmodelle erkennen, beschäftigt sie eine zentrale Frage: Welche der unendlich vielen Technologien ist am besten geeignet, um alle Daten zu verwalten und wertvolle Erkenntnisse zu gewinnen?

Der Data Lake ist eine Referenzarchitektur, die momentan zu Recht im Fokus steht. Doch es gibt eine Reihe von Mythen über den Data Lake, die Unternehmen und IT-Manager wertvolle Zeit bei der Suche nach der richtigen Lösung kosten.

Die 5 Mythen über den Data Lake

1 Der Data Lake ist ein Produkt, das sie kaufen können → Falsch

Der Data Lake ist kein Produkt, das man einfach kaufen kann. Setzen Sie Hadoop oder eine Data Warehouse-Lösung deshalb nicht mit einem Data Lake gleich.

Ein Data Lake ist eine Referenzarchitektur, die unabhängig von einer bestimmten Technologie ist. Mit diesem Ansatz rücken Unternehmen Daten ins Zentrum ihrer betrieblichen Abläufe rund um Governance, Qualität und Datenmanagement. Die Daten bilden die Basis für Self-Service Analytics, von der alle Datenkonsumenten profitieren.

2 Hadoop ist der einzige Data Lake → Falsch

Auch wenn der Begriff „Data Lake“ häufig mit Hadoop oder Hadoop-orientiertem Objektspeicher in Verbindung gebracht wird, kann ein Data Lake durchaus ohne Hadoop entwickelt und effizient eingesetzt werden. Noch effizienter ist ein Data Lake, der mit mehreren Managementsystemen für relationale Datenbanken verbunden ist.

Ein Data Lake kombiniert vielfältige Technologien in einem „System of Insight“, damit Data Scientists Geschäftsanforderungen mithilfe agiler Datenexploration umsetzen können.

3 Der Data Lake generiert einen Speicherauszug der Daten – ganz ohne Governance → Falsch

Software und Hardware sind die Schlüsselkomponenten einer Data Lake-Lösung, aber auch die Katalogisierung, Qualität, Governance und Verwaltung von Daten spielt eine wichtige Rolle.

Genau so wie Unmengen an Daten in manchen Data Warehouses in einem „schwarzen Loch“ verschwinden, kann auch ein Data Lake eine Flut von Daten produzieren, wenn wirksame Governance-Richtlinien fehlen.

In einem digitalen Unternehmen ist es wichtig, dass Daten im Data Lake katalogisiert, zugänglich, vertrauenswürdig und verwertbar sind. Deshalb sind Governance, Qualität und Informationsmanagement unverzichtbare Säulen eines Data Lakes.

4 Der Datenzugriff ist die wichtigste Kennzahl für den Erfolg eines Data Lakes → Falsch

Nur Daten an einem zentralen Ort zu sammeln, ist von einer echten Analyselösung sehr weit entfernt. Das Ziel sind Datenanalysen, die aussagekräftige Geschäftseinblicke für neue Umsatzströme, Kundenbindungsmodelle oder Produkterweiterungen liefern.

Aber die Daten müssen vertrauenswürdig, relevant und für alle Datenkonsumenten verfügbar sein. Ein Data Lake setzt einen intelligenten Metadatenkatalog voraus, der mit Geschäftsbegriffen verknüpft ist. So werden auch kryptische Daten ausgewertet und in einem verständlicheren Kontext präsentiert. Darüber hinaus berücksichtigt der Data Lake die Quelle und Qualität strukturierter wie unstrukturierter Datenbestände sowie die Governance-Fabric. Auf diese Weise verfügen Unternehmen über geschützte, standardisierte, effizient verwaltete und vertrauenswürdige Daten.

5 Der Data Lake ersetzt das Data Warehouse → Falsch

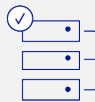
Der Data Lake kann mehrere Enterprise Data Warehouses (EDW) sowie weitere Datenquellen wie Social Media oder IoT-Daten zusammenführen. Alle diese Informationen fließen in den Data Lake, der mit Governance kombiniert werden kann, um die vertrauenswürdige Erkennung von Daten unternehmensweit zu vereinfachen.

Der Data Lake ist eine Ergänzung zu Ihrer EDW-Umgebung. Mit dieser Lösung können Data Scientists und Business Analysten Daten und weitere Assets schnell durchsuchen, um neue Einblicke zu gewinnen und Innovation und Wachstum voranzubringen.


Schluss mit den Mythen über den Data Lake

Räumen Sie auf mit den Mythen rund um den Data Lake. Kombiniert mit Governance bietet Ihnen der Data Lake Zugang zu komplexen strukturierten und unstrukturierten Daten aus dem ganzen Unternehmen sowie sichere und vertrauenswürdige Daten.


Passen Sie eine kontrollierte Data Lake-Lösung an Ihre Geschäftsanforderungen an, um schnellere Analytics-Ergebnisse und präzisere Einblicke zu erhalten:




Die Daten sind durch eine sichere und vertrauenswürdige Infrastrukturbasis geschützt.




Daten werden in Geschäftsterminologie ausgedrückt. Data Scientists können Daten also sofort verwerten, ohne über die Bedeutung kryptischer Begriffe zu grübeln.



Zuverlässige Daten aus kontrollierten Datenfeeds fließen in den Data Lake, wo Informationsassets, Metadaten und Geschäftskontext dokumentiert werden. Neue Daten kommen in Echtzeit hinzu.



Die Daten sind streng klassifiziert, geschützt und kontrolliert.



Qualität, Ursprung und Herkunft der Daten werden genau analysiert.

Führende Unternehmen wissen bereits, welche Mythen und Fakten wirklich stimmen und wie sie mit einem kontrollierten Data Lake einen echten geschäftlichen Nutzen aus ihrer Datenumgebung ziehen.

Überzeugen Sie sich von den Fakten

Mit dem IBM Unified Governance and Integration-Portfolio umgehen unsere Kunden gängige Probleme und Mythen und meistern alle kritischen Phasen auf dem Weg zu einem kontrollierten Data Lake.

Video ansehen 

Mit einem kontrollierten Data Lake hat jeder Mitarbeiter im Unternehmen Zugang zu vertrauenswürdigen Qualitätsdaten.

E-Book lesen 

Lernen Sie die wesentlichen Bausteine eines kontrollierten Data Lakes kennen.

Erfahren Sie mehr 

Besuchen Sie ibm.com/unified-governance-integration.



© Copyright IBM Corporation 2018. IBM, das IBM Logo und ibm.com sind eingetragene Marken der International Business Machines 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. Dieses Dokument ist zum Datum seiner Erstveröffentlichung aktuell und kann jederzeit von IBM geändert werden. Nicht alle IBM Angebote sind in jedem Land, in welchem IBM tätig ist, verfügbar.

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

Die Informationen in diesem Dokument werden auf der Grundlage des gegenwärtigen Zustands (auf „as-is“-Basis) ohne jegliche ausdrückliche oder stillschweigende Gewährleistung zur Verfügung gestellt, einschließlich, aber nicht beschränkt auf die Gewährleistungen für die Handelsüblichkeit, die Verwendungsfähigkeit für einen bestimmten Zweck oder die Freiheit von Rechten Dritter. Für IBM Produkte gelten die Gewährleistungen, die in den Vereinbarungen vorgesehen sind, unter denen sie erworben werden.

Hinweis: Jeder Kunde ist für die Einhaltung der geltenden Gesetze und Verordnungen, einschließlich der Datenschutz-Grundverordnung der Europäischen Union, selbst verantwortlich. Es obliegt allein den Kunden, sich von kompetenter juristischer Stelle zu Inhalt und Auslegung aller relevanten Gesetze und gesetzlichen Bestimmungen beraten zu lassen, die ihre Geschäftstätigkeit und die von ihnen eventuell einzuleitenden Maßnahmen zur Einhaltung dieser Gesetze und Bestimmungen betreffen. Die in diesem Dokument beschriebenen Produkte, Services und sonstigen Funktionen eignen sich nicht für alle Kundensituationen und sind möglicherweise nur eingeschränkt verfügbar. IBM erteilt keine Rechts- oder Steuerberatung und gibt keine Garantie bezüglich der Konformität von IBM Produkten oder Services mit den geltenden Gesetzen und gesetzlichen Bestimmungen.