

Arvato Systems Whitepaper

# Modern Data Management

## Die Schlüsselkompetenz zukunftsfähiger Unternehmen

**arvato**  
BERTELSMANN  
Arvato Systems

## Inhaltsverzeichnis

<b>1. Modern Data Management</b> Die Schlüsselkompetenz zukunftsfähiger Unternehmen	3
<b>2. Data Leveraging</b>	4-5
<b>3. Status Quo in deutschen Unternehmen</b>	6
<b>4. Die Nutzung der Cloud</b> Als Game Changer im modernen Datenmanagement	7-8
<b>5. Integration ist der Schlüssel zum Erfolg</b>	8-9
<b>6. Fazit</b>	11

# 1. MODERN DATA MANAGEMENT

## DIE SCHLÜSSELKOMPETENZ ZUKUNFTSFÄHIGER UNTERNEHMEN

Die Temperatur von Materie, die Geschwindigkeit von Objekten oder der CO<sub>2</sub>-Gehalt der Atmosphäre auf unserem Planeten – Daten gab und gibt es schon immer. Aber erst mit der Aufzeichnung durch Schriftsysteme konnten Daten auch dargestellt, abgespeichert und analysiert werden. Bis weit ins 20. Jahrhundert hinein erfolgten die Datenauswertungsprozesse analog. Dazu gab es Hilfsmittel wie Akten, Dossiers und riesige Archive, in denen die Daten in Papierform aufbewahrt werden konnten.

Mit der Digitalisierung veränderte sich die Sichtweise auf und der Umgang mit Daten grundlegend. Nicht nur, dass Daten seit dieser Zeit digital erfasst, abgespeichert und analysiert werden können. Auch die schiere Anzahl der Daten ist gleichzeitig explodiert. Aktuelle Schätzungen gehen davon aus, dass bis 2025 die weltweite Datenmenge auf 175 Zetabyte anwachsen wird – eine Zahl mit 21 Nullen.<sup>1</sup>

80 % dieses Datenvolumens entfallen auf weltweit agierende Unternehmen.

Egal ob Industrieunternehmen, Opernhaus oder StartUp:

In der Zukunft gilt, dass eine zeitgemäße Aufbereitung von Daten zentral für den Erfolg der Unternehmen sein wird. Damit wird das moderne Datenmanagement zur Schlüsselkompetenz für alle Unternehmen.

Entlang der Daten-Touchpoints von Unternehmen gibt es vier zentrale Datenkategorien:

1. Produktdaten
2. Kundendaten
3. Prozess- und Ablaufdaten
4. Mitarbeiterdaten

Ihre eigentliche Stärke entfalten diese vier unterschiedlichen Datenkategorien aber erst, wenn sie gemeinsam genutzt werden und daraus neue Erkenntnisse für die Ausrichtung der Unternehmen in der Zukunft gezogen werden können. Je zeitnaher dies geschieht, desto größer fällt der gewonnene Wettbewerbsvorteil aus. Damit wird das Thema „Data Leveraging“ zur zentralen Zukunftsaufgabe für die strategische Ausrichtung von Unternehmen.

<sup>1</sup> Quelle: <https://blog.wiwo.de/look-at-it/2018/11/27/weltweite-datenmengen-sollen-bis-2025-auf-175-zetabyte-wachsen-8-mal-so-viel-wie-2017/>

# 2. DATA LEVERAGING

Der Begriff Data Leveraging bezeichnet die Fähigkeit eines Unternehmens, die eigenen Daten als Hebel für die eigenen Geschäftsmodelle zu nutzen. Dabei stehen den Unternehmen zwei Optionen offen, wie sie die gewonnenen Daten gewinnbringend in die eigenen Geschäftsmodelle integrieren können: Sie können entweder bestehende Geschäftsmodelle verbessern oder neue Geschäftsmodelle entwickeln.

Ein Beispiel für die Verbesserung und Neuentwicklung von Geschäftsmodelle bieten die sogenannten Digital Twins. Der Begriff des digitalen Zwillings steht für die digitale Nachbildung realer Produktionsprozesse eines Unternehmens in Echtzeit. So können ganze Fertigungsanlagen digital abgebildet werden und der Prozess der Produktion dargestellt und analysiert werden.

Der öffentliche Personennahverkehr liefert ein Beispiel für den Einsatz solcher Digital Twins. Während noch vor wenigen Jahren die wenigsten Busse in Deutschland eine digitale Spur hinter sich herzogen, ist die Anzahl der übermittelten Daten inzwischen sprunghaft gestiegen. Anhand von Datenverbindungen und GPS-Systemen kann der genaue Standpunkt jedes einzelnen Busses einer städtischen Flotte erfasst werden.

Aber nicht nur das: Inzwischen werden die Ticketautomaten, Reifen und Motoren der Busse mit Sensoren ausgestattet, die einen Informationsaustausch zwischen dem Fahrzeug und den Datenspeicherungssystemen ermöglichen.

So lassen sich in Zukunft die Effizienz von Linienplänen optimieren aber auch mögliche Gefahren minimieren. Verbunden mit den Kundendaten, die beispielsweise von den mobilen Endgeräten der Nutzer stammen, können so auch neue Geschäftsmodelle entstehen. Längst sind beispielsweise in südostasiatischen Ländern Systeme im Einsatz, die via App das Einsteigen und Aussteigen eines Kunden aus dem Bus erkennen und anhand dessen den für ihn richtigen Tarif auswählen und direkt nach Abschluss der Fahrt abbuchen.

Das Prinzip hinter diesen Datenauswertungsstrategien wird als sogenannter Feedback Loop bezeichnet (siehe Abbildung 1). Dabei werden alle Informationen entlang der Wertschöpfungskette gesammelt ausgewertet und in die neue Erstellung des Produkts eingefügt. Im Zusammenspiel zwischen Kunden-, Prozess-, Produkt- und Mitarbeiterdaten können bisher unerkannte Zusammenhänge aufgedeckt und neue Lösungsansätze für Unternehmen entwickelt werden.<sup>2</sup>

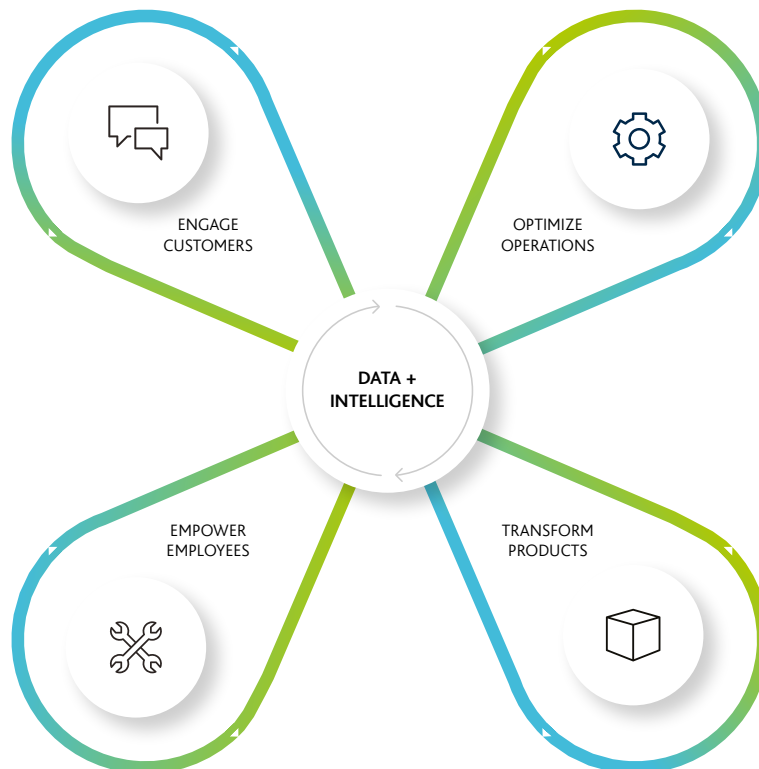


Abbildung 1: Der digitale Feedback Loop, Quelle: <https://www.hwzdigital.ch/files/2019/11/Digital-Feedback-Loop.jpg>

<sup>2</sup> Quelle: <https://www.inovex.de/blog/feedback-loop/>



Sogenannte Vendor Management Inventories gewähren beispielsweise Lieferanten einen Datenzugriff auf die Bestands- und Lagersituation von Produzenten und Händlern. Dadurch kann der Prozess der Fertigung und Bestellung von Vorprodukten besser gesteuert werden. Werden diese Datenbestände mit den Kundendaten am Point-of-Sale abgeglichen, können Konsummuster der Kunden in die Produktion der Vorprodukte mit einfließen und die Effizienz der Produktionsplanung deutlich gesteigert werden.<sup>3</sup>

**Diejenigen Unternehmen, die in der Lage sind, Daten besser zu generieren, zu speichern und auszuwerten, können Produkte und Dienstleistungen schneller, kostengünstiger und kundennäher an den Markt bringen. Unternehmen, die bei der Datenauswertung zurückbleiben, werden über kurz oder lang an Bedeutung verlieren.**

Vielen Anbietern fehlt schon allein der Zugriff auf die relevanten Datensätze. Buchverlage bedienen mit ihren Produkten beispielsweise einen klassischen Endkundenmarkt. Die Datensätze der Kundendaten liegen aber in der Regel nicht bei ihnen, sondern bei Buchhändlern und den großen E-Commerce-Anbietern. Daher besteht eine zentrale Aufgabe der Buchverlage in den kommenden Jahren darin, die entsprechenden Daten-Touchpoints zu den Endkunden bereitzustellen, um so von der Hebelwirkung der Daten auf das eigene Geschäftsmodell zu profitieren. Gleichzeitig ist den Unternehmen, die anhand eines digitalen Feedback Loops die eigenen Produkte verbessern wollen, gar nicht klar, wozu die abgespeicherten Daten von heute in der Zukunft verwendet werden können. Unternehmen stehen daher vor der Aufgabe, sich permanent auf einen wechselnden Datenbedarf bei wechselnden Fragestellungen einzustellen. Die Kundendaten, aber auch die Produktdaten, die Prozess- und Mitarbeiterdaten von heute können die Produktentwicklung von morgen dramatisch verbessern.

Die deutschen Unternehmen hinken in der Disziplin „digitaler Feedback Loop“ teilweise weit hinterher. Zu spät haben die Entscheidungsträger die zentrale Bedeutung des Datenmanagements erkannt. Um diesen Rückstand aufzuholen, ergeben sich ganz neue Herausforderungen: Wie können Unternehmen die zunehmende Datenflut beherrschen? Und wie können die gespeicherten Daten dazu beitragen, die Geschäftsmodelle der Unternehmen zu verbessern?

## → Die zentralen Trends der Datenverarbeitung

- Endkunden-, Prozess-, Mitarbeiter- und Produktdaten bilden das Fundament für gewinnbringende digitale Geschäftsmodelle.
- Daten besitzen daher einen Wert für das Unternehmen, der eventuell erst in den kommenden Jahren zum Tragen kommt.
- Daten entstehen nicht automatisch dort, wo der Umsatz entsteht.
- Da sie eine Ressource darstellen, sollten daten-kreierende Wertschöpfungsprozesse im Unternehmen unterstützt bzw. subventioniert werden.

<sup>3</sup> Quelle: <https://www.cegeka.com/de-at/blog/digital-feedback-loops-das-neue-sahnehaebchen>

# 3. STATUS QUO IN DEUTSCHEN UNTERNEHMEN

Der Status Quo der Datenverarbeitung in vielen deutschen Unternehmen ist weit davon entfernt, den Ansprüchen an ein modernes Datenmanagement-System zu genügen. Vielmehr sind in vielen Unternehmen noch überholte Technologien und Prozesse im Einsatz. Interne IT-Abteilungen verfahren nach dem Motto: „Never change a running system“. Ende der 1990er und Anfang der 2000er Jahre sind viele Unternehmen im Bereich der Datenverarbeitung auf die sogenannten Data-Warehouses umgestiegen. Der Begriff Data Warehouse steht dabei für zentral organisierte Datenbanksysteme, die es dem Unternehmen ermöglichen, neue Erkenntnisse aus den gespeicherten Datensätze zu gewinnen. Zu diesem Zweck werden beispielsweise OLTP, LOB, ERP und CRM-Systeme zum sogenannten „Enterprise Data Warehouse“ zusammengefasst. Die Auswertung dieser Daten erfolgt durch den Einsatz von Spezialisten. Diese ausgebildeten Datenbank-Entwickler sind in der Regel dafür verantwortlich, die Schnittstellen für die Aufbereitung der Daten für das Management zu programmieren („Business Intelligence“).<sup>4</sup>

Das grundlegende Problem dieser Systeme besteht darin, dass von den einströmenden Daten nur einige wenige Attribute verwendet und in das Data-Warehouse integriert werden. Der Rest der Informationen, die in den einströmenden Daten vorhanden sind, wird verworfen. Dadurch entstehen Datenbank-Systeme, die zumeist auf Tabellen basieren und die von der Datenstruktur her flach beziehungsweise zweidimensional aufgebaut sind.

Mit Blick auf Performance und Speicherbedarf ist dieses Vorgehen nachvollziehbar, richtig und auch zukünftig noch absolut relevant. Jedoch gehen viele Einsichtsmöglichkeiten und potenziell wichtige Details der tiefsten Datenebenen verloren. Damit können sie alleinstehend nicht mehr als Lösung gesehen werden.

Bildlich gesprochen lässt sich der Vorgang der Datenspeicherung in den klassischen Data-Warehouse-Systemen mit der Einlagerung von Wasserkisten vergleichen. Zunächst wird das einströmende Datenmaterial portioniert und normiert. Die Daten werden einzeln verpackt und gelabelt. Anschließend werden die Daten in die dafür vorgesehenen Behälter eingelagert. Die Datenbanktabellen entsprechen in dieser Analogie den Wasserkästen, die Flasche für Flasche befüllt werden.

Bevor die Daten aber analysiert werden können, müssen sie „an den richtigen Platz“ gebracht und gelagert werden. Möchte die IT eines Unternehmens sicherstellen, dass ein gespeicherter Datensatz an zwei Stellen im Unternehmen verwendet werden kann, so muss er in der Regel auch doppelt vorgehalten werden.

Möchte das Management auf die Daten zugreifen, dann ist der Einsatz von Spezialisten notwendig. Diese müssen die vorhandenen Daten so aufbereiten, dass andere Abteilungen im Unternehmen darauf zugreifen können. Ändert sich etwas am Datenbedarf müssen die Spezialisten immer wieder neue Schnittstellen zu den Systemen programmieren. Bildlich gesprochen, wird jede Flasche oder jeder Kasten mühsam einzeln aus dem Keller geholt.

<sup>4</sup> Quelle: <https://www.bigdata-insider.de/was-ist-ein-data-warehouse-a-606701/>

# 4. DIE NUTZUNG DER CLOUD

## GAME CHANGER IM MODERNEN DATENMANAGEMENT

**Da die Cloud zu einem rasanten Preisverfall bei Speicher und Rechenkapazitäten geführt hat, ergeben sich Möglichkeiten des modernen Datenmanagements, die noch vor wenigen Jahren weder technisch noch ökonomisch möglich waren.**

Mit dem Umstieg der Datenverarbeitung in die Cloud stehen Unternehmen eine Reihe von neuartigen Tools zur Bewältigung der Herausforderungen des modernen Datenmanagements zur Verfügung.

Ein fundamentales Tool ist der sogenannte Data Lake. Data Lakes unterscheiden sich von klassischen Data Warehouse-Systemen darin, dass alle einströmenden Daten in ihrer ursprünglichen Form – also als Rohdaten – gespeichert werden. Dabei spielen die Formate der einströmenden Daten keine Rolle und ein direkter Zugriff ist konstant möglich. Während bei traditionellen Data-Warehouse-Systemen weiterhin die Daten vorklassifiziert und einzeln abgespeichert werden, gehen im Data-Lake-Systemen keine Information mehr verloren.

### —> Unterschiede zwischen Data Lake und Data Warehouse

Ein Data Lake ist ein großer Pool mit Rohdaten, für die noch keine Verwendung festgelegt wurde. Bei einem Data Warehouse dagegen handelt es sich um ein Repository für strukturierte, gefilterte Daten, die bereits für einen bestimmten Zweck verarbeitet worden sind.

Diese Vorgehensweise erweitert enorm den Spielraum der Datenauswertung. Denn viele Unternehmen können heute noch gar nicht abschätzen, wann und wozu sie die gespeicherten Daten in der Zukunft gebrauchen können. In der Sprache des Sports sind Data-Lake-Systeme damit ein Game Changer, der die Art und Weise der modernen Datenspeicherung neu definiert.

Im Gegensatz zu Data-Warehouse-Systemen kann (die Zugriffserlaubnis vorausgesetzt) jeder Mitarbeiter permanent und in Echtzeit auf alle vom Unternehmen gespeicherten Daten zugreifen. Dies führt zu einer Demokratisierung der Datenverarbeitung. Aus zentral gesteuerten Datenabfragen, die sich auf die unterschiedlichen Datensilos im Unternehmen verteilen, werden dezentrale Zugriffsmöglichkeiten. Diese Demokratisierung und Dezentralisierung sind die Voraussetzung für das erfolgreiche Arbeiten in agilen Teams und ermöglicht es Fachabteilungen, schneller und gezielter Entscheidungen zu treffen.

## 5. INTEGRATION IST DER SCHLÜSSEL ZUM ERFOLG

Der Übergang von der Nutzung zentral organisierter Data-Warehouse-Systeme zu dezentral organisierten Data-Lakes gleicht damit dem Übergang von zentral organisierter Rechenleistung in Unternehmen in Form von On-Premises-Servern hin zur Bereitstellung von Rechenleistung an jedem Arbeitsplatz. Also einem Wandel, der vor ca. 30 Jahren nahezu jeden Arbeitsplatz der Welt mit einem PC ausstattete. Bildlich gesprochen funktionieren Data-Lakes wie kleine Wasserkraftwerke, die jede Abteilung einzeln mit dem gesamten „Datensee“ des Unternehmens verbinden, statt die Spezialisten einzeln zum Datenholen in den Keller zu schicken.

Möglich machen dies die riesigen und extrem kosteneffiziente Speicherkapazitäten, die Data-Lake-Anbieter im Netz für ihre Kunden bereitstellen. Besonders die großen Anbieter von Cloud-Diensten bieten darüber hinaus ein breites Ökosystem an Analyse-Diensten, mit denen Unternehmen die klassischen Anwendungen aber auch state-of-the-art Anforderungen abdecken können. Die Breite und die Tiefe dieser einfach zugänglichen Dienste eröffnen Unternehmen komplett neue Möglichkeiten.

Nur, wenn mit dem Data Lake auch eine automatisierte funktionierende Management-Plattform verbunden ist, kann ein erfolgreiches und sicheres Data-Management funktionieren. Erfolgreiches und sicheres Data-Management kann dabei nur funktionieren, wenn eine automatisierte Managementplattform angebunden wird. Technisch realisiert wird eine solche Datenaufbewahrung durch ein präzise ausgestaltetes System, an Speicher- und Analysetools, die bei richtiger Anwendung einen Verlust von Daten ausschließen.

# 5. INTEGRATION IST DER SCHLÜSSEL ZUM ERFOLG

Viele Unternehmen stehen bei der Umstellung vom On-Premises Datenmanagement auf Data-Management-Systeme in der Cloud vor großen Herausforderungen. Auch wenn die Bedienung eines modernen Data-Management-Systems einfach zu erlernen ist, so müssen doch die entsprechenden Mitarbeiter und Abteilungen für den Umgang mit der neuen Technologie geschult werden.

Die gute Nachricht dabei ist, dass die alten Systeme – d.h. die seit Jahren laufenden Data-Warehouses – nicht einfach abgeschaltet werden müssen, sondern parallel aber günstiger zu den neu aufzubauenden Data-Management-Strukturen betrieben werden können. Dadurch reduziert sich das Risiko, dass die alten Systeme abgeschaltet werden und anschließend nichts mehr funktioniert. Ein solches Szenario wäre für viele Unternehmen der Daten-Super-GAU.





Das Zauberwort für die Umsetzung einer solchen parallelen Datenstruktur lautet Integration. Data-Management-Plattformen in der Cloud sind durch ihren vergleichsweise simplen Aufbau so breit einsetzbar, dass eine schier unendliche Zahl von Anwendungsszenarien umgesetzt werden kann. So können Data-Lakes als Teil einer bereits bestehenden Data-Warehouse-Struktur installiert werden. Aber auch das Gegenteil ist denkbar und die bisherigen Data-Warehouse-Strukturen können als Teil des Data-Lakes des Unternehmens aufgesetzt werden. Mit Hilfe von solchen Lösungsansätzen entstehen neue Daten-Ökosysteme, die das Beste aus den beiden Welten miteinander vereinen: Sie bieten den Unternehmen die Sicherheit, dass alle Prozesse auch weiterhin so laufen wie bisher. Gleichzeitig können die Unternehmen mit Hilfe der neuen Datenarchitektur auf die Speicher- und Analyse Vorteile zurückgreifen, die sich aus der Speicherung der Daten in der Cloud ergeben.





# 6. FAZIT

---

Unternehmen, die auch in der Zukunft erfolgreich am Markt bestehen wollen, können sich den Anforderungen, die ein modernes Data-Management an die Aufnahme, Speicherung und Analyse der Daten stellt, nicht entziehen. Immer schneller wuchsen in den vergangenen Jahren die Datenmengen, die auf die Unternehmen zurollen. Und diese Entwicklung steht mit Blick auf 5G und IoT erst am Anfang. Gleichzeitig betrifft die Umstellung der Datenverarbeitung eines Unternehmens die Herzkammer eines jeden Unternehmens. Und ein ungeplanter Schuss ins Blaue könnte den Fortbestand des Unternehmens gefährden. Der Schritt in die neue Datenwelt muss daher gut geplant werden. Die neuen Datenstrukturen sollten gleichermaßen den Anforderungen an Verfügbarkeit, Skalierbarkeit und Sicherheit der Daten entsprechen.

Zudem sollte beim Aufsetzen der neuen Datenarchitekturen darauf geachtet werden, dass sie eine evolutionäre Anpassungsfähigkeit der Systeme in der Zukunft ermöglichen und dass technische Weiterentwicklungen ohne großen Aufwand eingebunden werden können. Das Ziel einer solchen Anpassungsmaßnahme könnte daher lauten: Schaffung einer strategischen Lage, die es ermöglicht, die Datenspeichersysteme an alle zukünftigen Bedürfnisse anpassen zu können.

Die Bewältigung dieser Aufgaben setzt ein großes Maß an Verständnis für die Probleme voraus, die auf dem Weg der Zielerreichung auftreten können. Das Management des gesamten Umstellungsprozesses sollte bei den Bedürfnissen des Unternehmens beginnen und daraus die Ziele für die Datentransformation festlegen.

Erst dann kann mit einer erfolgreichen Planung, Umsetzung und Integration der neuen Datenstrukturen begonnen werden. Letztlich wird sich der Mut, den Weg in die neue Cloud-gestützte Datenwelt zu beschreiten, in den meisten Fällen bezahlt machen. Alleine schon dadurch, dass das Überleben des eigenen Unternehmens in hochkompetitiven Märkten gesichert wird.

---





**arvato**  
**BERTELSMANN**  
Arvato Systems

Arvato Systems GmbH, An der Autobahn 200, D-33333 Gütersloh  
cloud@bertelsmann.de | arvato-systems.de