

Datenplattform: Was sie ist, welche Bausteine sie hat und wie du die richtige für dein Unternehmen findest

Ohne Daten keine fundierten Entscheidungen, keine Automatisierung, keine KI. Was vor zehn Jahren noch Strategiethema war, ist heute Realität in jedem Unternehmen, das wachsen will. Die deutsche Wirtschaft hat diesen Schritt längst gemacht: Laut dem [Bitkom Cloud Report 2025](#) nutzen 90 Prozent der deutschen Unternehmen ab 20 Beschäftigten Cloud-Dienste, und rund die Hälfte aller IT-Anwendungen läuft inzwischen in der Cloud.

Doch zwischen "wir nutzen Cloud-Tools" und "wir haben eine strategische Datenplattform" liegen Welten. Viele Mittelständler:innen stehen vor der gleichen Frage: Was brauchen wir wirklich, um aus unseren Daten Wert zu schöpfen? Dieser Artikel klärt, was eine Datenplattform ist, aus welchen Bausteinen sie besteht und wie du die richtige für dein Unternehmen auswählst.

Was ist eine Datenplattform?

Eine Datenplattform ist die zentrale technische Infrastruktur, die alle Daten eines Unternehmens zusammenführt, speichert, aufbereitet und für Auswertungen, Berichte und KI-Anwendungen bereitstellt. Sie ist das Fundament, auf dem Datenanalyse, Business Intelligence und KI-Projekte aufbauen.

Drei Begriffe werden oft synonym verwendet, beschreiben aber unterschiedliche Konzepte:

Datenplattform ist der Oberbegriff. Sie umfasst alle Werkzeuge und Schichten, die nötig sind, um Daten zu sammeln, zu speichern, zu verarbeiten und auszuwerten.

Datenarchitektur beschreibt das Gesamtdesign: Welche Datenquellen gibt es, wie fließen die Daten, wer hat Zugriff, welche Standards gelten? Sie ist die Blaupause, die Datenplattform die Umsetzung. Wer tiefer in das Thema einsteigen will, findet in unserem Artikel zu [Datenarchitektur](#) die wichtigsten Grundlagen.

Data Warehouse ist eine spezielle Form: Eine zentrale Datenbank, die strukturierte Daten für Reporting und Analyse aufbereitet. Modernes Data Warehouse (etwa Snowflake, Microsoft Fabric, Google BigQuery) ist Teil heutiger Datenplattformen, ersetzt aber nicht das Gesamtsystem.

Die fünf Bausteine einer modernen Datenplattform

Egal ob du eine Plattform neu aufbaust oder eine bestehende erweiterst: Die folgenden fünf Bausteine bilden den Kern jeder modernen Datenplattform.

1. Datenspeicher (Storage)

Hier landen alle Daten, strukturiert oder unstrukturiert. Moderne Plattformen setzen meist auf eine Kombination aus Data Lake (für Rohdaten, unstrukturierte Inhalte, Logs) und Data Warehouse (für strukturierte, analytisch aufbereitete Daten). Aktuelle Architekturen wie das "Lakehouse" verschmelzen beide Konzepte. Anbieter wie Snowflake, Databricks und Microsoft Fabric haben sich hier als die wichtigsten Optionen für Mittelständler:innen etabliert.

2. Datenintegration

Daten kommen aus dem ERP-System, dem CRM, der Produktionsanlage, dem Webshop, Sensoren, Excel-Listen und externen APIs. Sie müssen zuverlässig in die Plattform fließen, im richtigen Format und im richtigen Takt. Werkzeuge wie Fivetran, Airbyte oder Microsoft Data Factory automatisieren diese Pipelines und sorgen dafür, dass die Daten aktuell und konsistent bleiben.

3. Datenverarbeitung (Compute)

Sobald die Daten da sind, müssen sie aufbereitet werden: gefiltert, transformiert, angereichert, kombiniert. Diesen Schritt nennt man Data Engineering. Werkzeuge wie dbt (data build tool) oder Apache Spark sind heute Standard. Die Compute-Schicht ist der Ort, an dem aus Rohdaten verwertbare Information wird.

4. Daten-Governance und Sicherheit

Wer darf welche Daten sehen? Wie wird der Zugriff dokumentiert? Wie werden personenbezogene Daten geschützt? Diese Fragen sind kein Beiwerk, sondern zentraler Bestandteil. Microsoft Purview, Collibra oder Atlan bieten Governance-Lösungen, die Datenkatalog, Zugriffsverwaltung und Compliance-Dokumentation in einem Werkzeug bündeln. Gerade mit Blick auf DSGVO und EU AI Act ist Governance kein Nice-to-have mehr. Wer dort tiefer einsteigen will, findet in unserem Artikel zu [Datensicherheit im Unternehmen](#) das nötige Fundament.

5. Analyse und Visualisierung

Am Ende der Kette stehen die Werkzeuge, mit denen Daten genutzt werden: Power BI, Tableau und Looker für klassische Dashboards. Notebooks wie Jupyter oder Microsoft Fabric Notebooks für Data Science. Und zunehmend auch KI-Schnittstellen, die natürlich-sprachliche Abfragen auf die Datenplattform erlauben. Diese Schicht ist das Gesicht der Plattform für die Anwender:innen.

Die wichtigsten Datenplattform-Typen

Welche Architektur passt zu welchem Unternehmen? Hier die drei dominierenden Modelle.

Cloud-Datenplattform

Die Standardarchitektur für 2026. Anbieter wie Snowflake, Google BigQuery und Microsoft Fabric bieten vollständig verwaltete Datenplattformen in der Cloud. Vorteile: schnelle Skalierung, niedrige initiale Investition, Pay-per-Use, regelmäßige Updates ohne eigenen Aufwand. Nachteile: laufende Kosten je nach Nutzung, Abhängigkeit vom Anbieter. Für die allermeisten Mittelständler:innen ist dies heute der pragmatische Weg.

On-Premise-Datenplattform

Die Plattform läuft im eigenen Rechenzentrum. Vorteile: volle Datenhoheit, planbare Kosten, keine Abhängigkeit von externen Anbietern. Nachteile: hohe initiale Investition, eigenes Personal nötig, langsamere Skalierung. Für stark regulierte Branchen (Verteidigung, kritische Infrastruktur) und Unternehmen mit besonders sensiblen Daten manchmal weiterhin die richtige Wahl, aber zunehmend in der Minderheit.

Hybrid- und Multi-Cloud-Datenplattform

Eine Kombination aus eigener Infrastruktur und mehreren Cloud-Anbietern. Laut dem Bitkom Cloud Report 2025 setzen [bereits 29 Prozent der deutschen Unternehmen](#) auf Hybrid-Cloud und 41 Prozent auf Multi-Cloud. Vorteile: Flexibilität, Schutz vor Anbieter-Lock-in. Nachteile: höhere Komplexität, mehr Koordinationsaufwand. Für Mittelständler:innen oft die Antwort auf Anforderungen wie "manche Daten dürfen die EU nicht verlassen, andere können in der globalen Cloud liegen".

Datenplattform für den Mittelstand: 6 Fragen vor der Auswahl

Die Auswahl einer Datenplattform ist eine strategische Entscheidung mit langer Wirkung. Die folgenden sechs Fragen helfen dabei, eine fundierte Entscheidung zu treffen.

1. Welche Datenquellen müssen integriert werden?

Erstelle eine Liste aller Systeme, die heute oder in den nächsten zwei Jahren angebunden werden sollen: ERP, CRM, Produktionsdaten, Webshop, externe APIs, Excel-Tabellen, Sensoren. Plattformen unterscheiden sich stark darin, welche vorgefertigten Konnektoren sie bieten. Ein gutes Integrationsökosystem spart später viel Engineering-Aufwand.

2. Welche Datenmengen verarbeiten wir – heute und in zwei Jahren?

Eine Plattform für 10 Gigabyte Datenvolumen sieht anders aus als eine für 10 Terabyte. Wichtig ist nicht nur das aktuelle Volumen, sondern die Wachstumserwartung. Cloud-Plattformen skalieren elastisch, aber die Kostenkurve ist nicht linear. Wer früh modelliert, was wachsen wird, vermeidet später unangenehme Überraschungen.

3. Welche Anforderungen haben wir an Datenschutz und Compliance?

Verarbeitest du personenbezogene Daten? Bist du im regulierten Umfeld (Finanzdienstleistung, Gesundheit)? Müssen Daten in der EU bleiben? Der Bitkom Cloud

Report 2025 zeigt: Für [67 Prozent der Unternehmen](#) ist ein vertrauenswürdiges Herkunftsland des Cloud-Anbieters zwingende Voraussetzung. Diese Anforderungen müssen vor der Tool-Auswahl klar sein, nicht danach.

4. Welche Skills haben wir im Team?

Eine Datenplattform ist nur so gut wie die Menschen, die sie betreiben. Microsoft Fabric ist deutlich einfacher zu adoptieren, wenn das Team bereits in der Microsoft-Welt zuhause ist. Databricks lohnt sich für Teams mit Python-Affinität und Data-Science-Ambitionen. Snowflake bietet die niedrigste Einstiegshürde für klassische SQL-Anwender:innen. Wer ein Tool wählt, das niemand im Team beherrscht, kauft sich Beratungsabhängigkeit ein.

5. Wie sieht die Total Cost of Ownership aus?

Lizenz- und Hostingkosten sind nur ein Teil. Dazu kommen Implementierungskosten, Schulung, laufende Wartung, Erweiterungen. Eine realistische TCO-Rechnung über drei bis fünf Jahre vermeidet Fehlentscheidungen. Cloud-Plattformen sehen in Jahr eins günstig aus, können aber bei steigender Nutzung teurer werden als ein On-Premise-Setup.

6. Wie gut lässt sich die Plattform mit KI verbinden?

Datenplattform und KI gehen zunehmend Hand in Hand. Plattformen wie Snowflake, Databricks und Microsoft Fabric bieten heute integrierte KI-Funktionen: vom Embedding-Modell über Vektor-Suche bis zu LLM-Integration. Wer in den nächsten Jahren ernsthaft KI-Projekte fahren will, sollte diese Anbindung von Anfang an mitdenken.

Wie der Mittelstand Datenplattformen für KI aufbaut, ist eines der Kernthemen auf der [d:u27 vom 13. bis 14. April 2027 in Münster](#). Auf der Data Stage geht es um Datenstrategie, Infrastruktur, Governance und Sicherheit – mit echten Praxisbeispielen aus mittelständischen Unternehmen.

Praxistipps für die Einführung einer Datenplattform

Drei Tipps haben sich in der Praxis bewährt und tauchen in fast jeder erfolgreichen Datenplattform-Einführung auf.

Klein anfangen, früh Wert liefern

Die Versuchung ist groß, das große Wurfprojekt zu planen: Alle Datenquellen, alle Reports, alle Use Cases auf einmal. In der Praxis scheitern solche Mammutprojekte regelmäßig. Erfolgreiche Plattformen starten mit einem überschaubaren Anwendungsfall, oft im Vertrieb oder Controlling. Innerhalb von drei bis sechs Monaten gibt es erste sichtbare Ergebnisse, weitere Use Cases bauen darauf auf.

Data Governance von Anfang an mitdenken

Wer erst Daten sammelt und später Governance einführt, hat ein Problem. Zugriffsrechte, Datenkatalog, Lineage (also die Nachverfolgbarkeit, woher Daten kommen) sollten von Tag eins Teil der Plattform sein. Das spart später teure Aufräumaktionen.

In Datenkompetenz investieren

Eine Datenplattform schafft nur dann Wert, wenn Fachbereiche sie nutzen können. Schulungen, Daten-Communities im Unternehmen, klare Ansprechpartner:innen für Data Engineering und Analytics: Diese organisatorischen Investitionen sind genauso wichtig wie das Tooling. Mehr dazu in unserem Artikel zu [KI für Führungskräfte](#), in dem wir vertiefen, wie Entscheider:innen Datenprojekte sinnvoll begleiten.

Fazit: Die richtige Datenplattform ist die, die genutzt wird

Es gibt nicht die eine richtige Datenplattform. Es gibt nur die, die zu deinen Daten, deinen Anforderungen, deinem Team und deinen strategischen Zielen passt. Eine teure Plattform, die niemand nutzt, ist wertlos. Eine pragmatische Plattform, die im Alltag arbeitet und Entscheidungen verbessert, ist Gold wert.

Wer den Schritt von verteilten Datensilos hin zu einer integrierten Datenplattform geht, schafft die Grundlage für Automatisierung, KI und datengetriebene Entscheidungen. Und genau das ist 2026 keine strategische Kür mehr, sondern Pflicht.

Wie Mittelständler:innen den Schritt zur integrierten Datenplattform gehen, welche Architekturen sich in der Praxis bewähren und wie aus Daten echter Geschäftswert wird, zeigen Praktiker:innen auf der [d:u27 vom 13. bis 14. April 2027 in Münster](#). Auf der Data Stage, der Mittelstands Stage und in den Masterclasses teilen mittelständische Unternehmen ihre Erfahrungen – praxisnah, nahbar und mit konkreten Lessons Learned. [Sichere dir jetzt dein Ticket](#) und profitiere vom Austausch mit Menschen, die Datenplattformen bereits erfolgreich aufgebaut haben.