

*Datengetriebene
Prozesse und
Geschäftsmodelle*

Data Fabric - Basis für datengetriebene Geschäftsinnovationen

Marc Bastien



Definition(en) Data Fabric?

- Ein **Data Fabric** ist eine „subjektorientierte, nichtflüchtige, integrierte, zeitvariante Sammlung von Daten zur Unterstützung von Managemententscheidungen“.
- Beim **Data Fabric** handelt es sich um einen sehr großen Datenspeicher, der die Daten aus den unterschiedlichsten Quellen in ihrem Rohformat aufnimmt. Er kann sowohl unstrukturierte als auch strukturierte Daten enthalten und lässt sich für Big-Data-Analysen einsetzen.
- Ein **Data Fabric** ist ein Design Konzept, nach dem diese als integrierte Schicht für Daten und Verbindungsprozesse dient. Die Data Fabric ermöglicht durchgehende Analysen auf existierenden, erkennbaren und abgeleiteten Metadaten Assets, und erlaubt so Entwicklung, Verteilung und Nutzung von integrierten und wiederverwendbaren Daten über alle Umgebungen hinweg, ...

???

Quiz

Definition(en) Data Fabric?

- Ein **Data Warehouse** ist eine „subjektorientierte, nichtflüchtige, integrierte, zeitvariante Sammlung von Daten zur Unterstützung von Managemententscheidungen“. (nach **Bill Inmon**)
- Beim **Data Lake** handelt es sich um einen sehr großen Datenspeicher, der die Daten aus den unterschiedlichsten Quellen in ihrem Rohformat aufnimmt. Er kann sowohl unstrukturierte als auch strukturierte Daten enthalten und lässt sich für Big-Data-Analysen einsetzen. (**Big Data Insider**)
- Ein **Data Fabric** ist ein Design Konzept, nach dem diese als integrierte Schicht für Daten und Verbindungsprozesse dient. Die Data Fabric ermöglicht durchgehende Analysen auf existierenden, erkennbaren und abgeleiteten Metadaten Assets, und erlaubt so Entwicklung, Verteilung und Nutzung von integrierten und wiederverwendbaren Daten über alle Umgebungen hinweg, ... (**Übersetzt von Gartner**)
- BTW: „Fabric“ -> „Stoff“ / „Gewebe“ / „Tuch“ 😊 (nicht „Fabrik“ 😞)

Alles klar?

*Datengetriebene
Prozesse und
Geschäftsmodelle*

Definition(en)

- Gartner defines data fabric as a design concept that serves as an integrated layer (fabric) of data and connecting processes. A data fabric utilizes continuous analytics over existing, discoverable and inferred metadata assets to support the design, deployment and utilization of integrated and reusable data across all environments, including hybrid and multi-cloud platforms. ... (Gartner)
- Data Fabric ist eine einheitliche Architektur, die aus darauf laufenden Diensten oder Technologien besteht und verschiedene Cloud- und On-Premises-Umgebungen miteinander verbindet. Sie unterstützt Unternehmen mittels übergreifender Funktionalität bei der Verwaltung ihrer Daten. Das Ziel einer Data Fabric ist es, den Wert der Daten zu maximieren und die digitale Transformation zu beschleunigen. (Talend)
- Data Fabric ist ein Architekturansatz, der den Datenzugriff im Unternehmen und somit eine Self-Service-Datennutzung erleichtert. Diese Architektur ist unabhängig von Datenumgebungen, Prozessen, Dienstprogrammen und geografischen Gegebenheiten und bietet gleichzeitig durchgängige Funktionalitäten für das Datenmanagement. Data Fabric automatisiert die Datenerkennung, -verwaltung und -nutzung und ermöglicht es Unternehmen, Daten zur Optimierung ihrer Wertschöpfungskette zu nutzen. Mit Data Fabric können Unternehmen den Wert ihrer Daten steigern, indem sie die richtigen Daten zur richtigen Zeit und unabhängig von ihrem Speicherort bereitstellen. (IBM)

Alles klar?

Datengetriebene
Prozesse und
Geschäftsmodelle

Definition(en)

- Gartner defines data fabric as a design concept that serves as an **integrated** layer (fabric) of data and connecting processes. A data fabric utilizes continuous analytics over existing, discoverable and inferred metadata assets to support the design, deployment and utilization of integrated and reusable **data across all environments**, including hybrid and multi-cloud platforms. ... (Gartner)
- Data Fabric ist eine **einheitliche** Architektur, die aus darauf basierenden Diensten oder Technologien besteht und **verschiedene Cloud- und On-Premises-Umgebungen** miteinander verbindet. Sie unterstützt Unternehmen mittels übergreifender Funktionalität bei der Verwaltung ihrer Daten. Das Ziel einer Data Fabric ist es, den **Wert der Daten zu maximieren** und die digitale Transformation zu beschleunigen. (Talend)
- Data Fabric ist ein Architekturansatz, der den Datenzugriff im Unternehmen und somit eine Self-Service-Datennutzung erleichtert. Diese Architektur ist **unabhängig von Datenumgebungen**, Prozessen, Dienstprogrammen und geografischen Gegebenheiten und bietet gleichzeitig **durchgängige** Funktionalitäten für das Datenmanagement. Data Fabric automatisiert die Datenerkennung, -verwaltung und -nutzung und ermöglicht es Unternehmen, Daten zur Optimierung ihrer Wertschöpfungskette zu nutzen. Mit Data Fabric können Unternehmen den **Wert ihrer Daten steigern**, indem sie die richtigen Daten zur richtigen Zeit und unabhängig von ihrem Speicherort bereitstellen. (IBM)

Data Fabric Basics

*Datengetriebene
Prozesse und
Geschäftsmodelle*

Ziele im Unternehmen

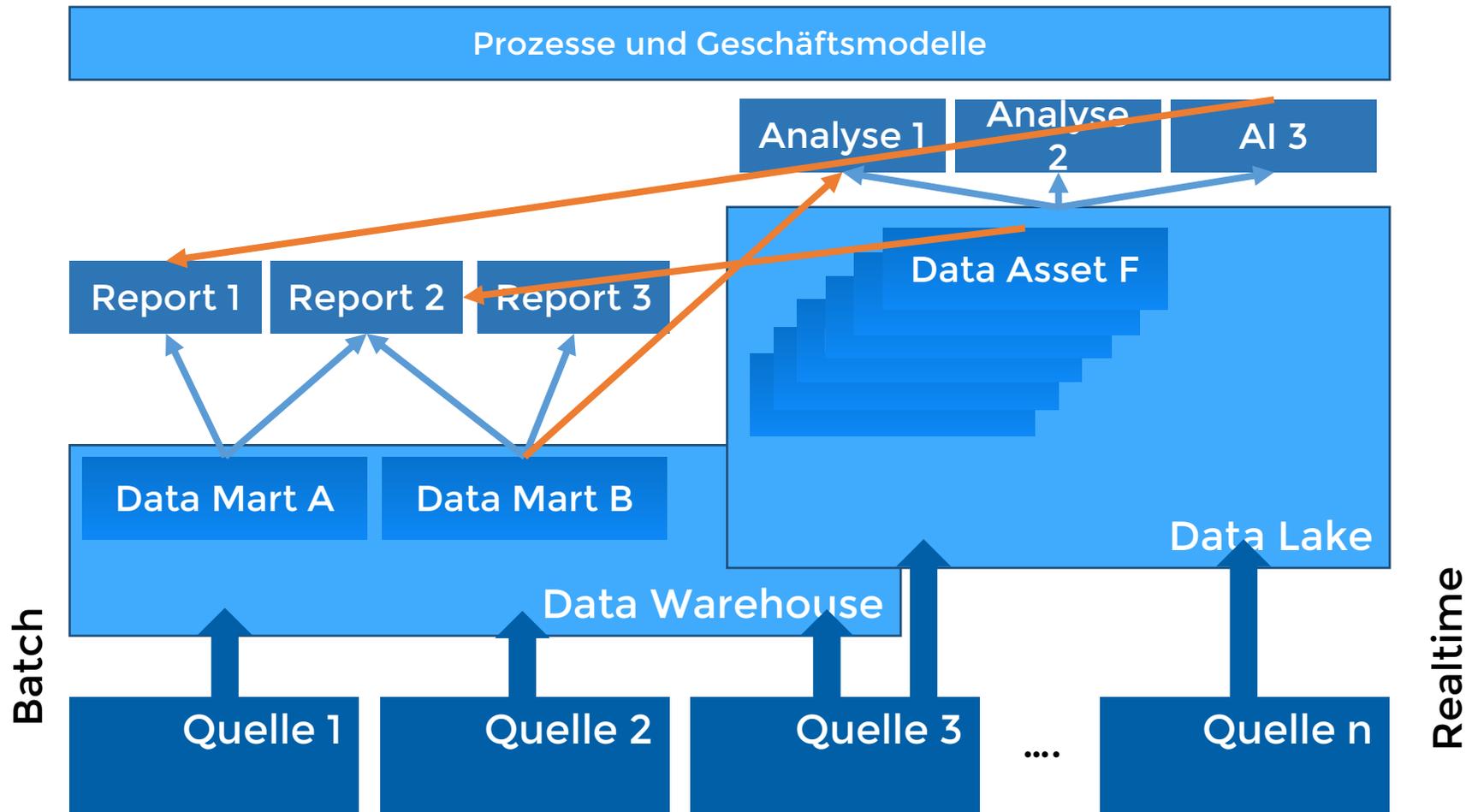
- **Fachlich: zentrale Definition der wesentlichen Geschäftsobjekte inkl. Zugriff darauf**
 - Nutzung des kompletten Datenbestandes, Potenziale voll nutzen
 - Bessere, umfassendere, aktuellere Informationen, schnellere Entscheidungen
 - Umfassende Unterstützung der Kernprozesse im Unternehmen!
- **Technisch**
 - Integration, Integration, Integration
 - Auflösen von Silos
 - Erweiterung des analytischen Stacks (Self-Service, AI bis in die Geschäftsprozesse)
 - Governance: einheitliche Begriffe, Schutz, Überwachung, Datenqualität

Data Fabric Basics

„Haben wir das nicht schon?“

➤ Data Warehouse, Data Lake, Data Mart, Data xxxx

*Datengetriebene
Prozesse und
Geschäftsmodelle*



Data Fabric Eigenschaften

*Datengetriebene
Prozesse und
Geschäftsmodelle*

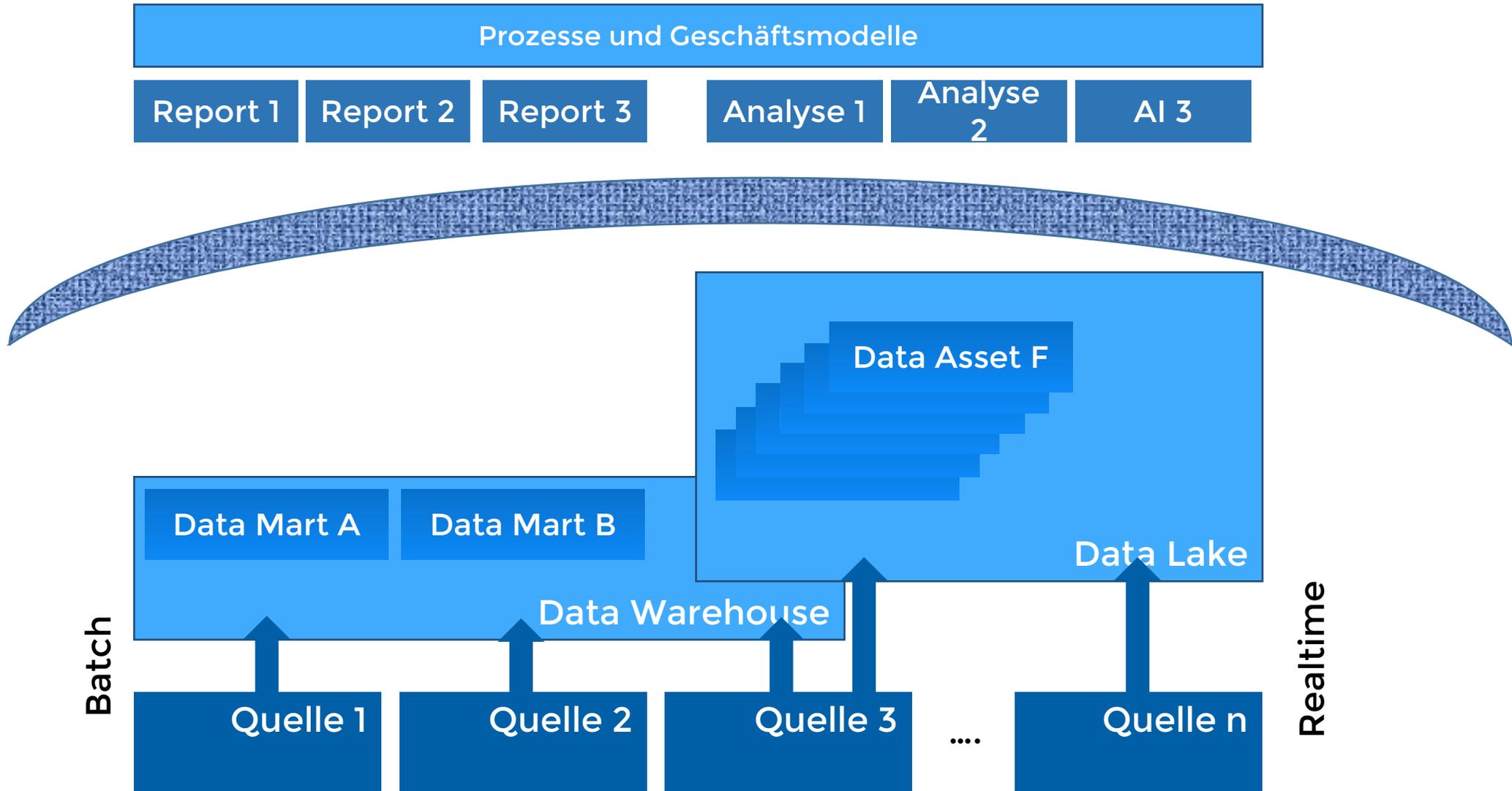
Problem bekannt?

- Bekannte Lösungsansätze neu interpretiert
 - Noch ein „Enterprise DWH / DataLake“?
 - (+) komplette Integration aller Unternehmensdaten
 - (-) steigende Komplexität und sinkende Agilität
 - (-) Exponentielle Kosten mit jeder neuen Quelle
 - Data Hub?
 - (+) verteilte Daten, gemeinsamer Zugriff
 - (-) Sicherstellung von Datenqualität und Vertrauen
- Balance herstellen zwischen De-zentralisierung und Globalisierung
 - Virtuelle Verbindungen zu Datenquellen erlauben leichtere, nicht-physische Integration
 - Globaler Katalog stellt Erkenntnisse fachlich gemeinsam zur Verfügung

Data Fabric Basics

*Datengetriebene
Prozesse und
Geschäftsmodelle*

MIT Netz (Fabric) und doppeltem Boden!



Data Fabric Eigenschaften

*Datengetriebene
Prozesse und
Geschäftsmodelle*

Erforderliche Fähigkeiten

- Gemeinsamer, transparenter Zugriff auf Daten in unterschiedlichem Format, Lokationen, Bedeutungen
- Integration, Transformationen, Zugriff auf Daten aller Art
- Konzept von fachlichen Assets, die für verschiedene Reports/Analysen zur Verfügung stehen
- Regeln unterstützen Zugriffe, Anforderungen und ggfls. Security
- Lineage, Abhängigkeiten, Datenqualität und Dokumentation
- Automatisierung („AI“): Nice-to-have?
 - Zusammenführung von Assets, Objekten, Logik
 - Abhängigkeiten zwischen Assets und Security
 - Workflows zur Freigabe
 - Auditierung

Data Fabric Eigenschaften

*Datengetriebene
Prozesse und
Geschäftsmodelle*

Fähigkeiten = Funktionen

- Datenintegration-> erweitert um Data Pipelining und echte Wahl zwischen ETL/ELT/EAI und Batch/(Near-)Realtime - **von der Quelle bis zum Prozess!**
- Transparenter Zugriff, Konnektivität -> Daten Virtualisierung
- Fachliche und technische Assets -> Datenkatalog
- Security usw. -> als Konsequenz aus der Integration der o.g. Funktionen („Plattform“)
- Lineage usw. -> als Konsequenz aus der Plattform
- Darüber hinaus möglicherweise noch
 - Datenhaltung (DBMS, S3 usw.)
 - Business Intelligence & weitere Tools

Data Fabric - Zusammenfassung

- Data Fabric = Architekturansatz <> Technologie
- Definition ist von den Herstellern / Analysten geprägt, muss sich noch bewähren
- Idee nachvollziehbar, Praxisbezug durchaus realistisch
- Aufwand vs. Nutzen muss individuell bestimmt werden
- **Analytics ist das Rückgrat der Digitalisierung!**

*Datengetriebene
Prozesse und
Geschäftsmodelle*



Kundenvortrag:

Entwicklung des Data Warehouse bei TRUMPF Photonic Components: Gemeinsam zur zentralen, qualitätsgesicherten Informationsplattform

Thomas Göttl, IT-Leiter TRUMPF Photonic Components