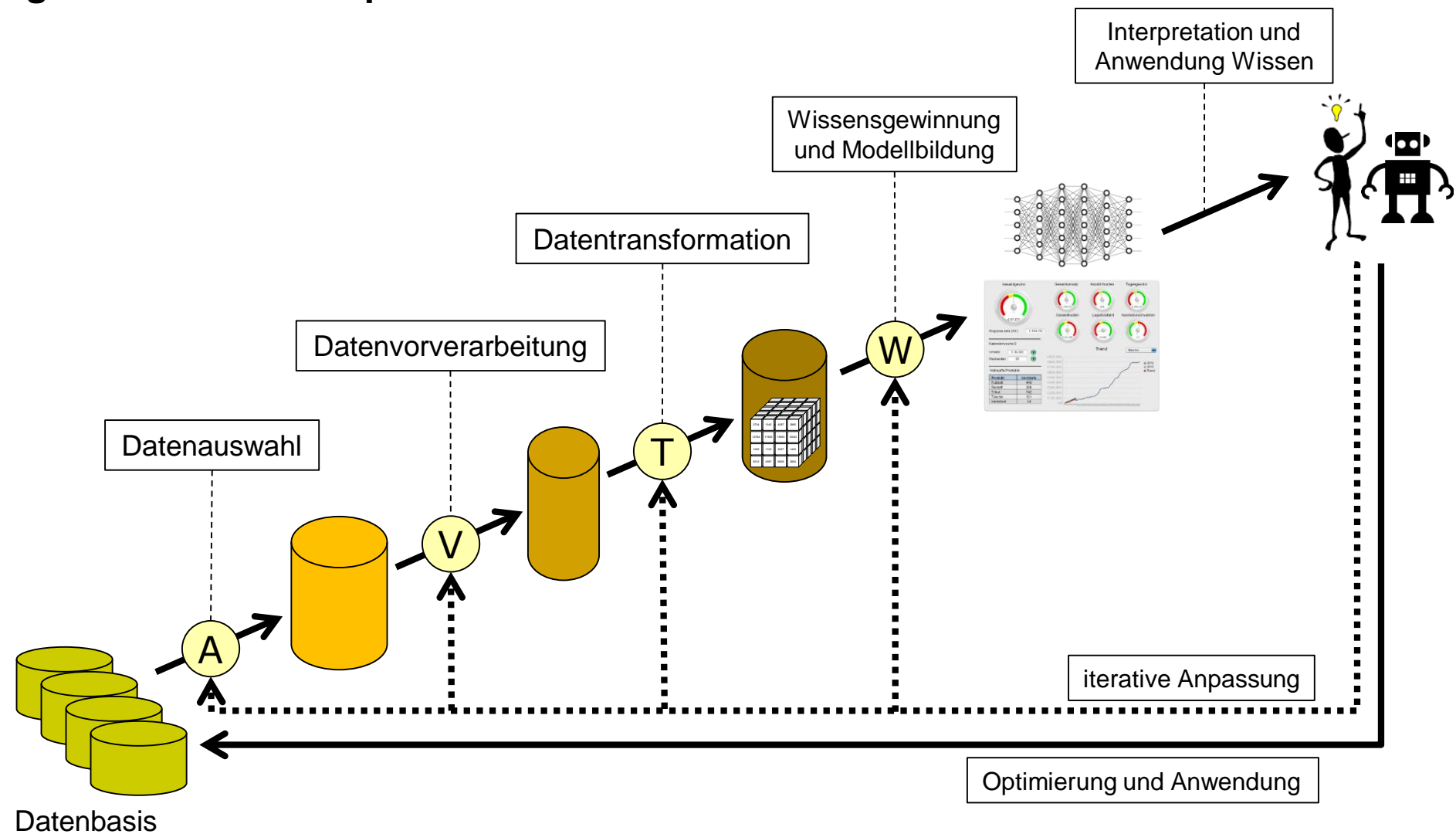


Data Science Systeme und Architekturen

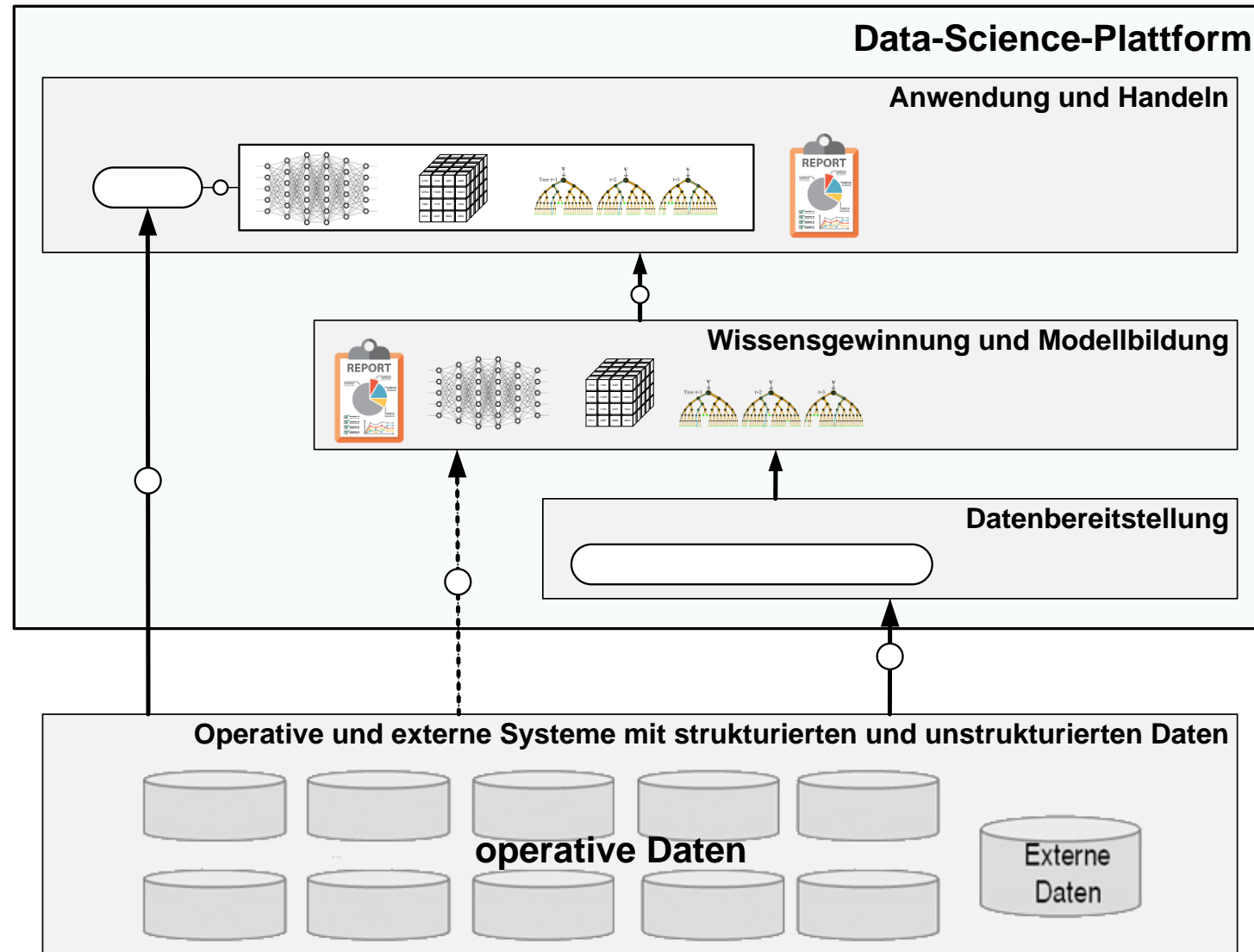


Data Science als Prozess (logische Sicht)

Zur Erinnerung: Data Science ist prozessorientiert

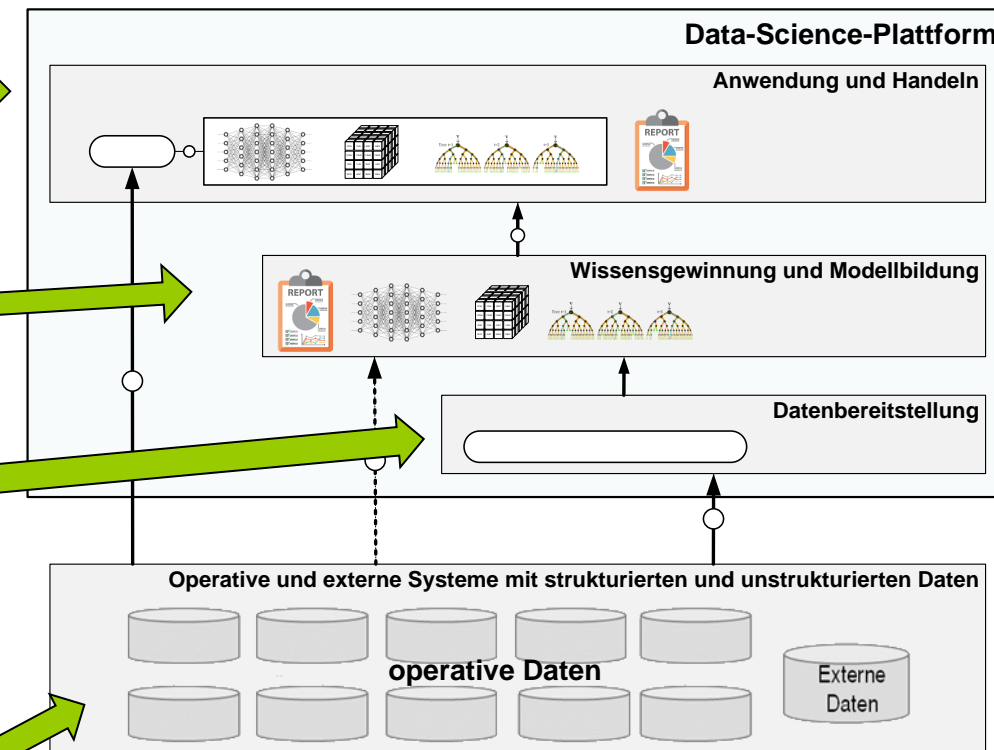


Data-Science-Plattformen (konzeptionell)



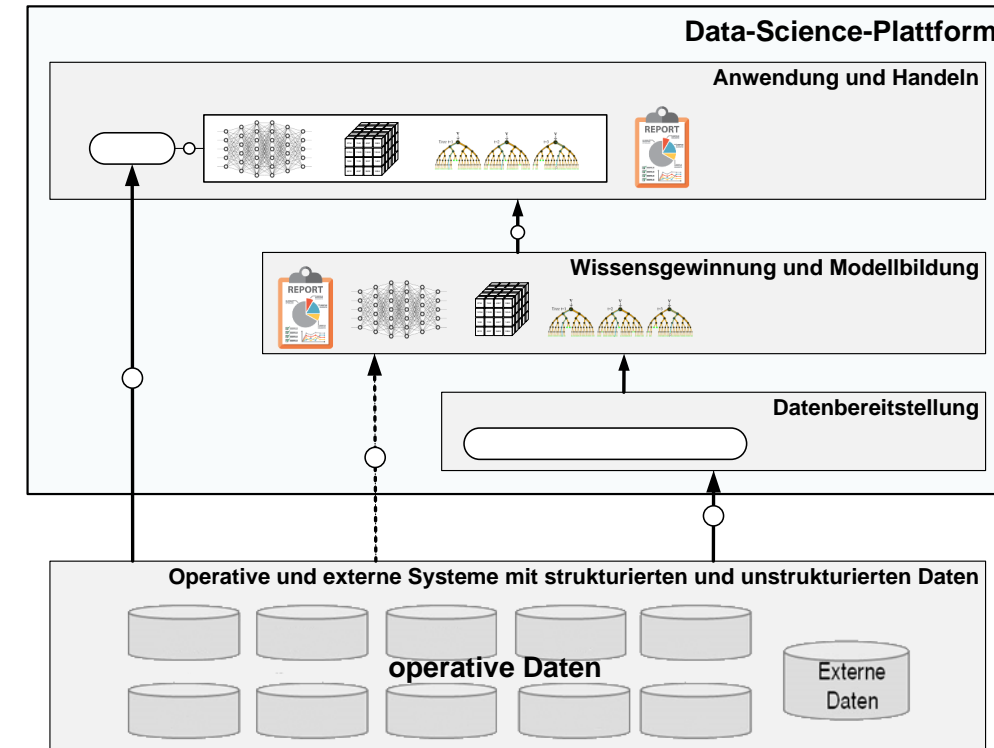
Data-Science-Plattformen

- Anwendung
 - der Modelle und Wissen
 - durch Schnittstellen (Benutzer, Systeme)
- Wissensgewinnung und Modellbildung
 - Analyse- und Miningsysteme / Methoden
- Datenbereitstellung
 - Generierung konsistenter **dispositiver** Daten aus Quellsystemen
 - Strukturiert, semi-strukturiert, unstrukturiert
 - aggregiert
 - aufbereitet, homogen, skaliert, angereichert
- Datenbasis: Operative und externe Quellsysteme
 - **operative** und externe Daten
 - strukturierte und unstrukturierte Daten

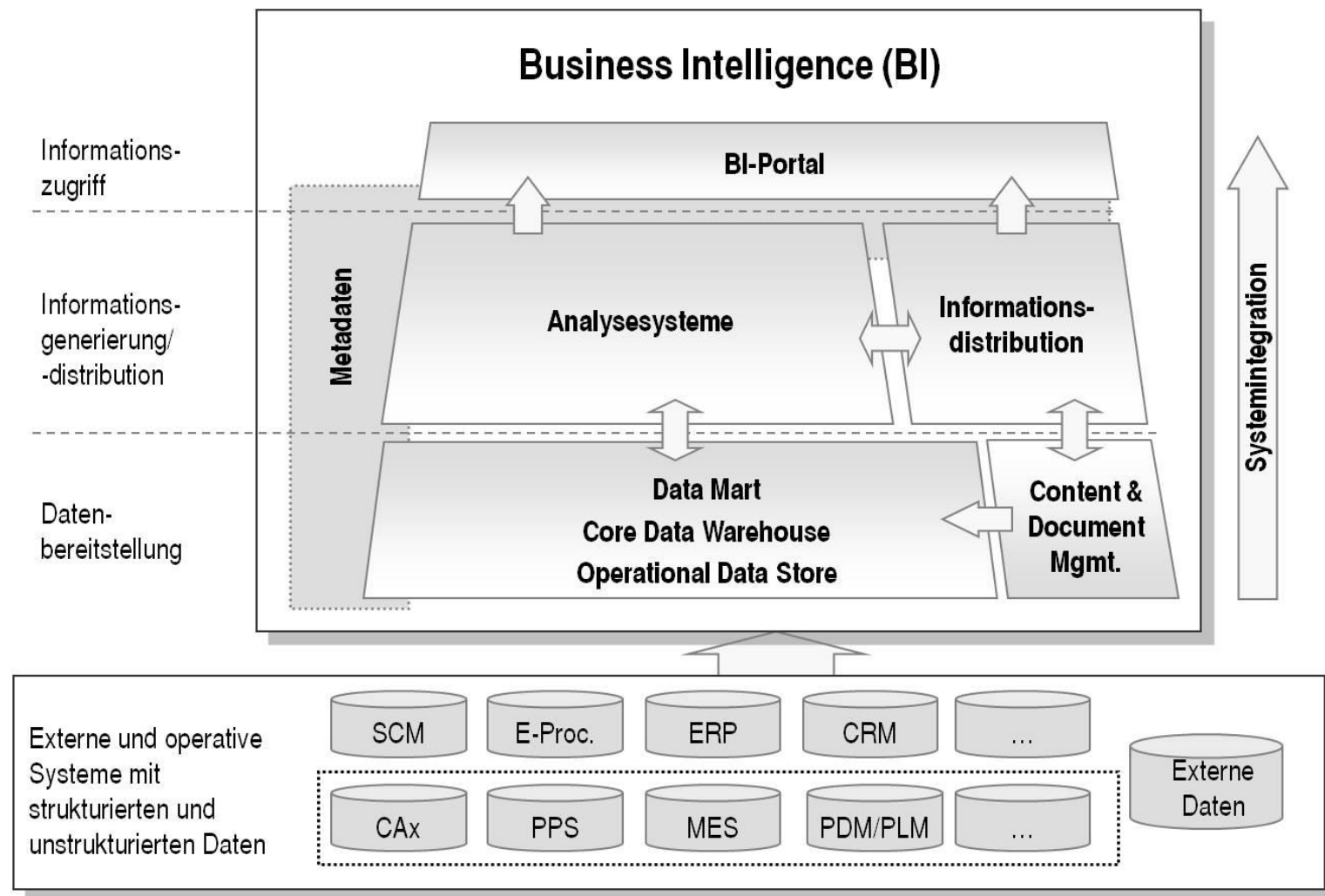


Beispiele/Lösungen für Data-Science-Plattformen

- λ - und κ -Architekturen für Big Data und IoT
- Data Science Labs: Lokale, abgegrenzte Plattformen für Research & Experiment
- Business-Intelligence-Systeme mit Data Warehouses



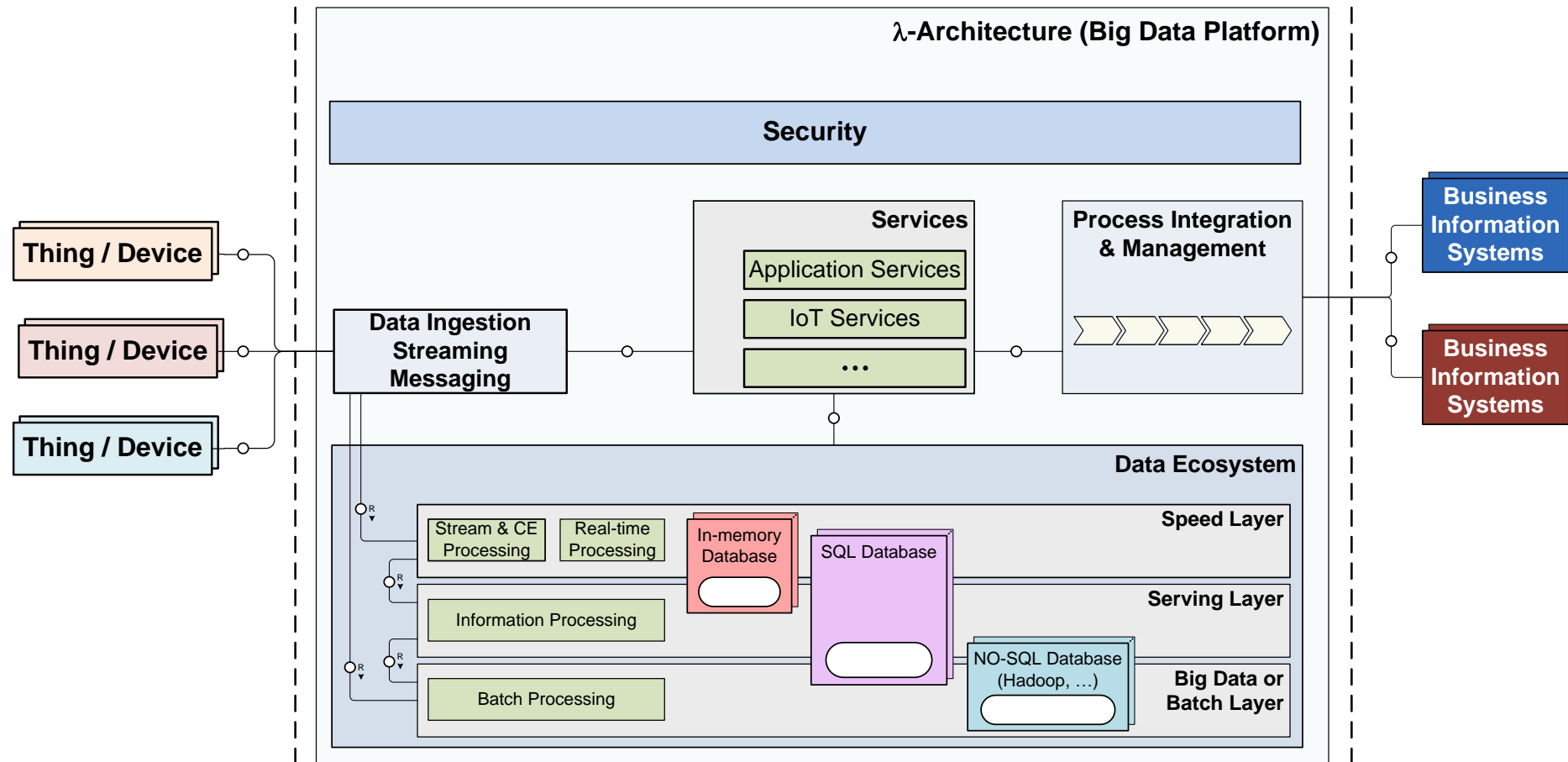
Business-Intelligence-Systeme mit Data Warehouses



©Kemper, Mehanna, Baars: Business Intelligence, Vieweg 2010, ISBN 978-3-8348-0719-9

λ-Architekturen für Big Data und IoT (exemplarisch)

- Details: Vorlesungen *Big Data Programming* und *Big Data Storage*



IT-basierter, integrierter Gesamtansatz zur Wissensgewinnung und Modellbildung und Handlungsunterstützung

- Basierend auf Gesamtheit der relevanten Daten in den operativen Informationssystemen und externer Datenquellen
 - (Geschäfts-)Daten
 - (Geschäfts-)Prozesse
 - Externe Daten

- Entwicklung moderner Technologien und Architekturen
 - In-memory- und NO-SQL-Datenbanken
 - Big-Data- und IoT-Architekturen, Streaming-Technologien
 - Cloud Computing, Distributed/Parallel/Scalable Computing, High-Performance Computing

Data Science aus konzeptionell algorithmischer Sicht

Methodischer Gesamtansatz zur Wissensgewinnung, Modellbildung und Handlungsunterstützung

- Moderne Methoden aus den Bereichen
 - Mathematik: Statistik, Analysis, diskrete Mathematik, Graphentheorie, etc.
 - Data Mining, Regression, Clustering, Assoziationsanalyse, Machine & Deep Learning, KI, Predictive Analytics, etc.
- Entwicklung skalierbarer und effizienter Algorithmen sowie Methoden des Distributed Computing

- Data Science auf allen Ebenen: operativ, taktisch, strategisch
 - im Kontext der gegebenen Anwendungs- oder Geschäftsdomäne

Fragen?

