

Business Intelligence und Data Mining

Dr. Robert Stahlbock

Universität Hamburg

Institut für Wirtschaftsinformatik (IWI)

- Was bedeutet Entscheidungsunterstützung bzw. Unterstützung von Führungsaufgaben ?
- Welche Mittel stellt die Wirtschaftsinformatik bereit, um Managementaufgaben zu unterstützen ?
- Was bedeutet Business Intelligence (BI) ?
- Was versteht man unter BI-Systemen ?



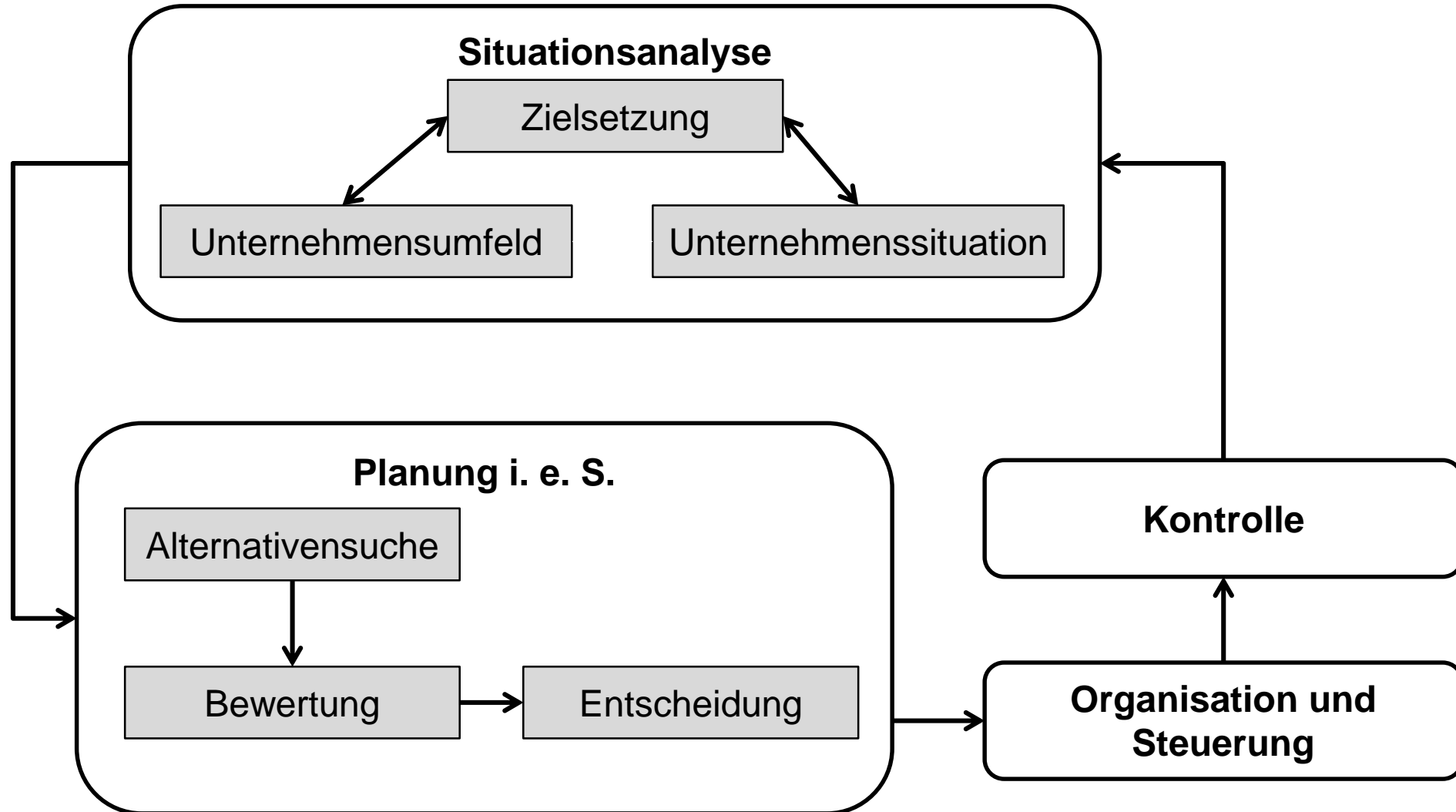
■ **Managementaufgaben und -prozesse**

■ Managementunterstützungssysteme

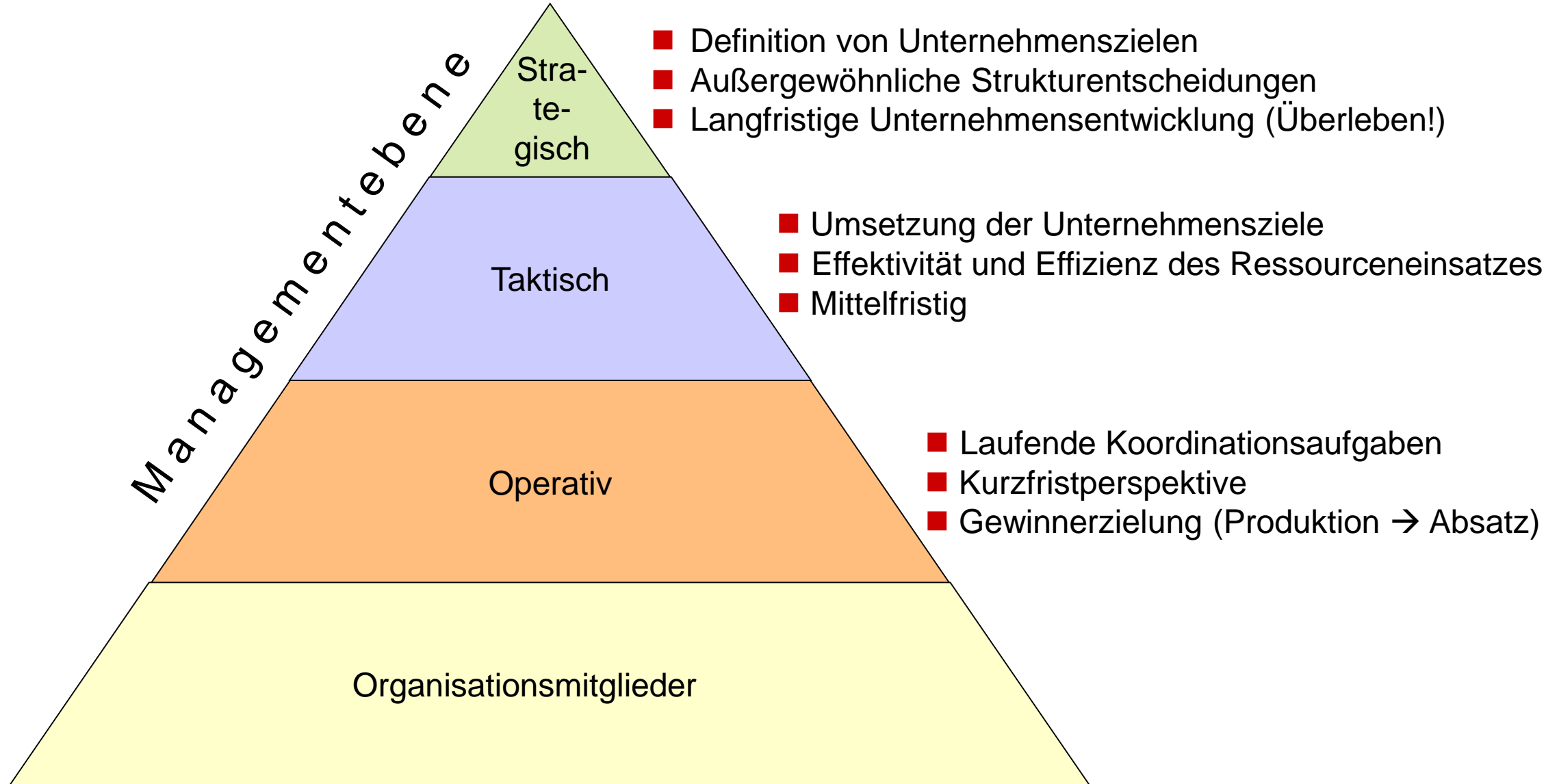
■ Business Intelligence

- Begriff
- Basistechnologien
- Nutzergruppen
- Architekturkonzepte
- Reifegradmodelle
- Systeme





Anthony's (1965): Taxonomie von Managementaktivitäten



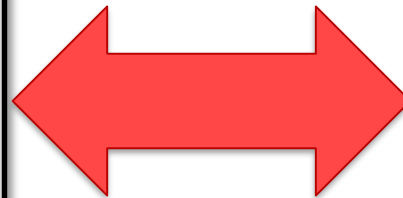
Managementphasen		Managementebenen		
		Operativ	Taktisch	Strategisch
Situations- analyse	Zielsetzung			
	Unternehmensumfeld			
	Unternehmenssituation			
Planung i.e.S.	Alternativensuche			
	Bewertung			
	Entscheidung			
Organisation & Steuerung				
Kontrolle				

Vielfältige Anforderungen hinsichtlich Management-
unterstützung bzw. an Managementunterstützungssysteme

- Phasenschema gilt prinzipiell für alle Managementaufgaben
- Unterschiede zwischen Aufgaben verschiedener Managementebenen
 - Abstraktheit der Ziele
 - Reichweite des Planungszeitraums
 - Bedeutung von Entscheidungen
 - Detaillierungsgrad von Kontrolldaten

Strukturierte/Programmierbare Entscheidungen

- Routineaufgaben
- Wiederkehrend
- Definierte Prozeduren



Unstrukturierte/Nicht-Programmierbare Entscheidungen

- Neu, ggf. einmalig
- Keine bekannte Methoden
- Problemadäquater Lösungsansatz ist zu entwickeln

Dimensionssystem für Führungsaufgaben

Managementphasen

<i>Unstrukturiert</i>		Operativ	Taktisch	Strategisch
<i>Semi-strukturiert</i>		Operativ	Taktisch	Strategisch
<i>Strukturiert</i>		Operativ	Taktisch	Strategisch
Situationsanalyse	Zielsetzung			
	Unternehmensumfeld			
	Unternehmenssituation			
Planung i.e.S.	Alternativensuche			
	Bewertung			
	Entscheidung			
Organisation & Steuerung				
Kontrolle				

Entscheidungsarten

Managementebenen

- Schnelle, exakte und umfassende personen- und sachgerechte Zurverfügungstellung und übersichtliche Aufbereitung des relevanten Informationsmaterials
- Unterstützung in allen Planungs- und Entscheidungsphasen (Analyse-, Planungs-, Realisierungs- und Kontrollphase)
- Zugriff auf historische und aktuelle interne und externe Informationen
- Möglichkeit zur permanenten Überwachung interner und externer Prozesse
Feststellung und Analyse von Abweichungen von Sollzuständen
- Unterstützung bei der Problemstrukturierung (Problemerkennung und -aufbereitung)
- Möglichkeit zur Erklärung und Interpretation der Informationen und Prozesse mit entsprechenden Simulations-, Optimierungs- und Analysemöglichkeiten

- Auswertung der Informationen nach unterschiedlichen Kriterien
- Methodische Aufbereitungsmöglichkeiten der Information (z.B. in Form von Berechnungen, Optimierungen, Simulation, Aggregationen und Disaggregationen)
- Zugriff auf gesamtes relevantes Wissen (Fakten- und Erfahrungswissen)
- Unterstützung auf allen Ebenen und in allen Bereichen mit entsprechenden Koordinations- und Integrationsmöglichkeiten
- Anpassbarkeit an betriebsspezifische Strukturen und Prozesse sowie benutzerindividuelles Problemlösungsverhalten
- Unterstützung der Kommunikation unabhängig von Zeit und Ort

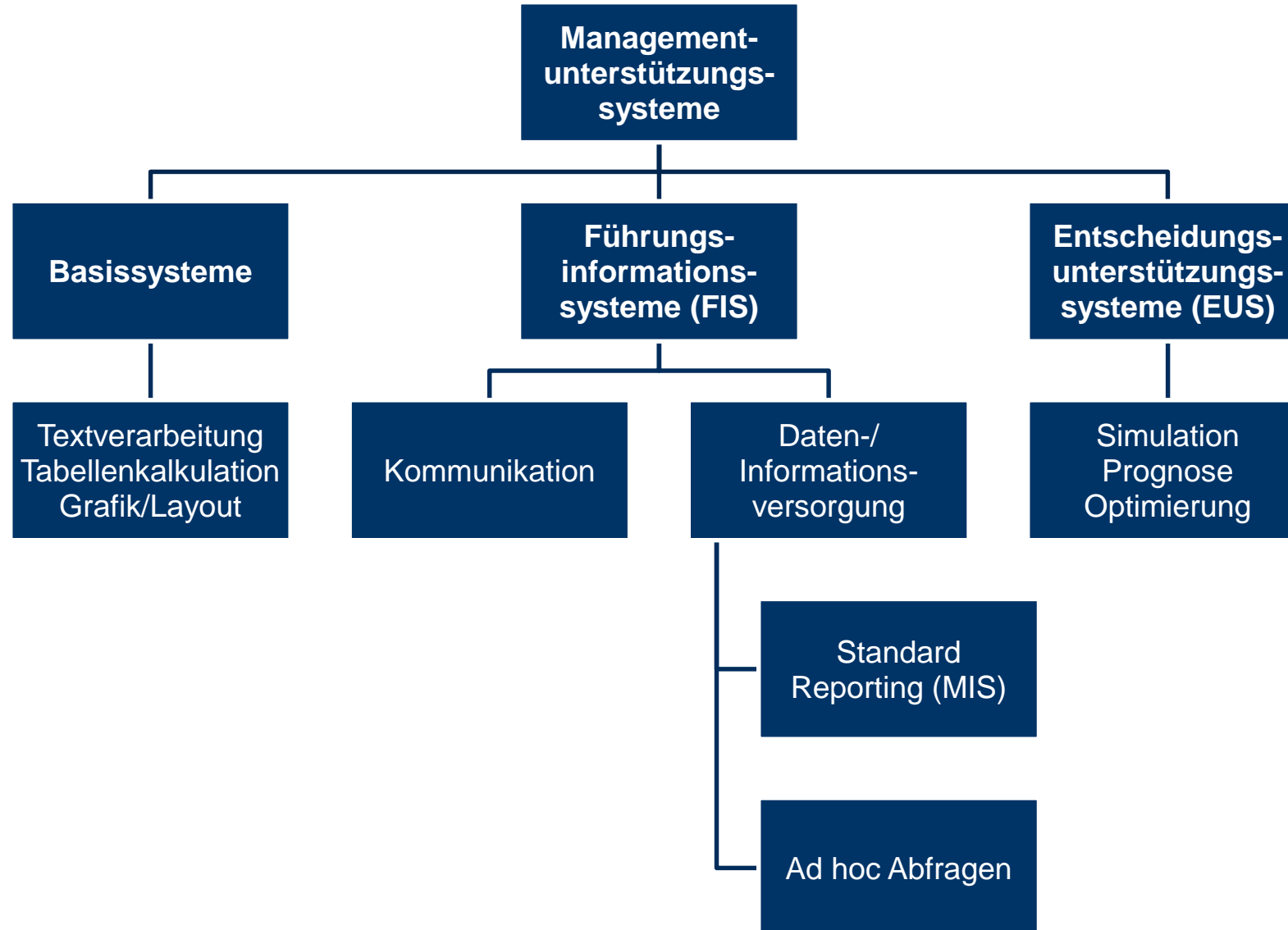
- Managementaufgaben und -prozesse
- **Managementunterstützungssysteme**
- Business Intelligence
 - Begriff
 - Basistechnologien
 - Nutzergruppen
 - Architekturkonzepte
 - Reifegradmodelle
 - Systeme



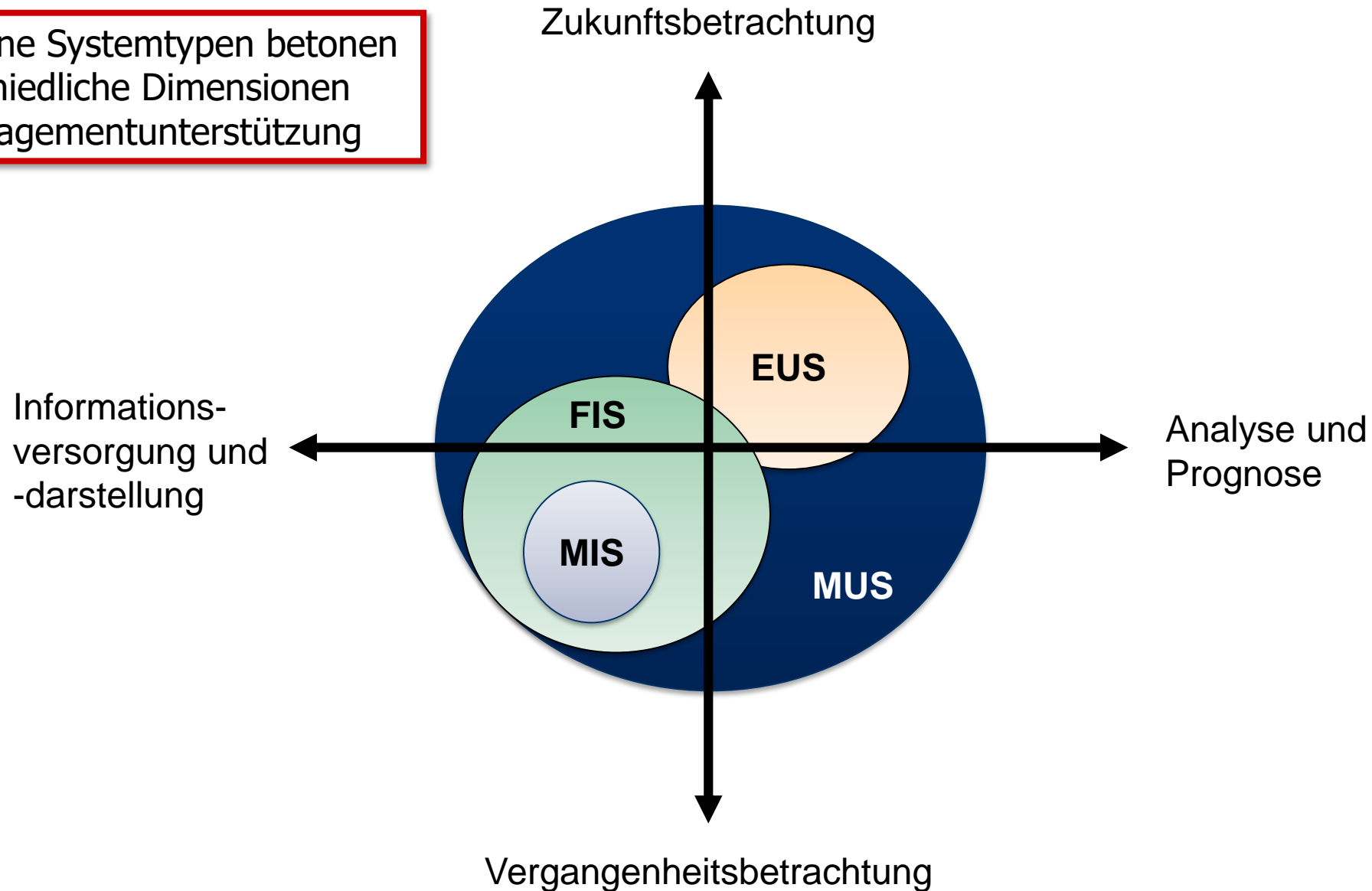
- **Managementinformationssysteme (MIS)**
 - Aufgabe: Auf Basis operativer Datenbestände Managern verschiedener Hierarchieebenen detaillierte und verdichtete Informationen, insb. in Form standardisierter Berichte, zur Informationsgewinnung bereitstellen
- **Entscheidungsunterstützungssysteme (EUS)**
 - Formale Modelle und Methoden zur Entscheidungsunterstützung
 - Spezialisierung auf Teilprobleme (z.B., Maschinenbelegung)
 - Häufig: Effektivität der Entscheidungsfindung verbessern
- **Führungsinformationssysteme (FIS)**
 - Interne und externe verdichtete Daten
 - Ausgeprägte Kommunikationselemente (Dialogorientierung)
 - Intuitive Benutzerschnittstellen

Vgl. Vorlesung
Informationsmanagement
von Prof. Voß

Funktionen und historische Ausprägungen von Management-Unterstützungssystemen



Verschiedene Systemtypen betonen unterschiedliche Dimensionen von Managementunterstützung



- Managementaufgaben und -prozesse
- Managementunterstützungssysteme
- **Business Intelligence**
 - Begriff
 - Basistechnologien
 - Nutzergruppen
 - Architekturkonzepte
 - Reifegradmodelle
 - Systeme



- Business Intelligence is a set of methods, processes, architectures, applications, and technologies that gather and transform raw **data** into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and **decision-making** (to drive business performance).
Forrester Report "Topic Overview: Business Intelligence", 2008
- Begriffliche Klammer, die unterschiedliche Technologien und Konzepte im Umfeld entscheidungsunterstützender Systeme zusammenführt und dabei eine **entscheidungsorientierte Sammlung und Aufbereitung von Daten** über das Unternehmen und dessen Umwelt sowie deren Darstellung in Form von **geschäftsrelevanten Informationen für Analyse-, Planungs- und Steuerungszwecke** zum Gegenstand hat.
Gluchowski et al., 2008

- Sammelbegriff für den IT-gestützten Zugriff auf Informationen, sowie die IT-gestützte Analyse und Aufbereitung dieser Informationen. Ziel dieses Prozesses ist es, aus dem im Unternehmen vorhandenen Wissen, **neues Wissen zu generieren**. Bei diesem neu gewonnenen Wissen soll es sich um **relevantes, handlungs-orientiertes Wissen handeln, welches Management-entscheidungen zur Steuerung des Unternehmens unterstützt**.

Gabler Wirtschaftslexikon, online 2018

- (mindestens) sieben unterschiedliche Deutungen von BI:
 - BI als Fortsetzung der Daten- und Informationsverarbeitung für Unternehmensleitung
 - BI als Filter in der Informationsflut: Informationslogistik
 - BI = Management-Information-System (MIS), aber besonders schnelle/flexible Auswertung
 - BI als Frühwarnsystem (Alerting)
 - BI = Data Warehouse
 - BI als Informations- und Wissensspeicherung
 - BI als Prozess: Symptomerhebung → Diagnose → Therapie → Prognose → Therapiekontrolle

BI konzentriert sich auf "analytische Datenverarbeitung".



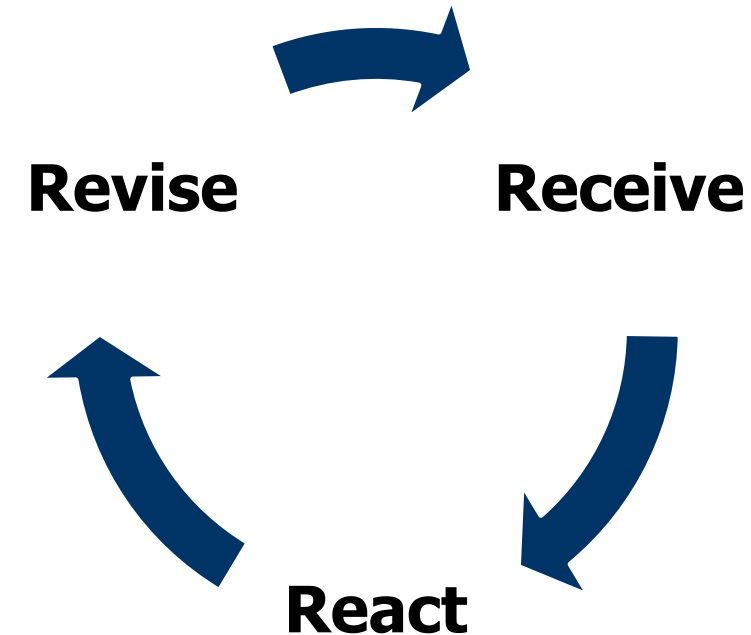
Verständnis von „Intelligence“

■ Merkmale von Intelligenz/Intelligence

- ❑ Komplexe Informationen rasch erfassen, Zusammenhänge verstehen
- ❑ Auf Information adäquat reagieren und Erkenntnisse umsetzen (Flexibilität und Kreativität)
- ❑ Rasch lernen und weitere Konsequenzen ziehen (Lernfähigkeit)

■ „Intelligence“ Konzepte

- ❑ Betonung der Informationslogistik
- ❑ Gezielte Sammlung von Daten
- ❑ Verdichtung und Anreicherung
- ❑ Gewinnung von Erkenntnissen, Abschätzungen, Meinungen



■ Geburtsstunde: 1958

- ❑ HP Luhn: A Business Intelligence System. IBM Journal 1958, 314-319.
- ❑ *"Business is a collection of activities carried on for whatever purpose, be it in science, technology, commerce, industry, law, government, defense, etc."*
- ❑ *"Intelligence is ... the ability to apprehend the interrelationships of presented facts in such a way as to guide towards a desired goal".*
- ❑ *„An automatic system is being developed to disseminate information to the various sections of any industrial, scientific or government organization. This intelligence system will utilize data-processing machines for auto-abstracting and auto-encoding of documents and for creating interest profiles for each of the 'action points' in an organization. Both incoming and internally generated documents are automatically abstracted, characterized by a word pattern, and sent automatically to appropriate action points."*

■ Geburtsstunde: 1958

- ❑ HP Luhn: A Business Intelligence System. IBM Journal 1958, 314-319.
- ❑ *"Business is a collection of activities carried on for whatever purpose, be it in science, technology, commerce, industry, law, government, defense, etc."*
- ❑ *"Intelligence is ... the ability to apprehend the interrelationships of presented facts in such a way as to guide towards a desired result."*
- ❑ *„An automatic system is being developed which will utilize data-processing machines for automatically processing incoming documents and for creating interest profiles for each of the documents. Documents in incoming and internally generated documents are automatically abstracted, characterized by a word pattern, and sent automatically to appropriate action points."*

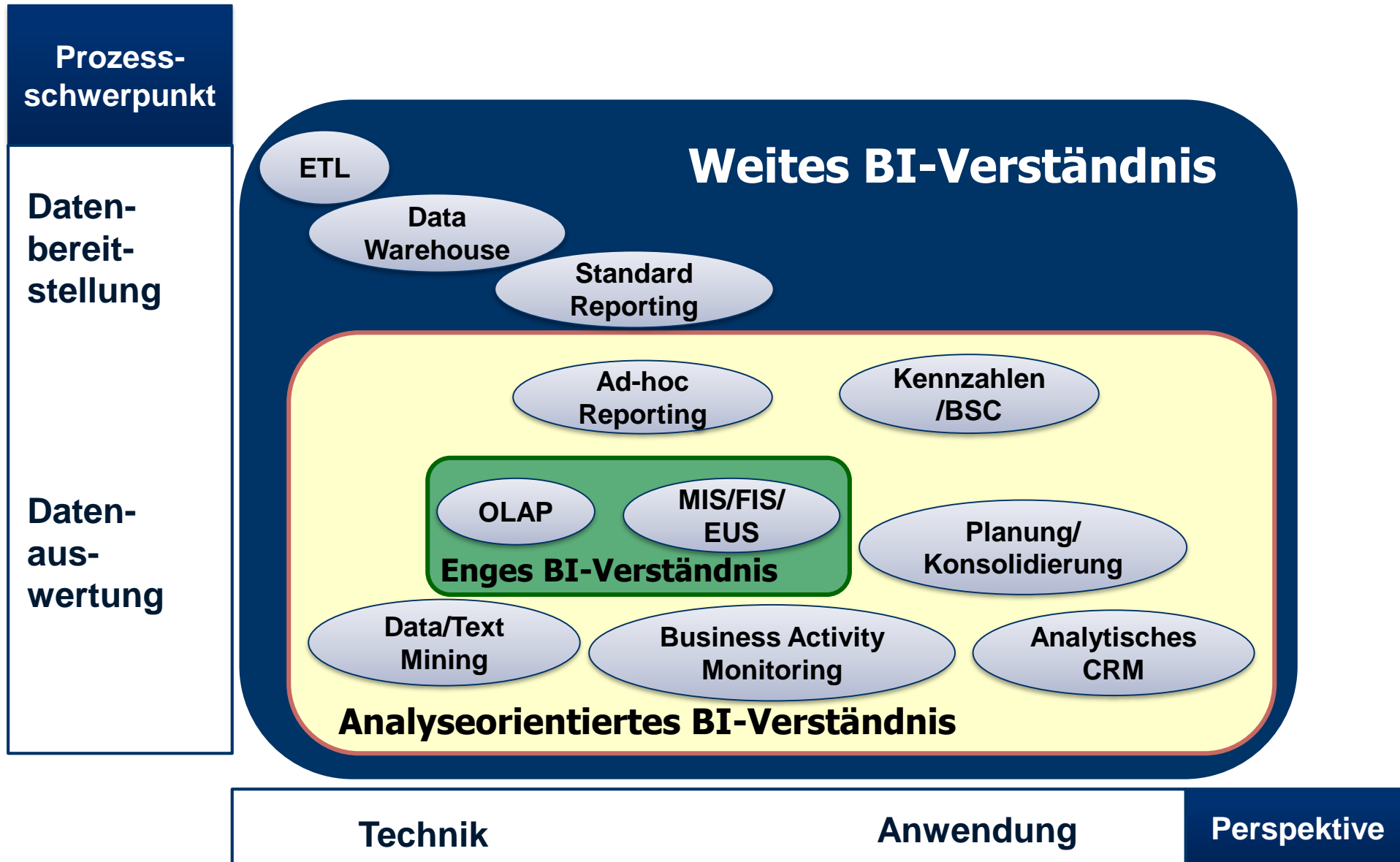
Dokumentenmanagement:

- Automatische Indizierung
- Speicherung
- Aktive Distribution

Def.

Begriffliche Klammer, die unterschiedliche Technologien und Konzepte im Umfeld entscheidungsunterstützender Systeme zusammenführt und dabei eine entscheidungsorientierte Sammlung und Aufbereitung von Daten über das Unternehmen und dessen Umwelt sowie deren Darstellung in Form von geschäftsrelevanten Informationen für Analyse-, Planungs- und Steuerungszwecke zum Gegenstand hat.

[Gluchowski et al. 2008]



- ETL: Extract, Transport & Loading. SW-Werkzeuge, die benutzt werden, um ein Data Warehouse mit Daten zu befüllen.
- Analytisches CRM: Derjenige Teil des Kundenbeziehungsmanagements (→ CRM), der die Speicherung und Auswertung kundenbezogener Daten zum Gegenstand hat.
- BSC: Balanced Score Card. Mehrdimensionales, hierarchisches Kennzahlensystem zur Steuerung und Kontrolle von z.B. Unternehmen oder Geschäftseinheiten
- OLAP: Online-Analytical-Processing. Interaktives Analysekonzept für multi-dimensionale Auswertungen.

Unterteilungen, Typen

- ❑ numerisch, Text, grafisch, ...
- ❑ strukturiert, unstrukturiert
- ❑ Standardformat, proprietäres Format
- ❑ intern, extern
- ❑ Datendatei, Datenbank
- ❑ Fakten, Schätzungen, Pläne, Prognosen
- ❑ disaggregiert, aggregiert
- ❑ ...

■ Verfügbarkeit

- nicht vorhanden
- benötigt (evtl. sehr aufwendige) Verarbeitungsschritte

■ Überfrachtung

- undurchschaubare „Flut“
- Datenorganisation komplex; effektive/effiziente Zugriffe?

■ Verteilung

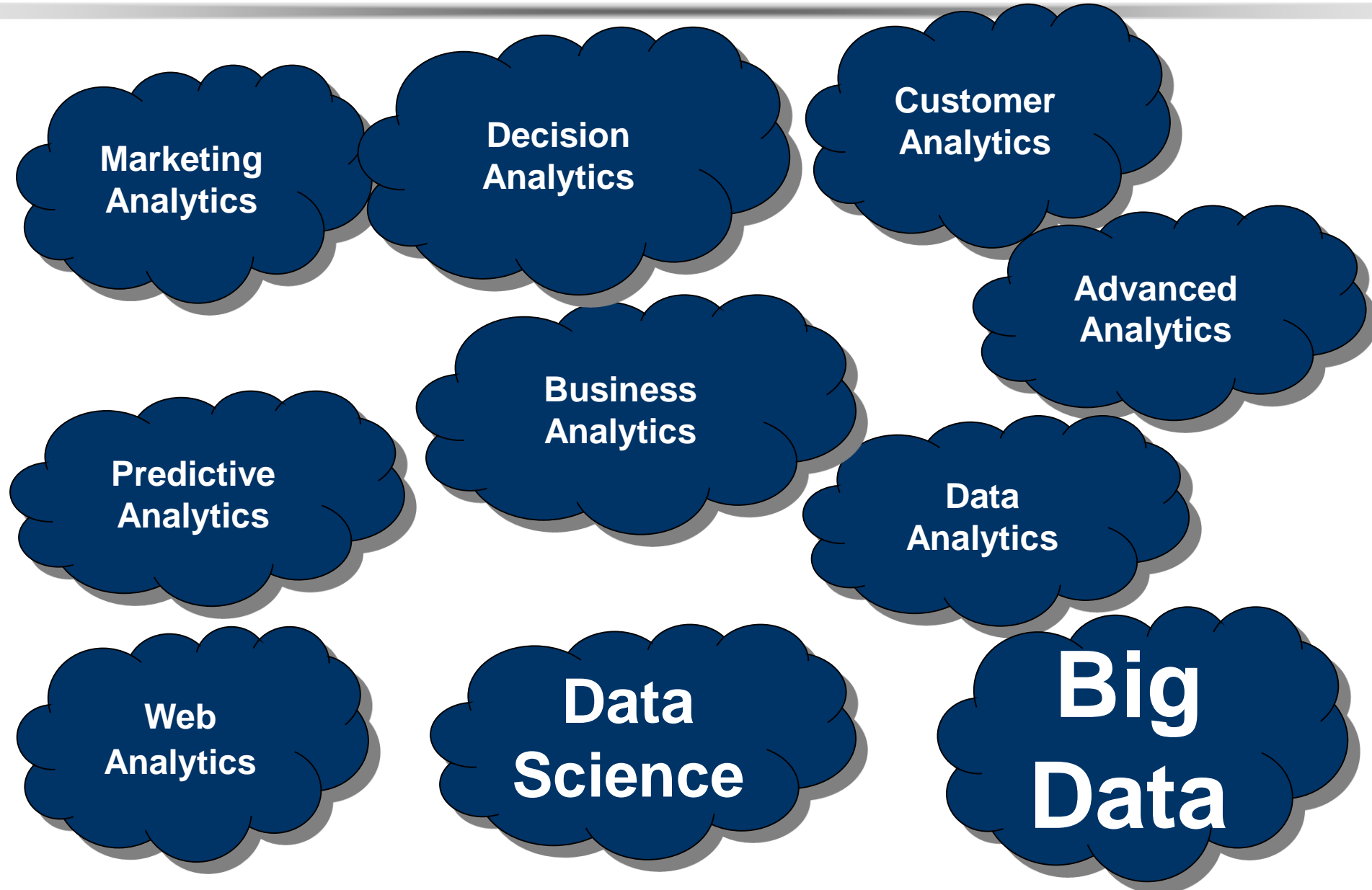
- „überall“
- Verteilte Systeme und Quellen
- Zusammenfassung, Zusammenführung?

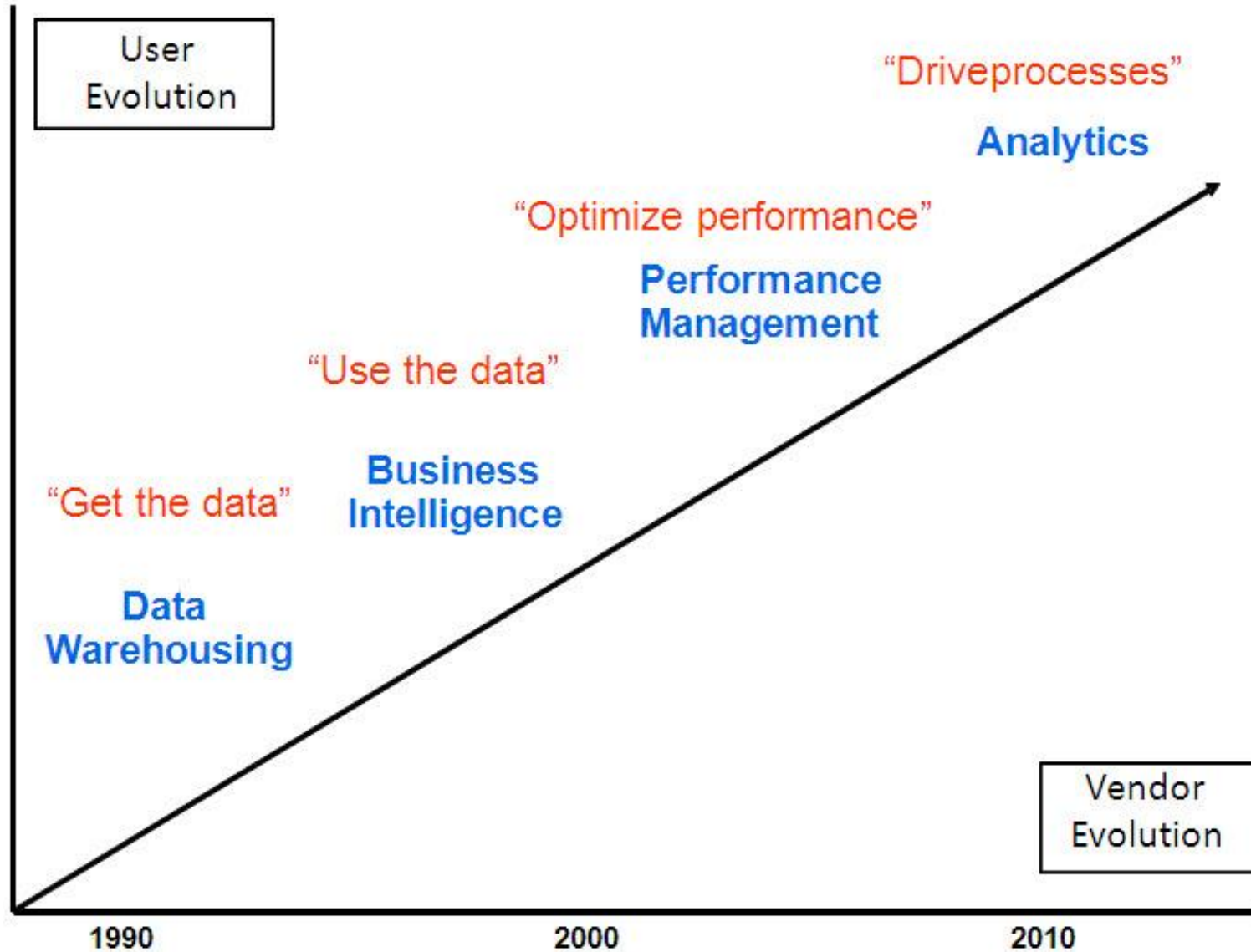
■ Zugriff

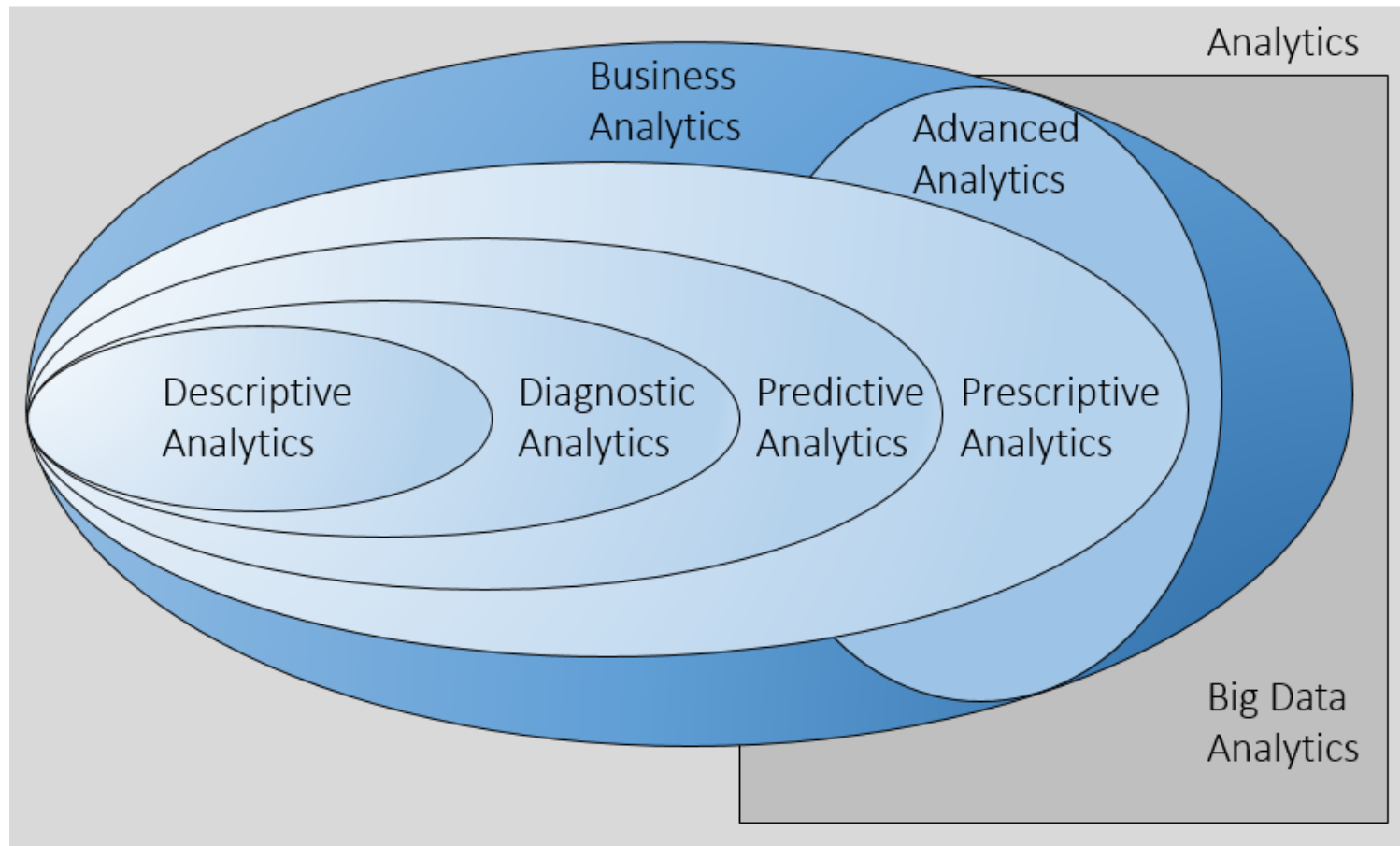
- technische und/oder administrative Hürden

- Entscheidungen werden getroffen auf Basis von
 - ❑ Fakten, Daten
 - ❑ Simulationen, Modellen
 - ❑ Intuition, ...
 - ❑ Gruppendiskussionen, -entscheidungen
- Traditionell wird BI als Entscheidungsunterstützung(ssystem) [EUS; oder: Decision Support System, DSS] angesehen. In Abgrenzung zu modellgetriebenen DSS spricht man hier von datengetriebenen DSS.
- Problem:
 - ❑ Lücke zwischen Daten und Wissen
 - ❑ Intuition bei dispositiver und auch operativer Arbeit
 - ❑ Feedback-Systeme? Anpassungen? Verbesserungen?
 - ❑ Voraussetzung: gute analytische Modelle und Verarbeitung

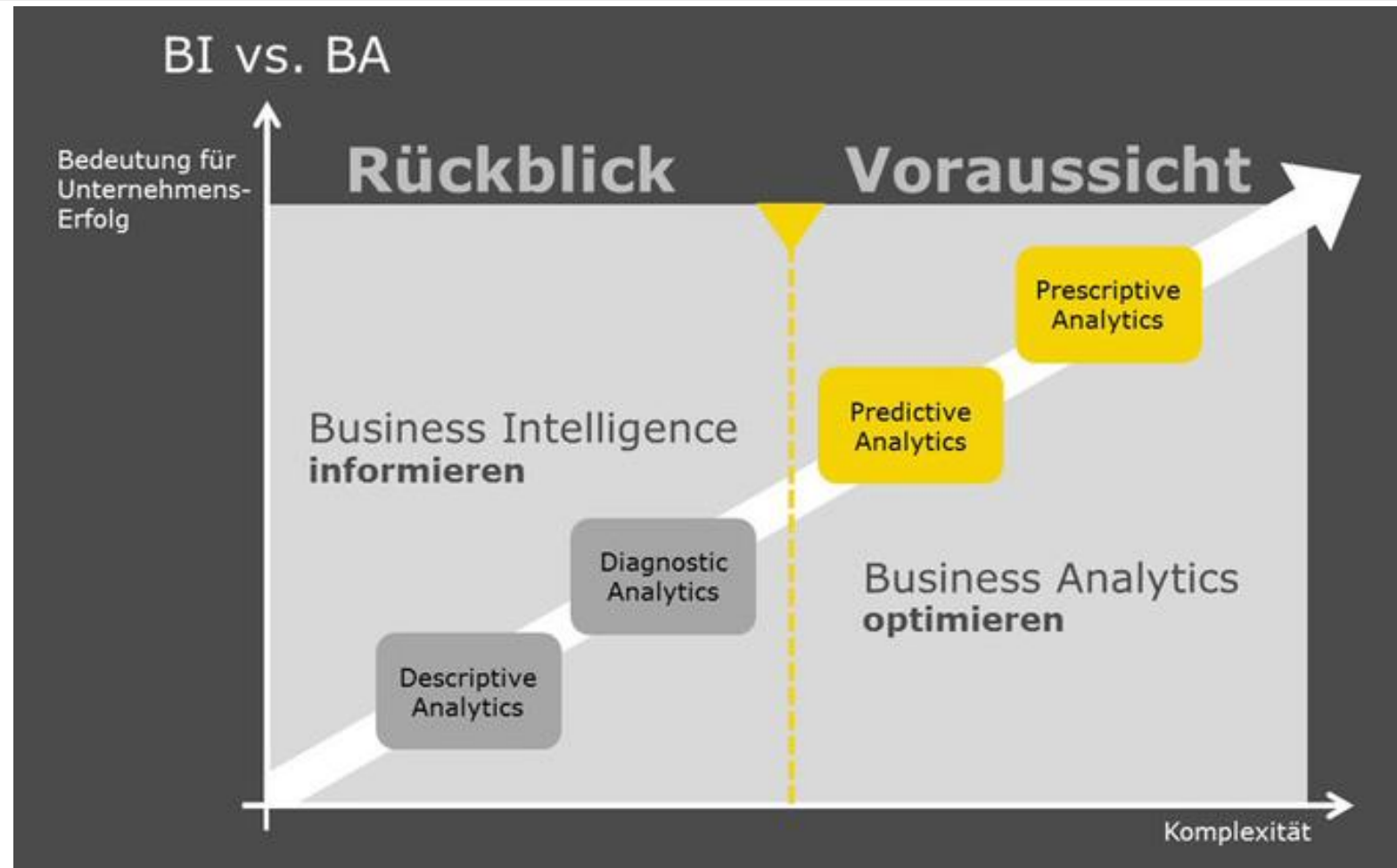








- Verarbeitung digitaler Daten
- Nutzung von statistischen Methoden und quantitativen Modellen
- Mustererkennung durch analytische Datenbetrachtung
- Prognose

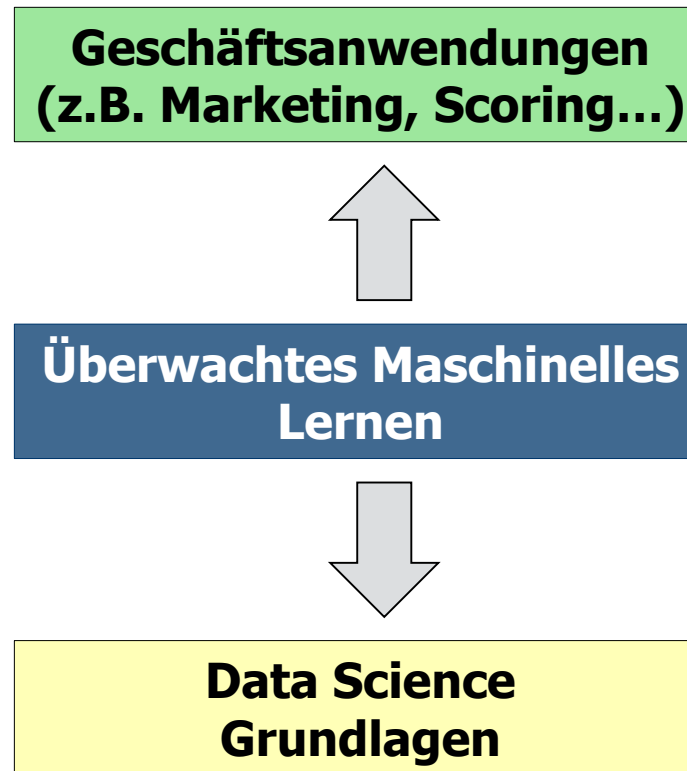
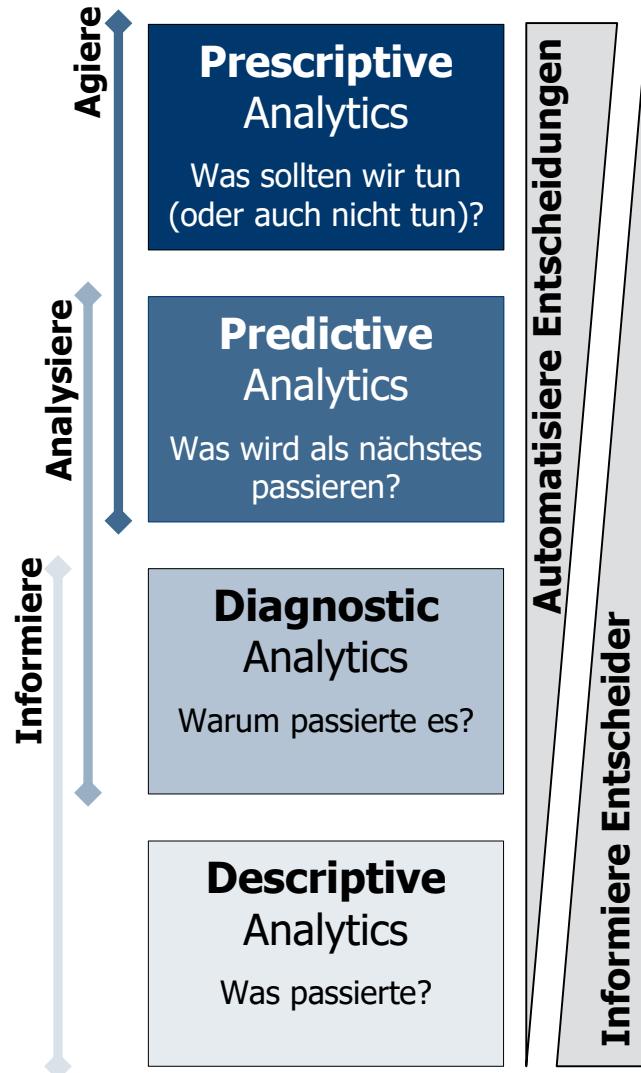


- BI: eher (noch) ausdrücklicher Bezug zu klassischen IT-Themen wie Datenintegration und -management
- BA: Verstärkung der Sicht zur eigentlichen Auswertung durch Fachanwender

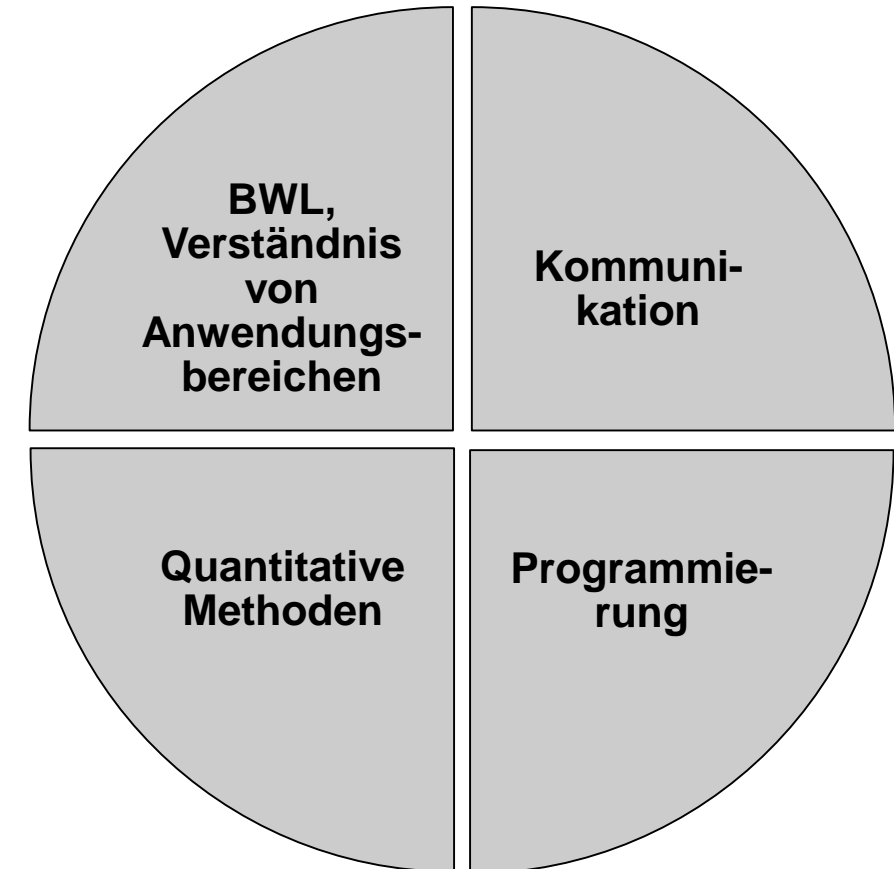
- Risikomanagement; oder viel besser: Abschätzung von Auswirkungen und Reaktionen
- Betrugserkennung
- Vorhersage von Wartungszeitpunkten
- Absatzprognose → Disposition;
- Kapazitätsauslastungsprognose
- ...

Wetteränderung beeinflusst Abverkauf von Garnelen

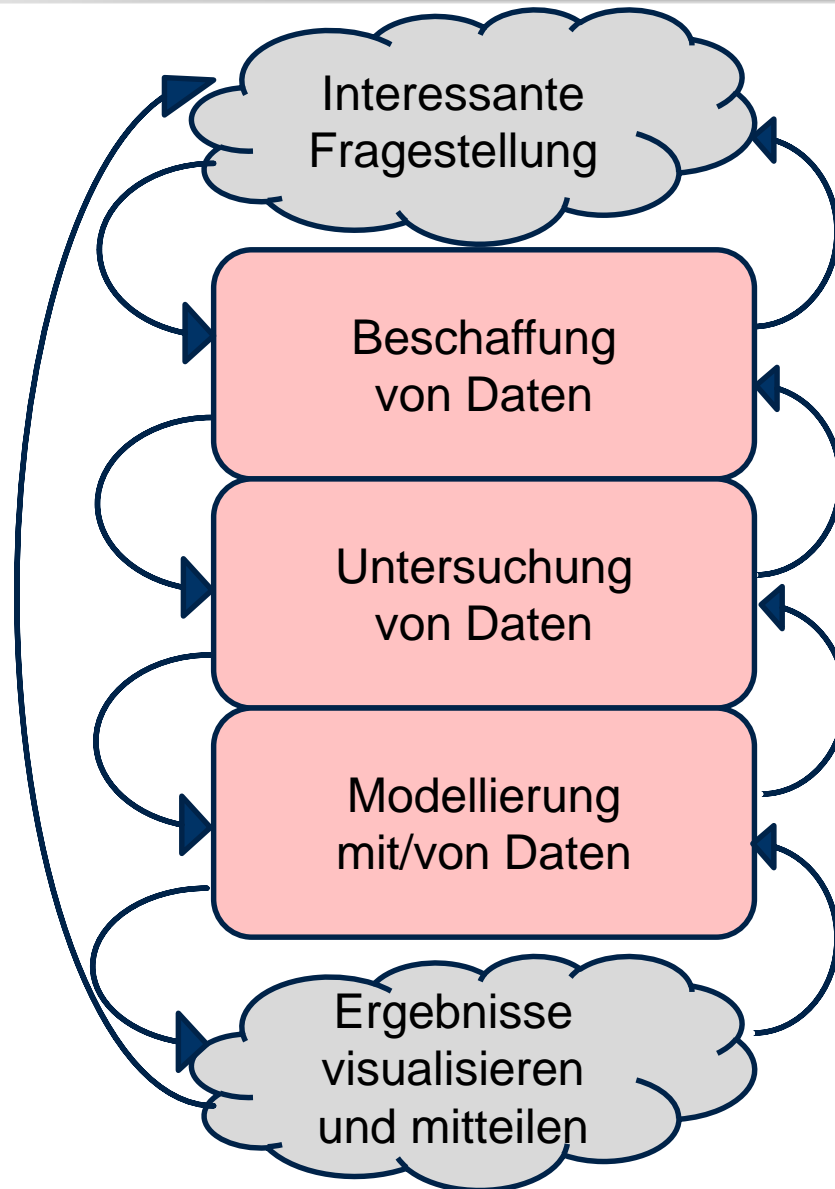




Business Analytics „Skill Set“



- Ist alles „fancy“ und „cool“? Was bedeutet „Powered by AI“?
- Müssen Dinge alle komplex sein? Ist Data Science zwingend komplex?
- Dinge, die „fancy“ und „komplex“ sind, erlangen Aufmerksamkeit! ... Marketing...
- Es geht nicht um das Finden möglichst vieler passender Convolutional Neural Networks oder um „deepest learning“ oder um „künstliche Intelligenz“ auf höchsten Leveln.
- Bei Data Science geht es vor allem um das Benutzen des passenden Tools, der passenden Methode für das Lösen eines Problems, das Erledigen einer Aufgabe in der realen Welt.
- Ein guter Data Scientist schafft das regelmäßig:
 - Konzentrieren auf Praxisprobleme
 - Finden/Nutzen des am besten geeigneten Werkzeugs (oftmals ist das einfachere unter vielen das bessere)
 - Lösungen erzeugen und bewerten



- Wissenschaftliches oder praktisches Ziel?
- Was würde man tun, wenn man alle Daten hat?
- Was soll geschätzt/prognostiziert werden?

- Wie sind Daten entstanden und gesammelt?
- Wie sind Daten gefiltert, zusammengesetzt?
- Welche Daten sind relevant?
- Gibt es Datenschutzprobleme?

- Visualisierung von Daten
- Gibt es Anomalien?
- Gibt es Muster?

- Erstellen eines Modells
- Anpassen des Modells
- Evaluation/Validierung des Modells

- Lernerfolg?
- Sind Ergebnisse sinnvoll und hilfreich?
- Gibt es eine passende „Geschichte“ dazu?

■ Phasen der Entscheidung

□ Intelligence

- Definition des Problems
- Ursachen, Ziele, Nebenbedingungen

□ Design

- Entwerfen von Lösungsalternativen

□ Choice

- Analyse und Bewertung der Handlungsalternativen
- Auswahl
- Implementierung

□ Revise

- Kontrolle
- Revision

■ „Programmierbare“ Entscheidungen

□ Routineaufgaben

□ Wiederkehrend

□ Definierte Prozeduren

■ „Nicht-Programmierbare“ Entscheidungen

□ Neu, unstrukturiert

□ Keine bekannte Methoden

□ Problemadäquater Lösungsansatz ist zu entwickeln

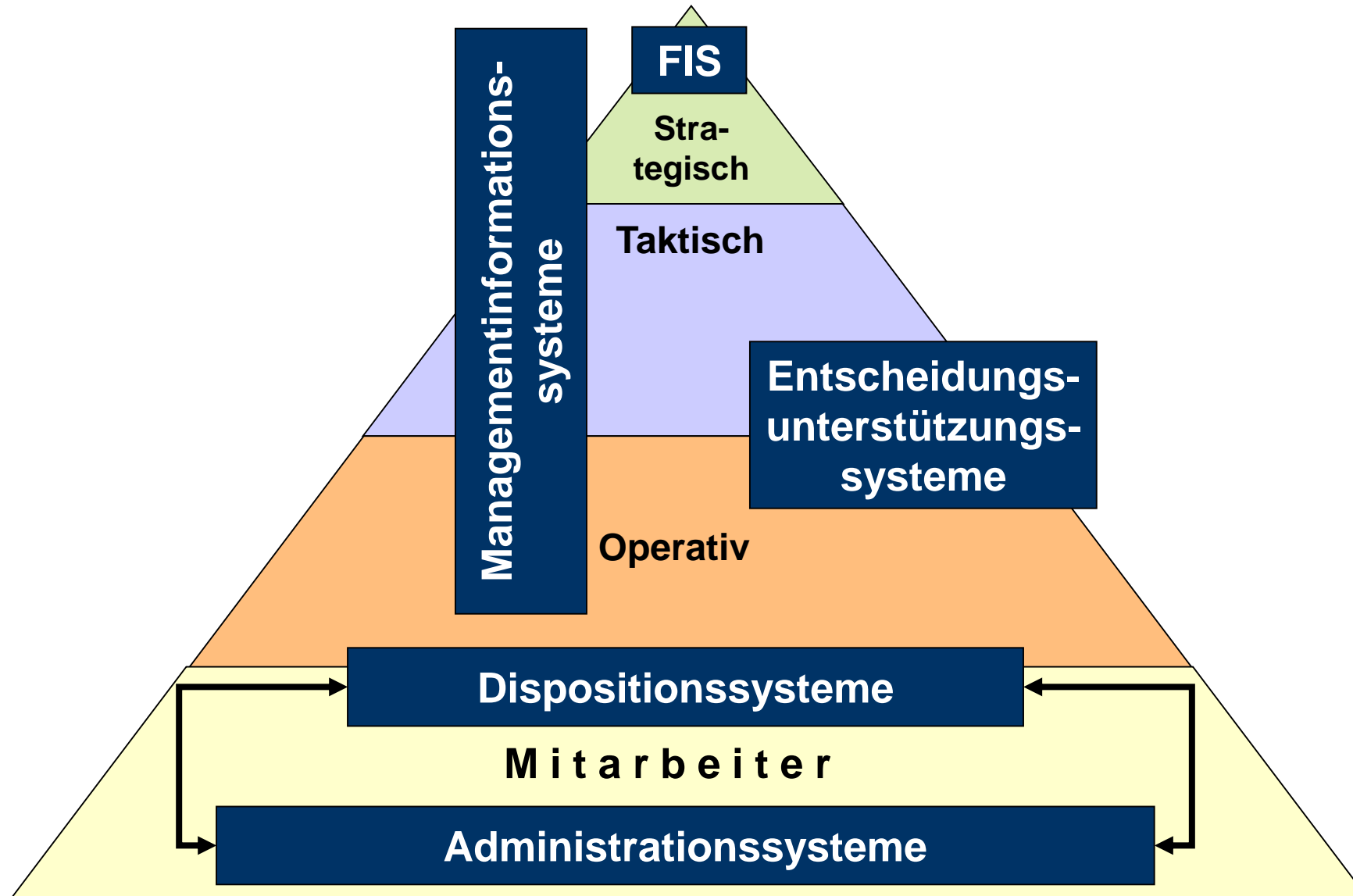
■ Keine klare Trennbarkeit

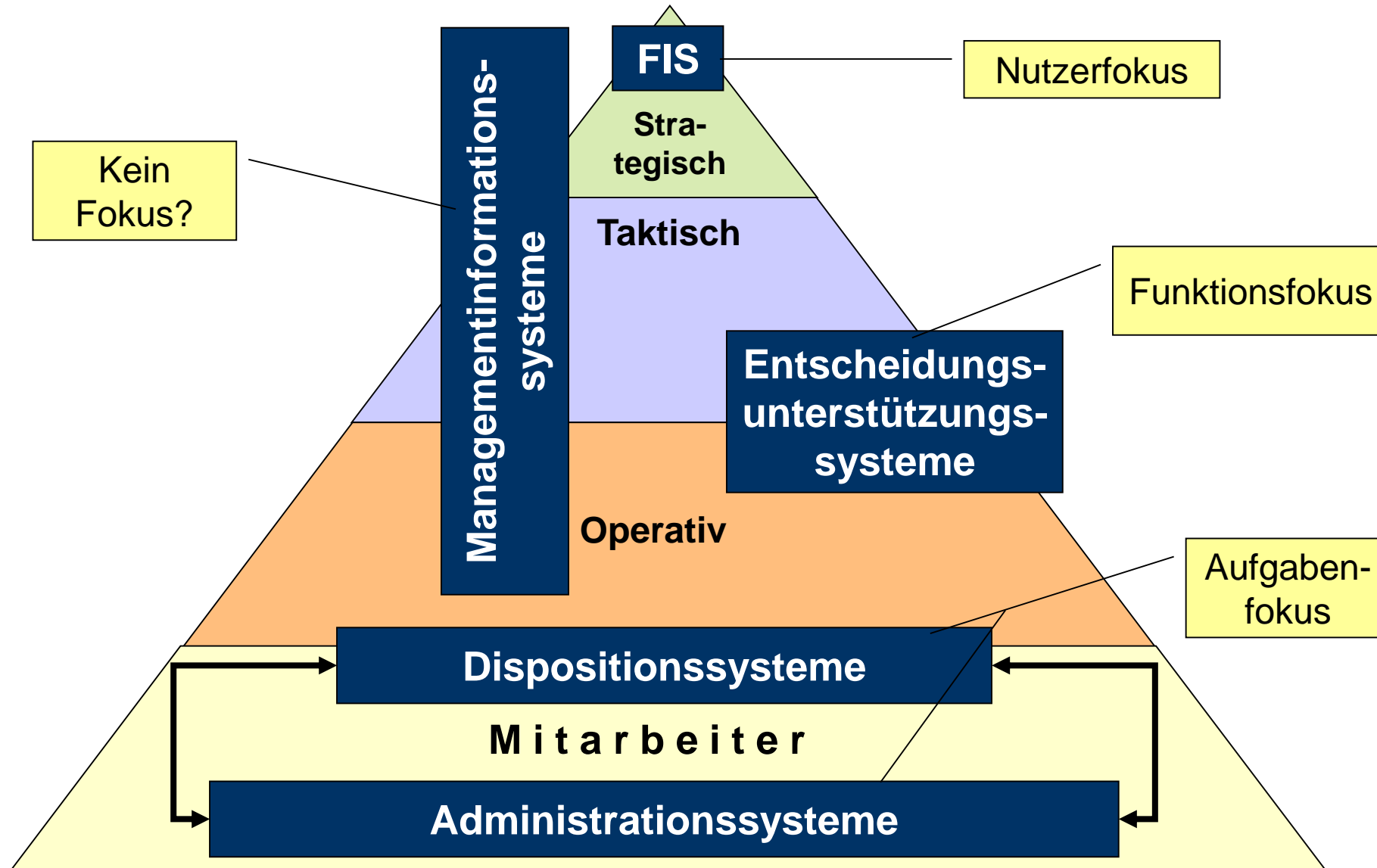
■ Kontinuum von programmierbaren bis zu nicht-programmierbaren Entscheidungen

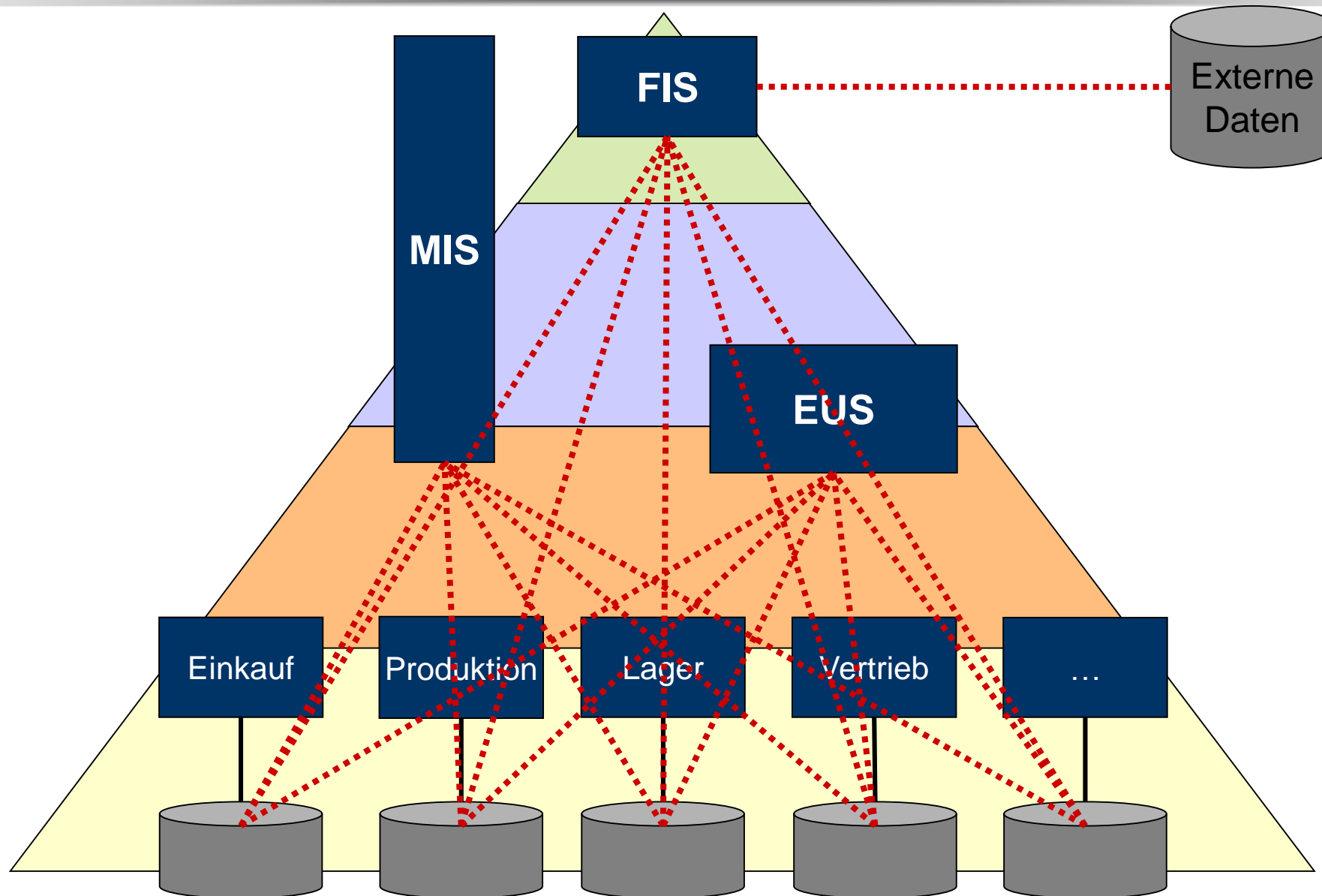
	Operativ	Taktisch	Strategisch
Strukturiert			
Semi-Strukturiert			
Unstrukturiert			

- Kombination von Simons Prozess rationaler Entscheidungen und Anthony's Managementhierarchie
 - Programmierbar → Strukturiert
 - Ein Entscheidungssituation ist strukturiert, wenn alle Phasen des Entscheidungsprozesses strukturiert sind.
 - Standardisierte Methoden
 - Klare Zielvorgabe
 - Definierte In-/Outputs

	Operativ	Taktisch	Strategisch
Strukturiert	Auftrags- erfassung, Bestellwesen	Absatzplanung Kurzfristprognosen	Investitionsplanung Finanzierung
Semi-Strukturiert	Ablaufplanung, Lagerhaltung	Projektplanung Layoutplanung Anreizsysteme	M&A Neuproduktplanung Standortplanung
Unstrukturiert	Kreditbearbeitung	Recruiting von Führungspersonal	Corporate Governance FuE







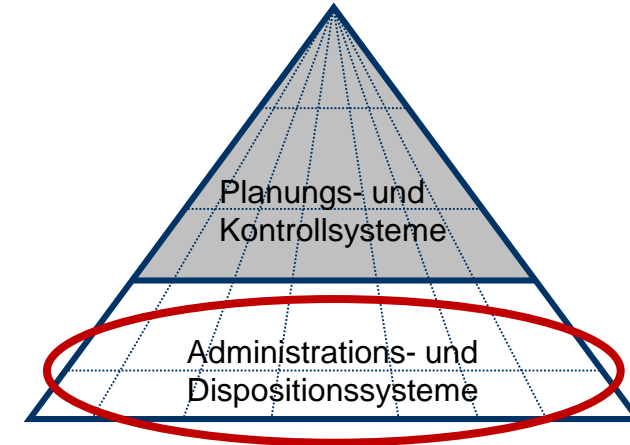
	Operativ	Taktisch	Strategisch
Quelle	Intern	↔	Intern und Extern
Umfang	Gering, klar definiert	↔	Weit
Aggregationsniveau	Detailliert	↔	Aggregiert
Zeitbezug	Vergangenheit	↔	Zukunft
Aktualität	Real-time	↔	Verzögerungen akzeptabel
Genauigkeit	Hoch	↔	Schätzungen akzeptabel
Nutzungsfrequenz	Hoch	↔	Unregelmäßig, Ad hoc

■ Beispiele

- ❑ Flugbuchungssystem
- ❑ Verarbeiten von Bestellungen
- ❑ Ein-/Verkauf von Waren

■ Charakterisierung

- ❑ Hoher Parallelitätsgrad (Concurrency Control)
- ❑ Zugriff auf jüngsten, aktuellsten Datenbestand
- ❑ Viele (Tausende pro Sekunde) kurze Transaktionen
- ❑ Sehr kurze Antwortzeiten gewünscht (ms bis s)
- ❑ Bearbeitung kleiner Datenvolumina pro Transaktion; Zugriff auf begrenzte Datenmenge
- ❑ „Mission-critical“ für das Unternehmen
- ❑ Hohe Verfügbarkeit muss gewährleistet sein
- ❑ Datenspeicherung
 - Normalisierte Relationen (möglichst geringe Update-Kosten)
 - Nur wenige Indizierungen



■ Technisch

□ Managementinformationssystemen

- Automatisierung des Standardberichtwesens
- Fehlende Interaktivität und Dialogorientierung
- Datenversorgung vs. Entscheidungsunterstützung

□ Entscheidungsunterstützungssysteme

- Lokale Ausrichtung der Werkzeuge erschwert Integration in unternehmensweites IT-Konzept
- Ggf. widersprüchliche Ergebnisse, z.B. zwischen Abteilungen

□ Führungsinformationssysteme

- Beschränkung auf Top-Management
- Zunehmende Entscheidungskompetenz auf unteren Managementebenen (z.B. „Lean-Management“)

■ Organisatorisch

□ Aufbauorganisation

□ Ablauforganisation

- Partielles Scheitern der o.g. Systeme führte zu Weiterentwicklungen, die heute unter dem Begriff der Analytischen Informationssysteme geführt werden.
- Hierzu gehören:
 - Data Warehouse Systeme
 - Berichtssysteme bzw. Reportgeneratoren
 - Online Analytical Processing Systeme (OLAP)
 - Data Mining
- Näherungsweise gilt:
BI-Systeme = Analytische Informationssysteme



■ Datenbanken

- ❑ Effektive und effiziente Speicherung großer Datenmengen
- ❑ Klassische relationale Datenbankmanagementsysteme
- ❑ Spaltenorientierte und NoSQL Datenbanken

■ Tabellenkalkulation

- ❑ Interaktive Auswertung und Darstellung (z.B. Diagramme) kleiner bis mittelgroßer Datenmengen
- ❑ What-If-Rechnungen
- ❑ Pivot-Tabellen

■ Kommunikations- und Kooperationssysteme

- ❑ Rechnernetze
- ❑ Asynchrone und synchrone Kommunikation
- ❑ Workgroup Management Systeme für Teamarbeit

■ Führungskräfte **und** Fachanwender mit analyseorientiertem Tätigkeitsfeld

■ Idealtypische Anwendergruppen

□ Informationskonsumenten

- Standardisierte Sichten auf definierten Datenbestand (→Informationsbedarfsanalyse)
- Periodische Aktualisierung
- Bereitstellung
 - Einfacher, ortsungebundener Zugriff mit verschiedenen Endgeräten
 - Vergleichsdarstellungen mit Abweichungen (Soll/Ist, Jahr/Vorjahr)
 - Zeitreihen

□ Analytiker

□ Spezialisten



- Führungskräfte **und** Fachanwender mit analyseorientiertem Tätigkeitsfeld
- Idealtypische Anwendergruppen
 - Informationskonsumenten
 - Analytiker
 - Lösung semi- bis unstrukturierter Probleme
 - Freie Navigation durch den Datenraum, selbstständige Auswahl/Kombination verfügbarer Daten (Aggregation/Disaggregation)
 - Flexible Darstellungsmöglichkeiten (z.B. Diagramme)
 - Einfache Analysemethoden (Durchschnittsbildung, Anteilsberechnung, ...)
 - Spezialisten



■ Führungskräfte **und** Fachanwender mit analyseorientiertem Tätigkeitsfeld

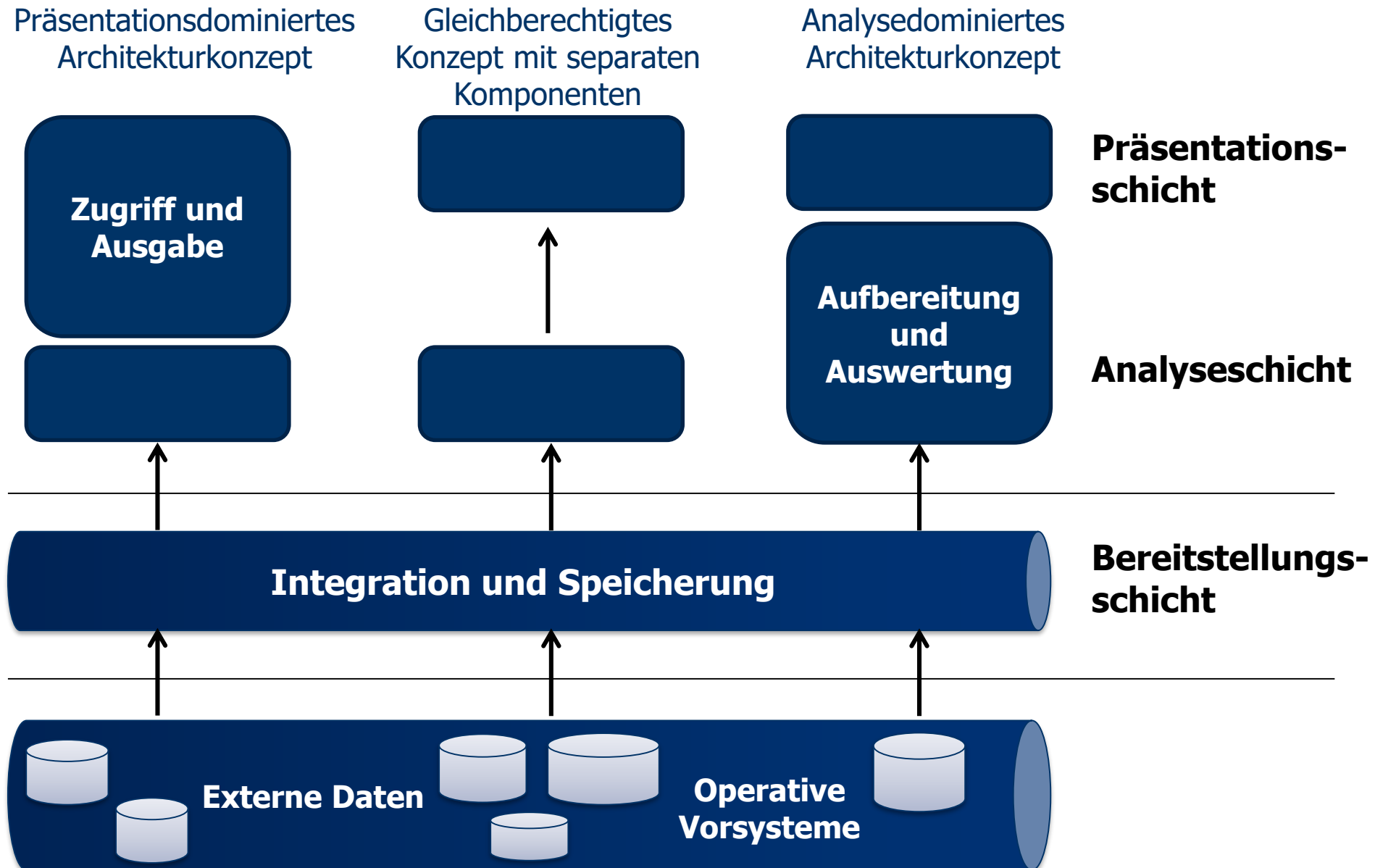
■ Idealtypische Anwendergruppen

- ❑ Informationskonsumenten
- ❑ Analytiker
- ❑ Spezialisten
 - Komplizierte Aufgabenstellungen
 - Ursache-Wirkungszusammenhänge aufdecken
 - Ziel-Mittelbeziehungen erklären
 - Anspruchsvolle Methoden / SW-Tools
 - Statistik
 - Operational Research
 - Ggf. eigene Lösungsentwicklung
 - Z.B. SQL-Programmierung

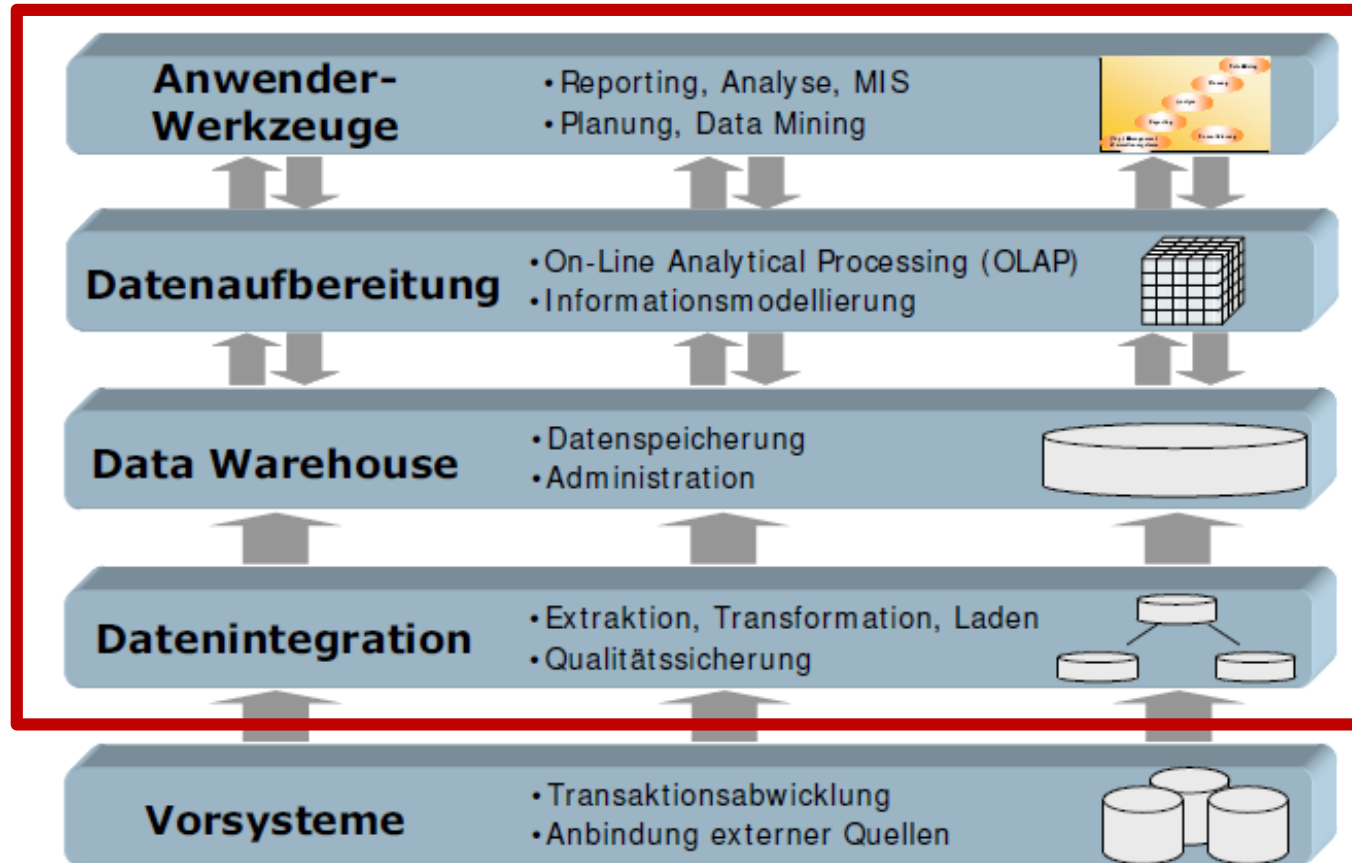


Hausaufgabe:

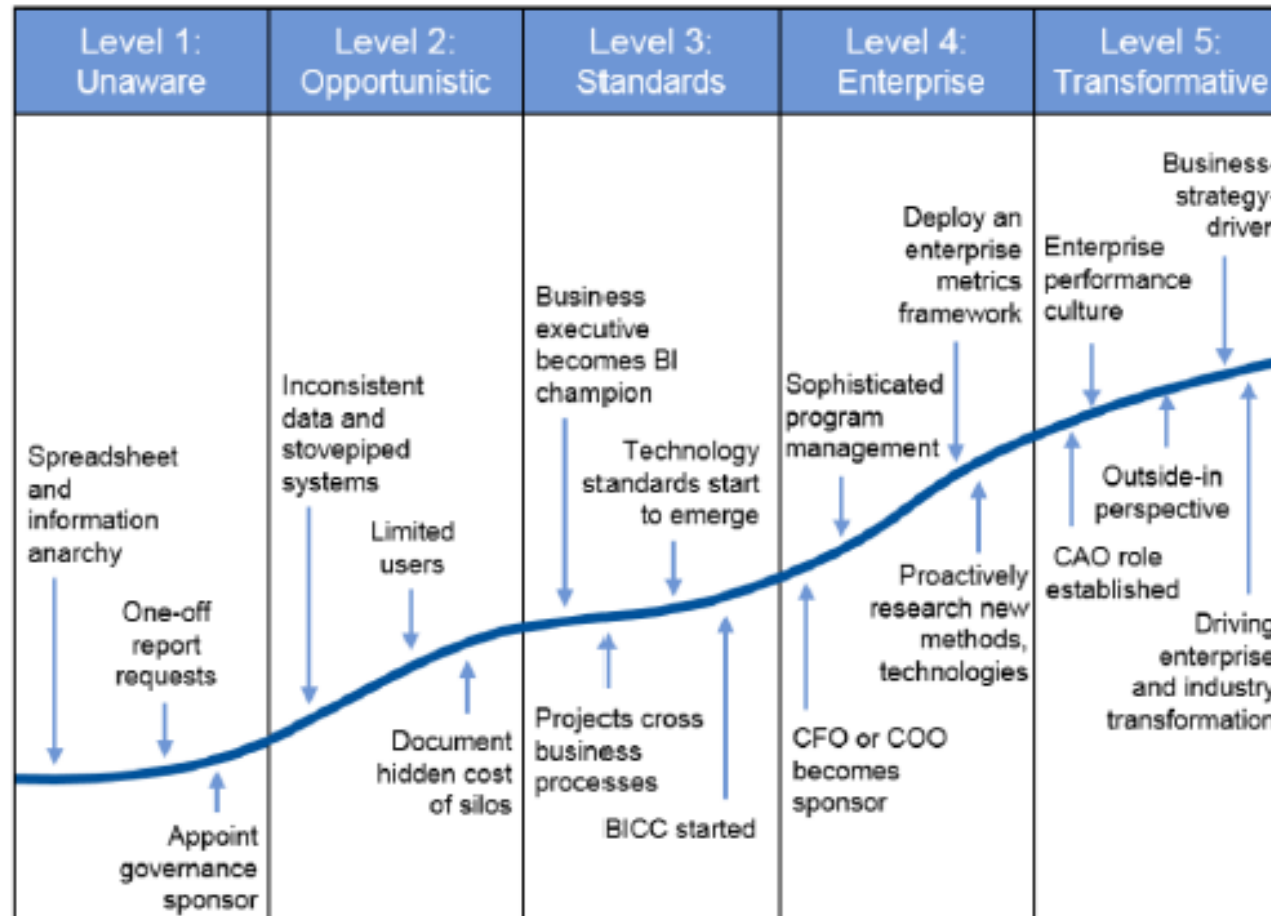
Recherchieren Sie zu „Self Service BI“ (Begriff, Inhalt, Unterschiede zu „herkömmlicher BI“, Besonderheiten...)



- Dispositive Informationssysteme zur Unterstützung von BI-Aufgaben und -Prozessen auf allen/ausgewählten Ebenen der BI-Referenzarchitektur



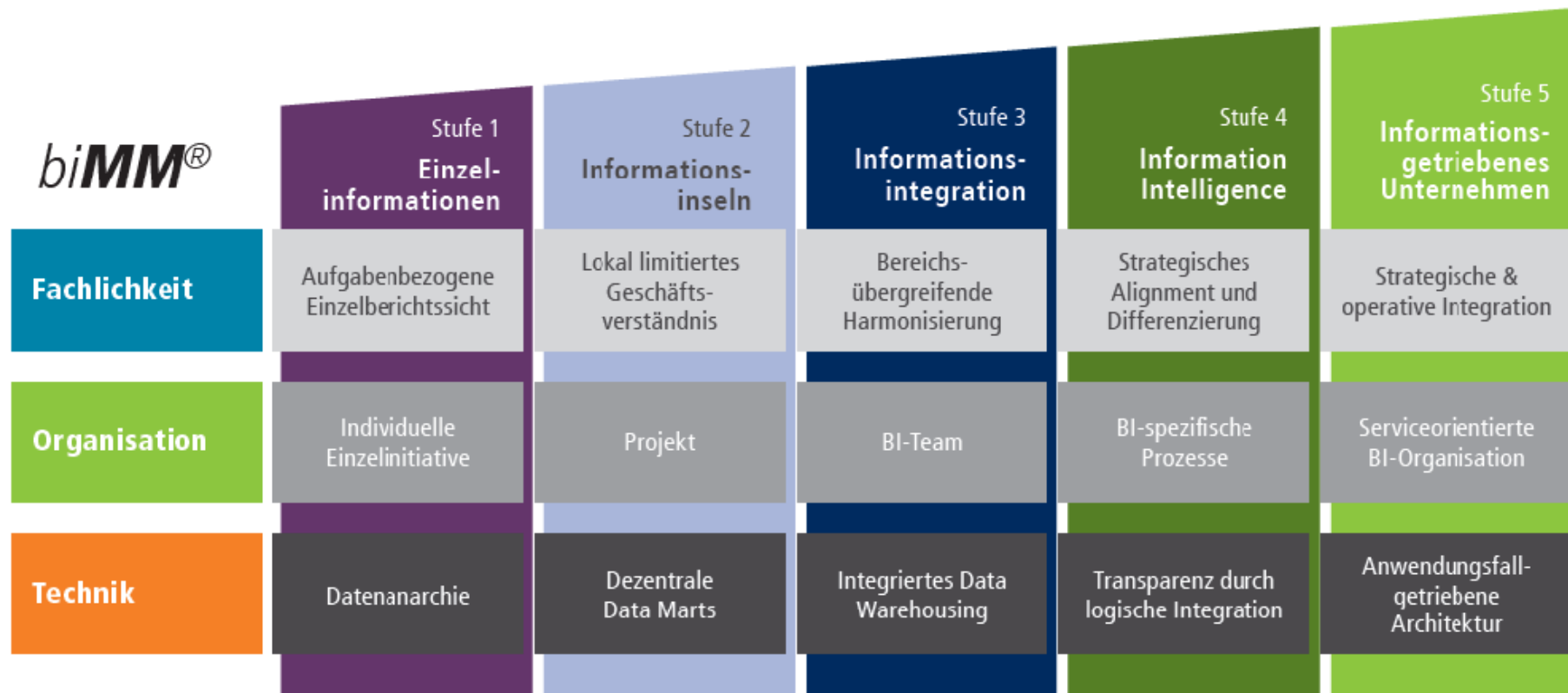
■ BI and Analytics Maturity Model (Gartner, 2015)

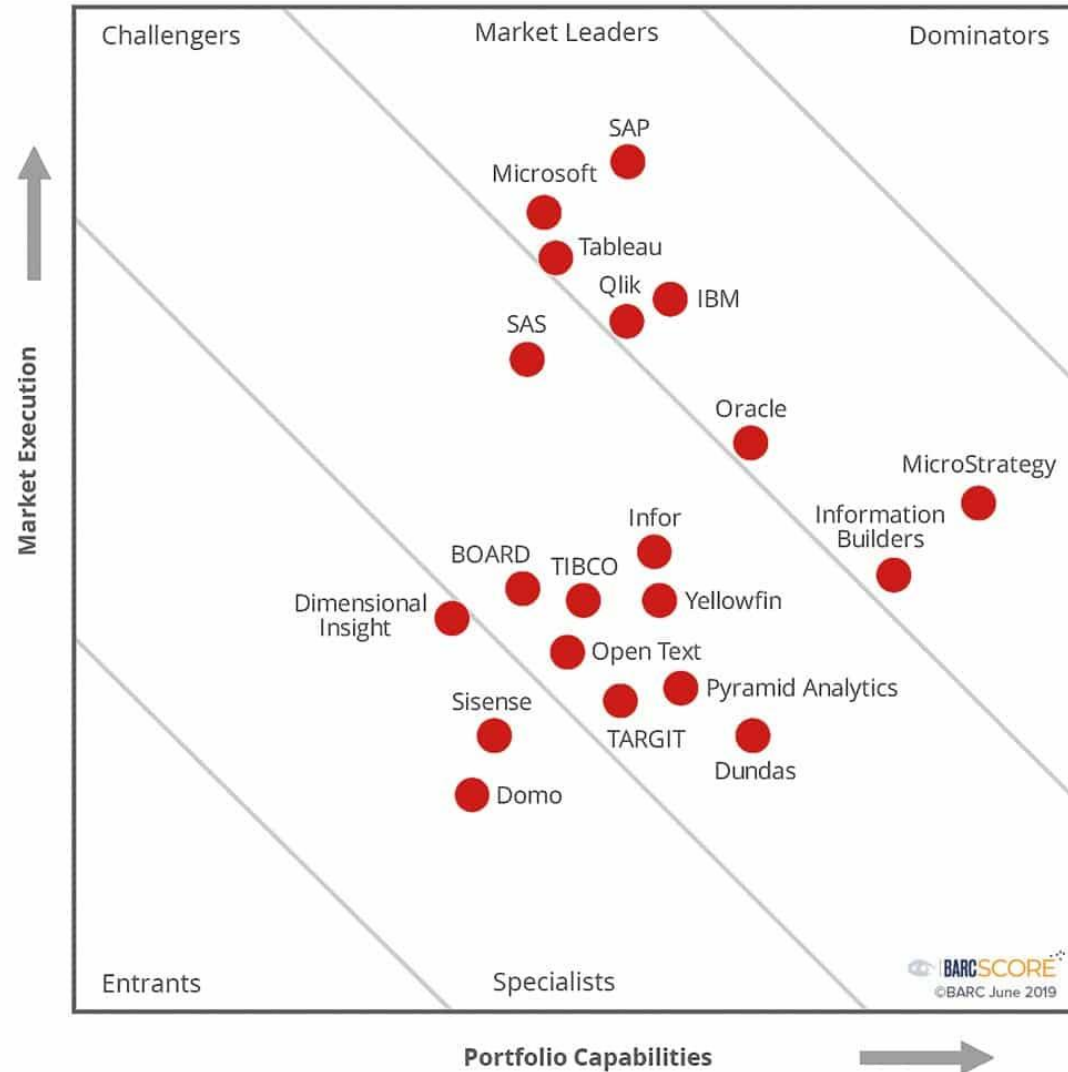


BI = Business intelligence

BICC = BI competency center

■ BI Maturity Model (Steria Mummert Consulting 2013)





Positionierung der Anbieter im BARC Score Enterprise BI and Analytics Platforms.

■ Schwierig

- ❑ BI-System vs. BI-Anbieter
- ❑ Viele Lösungen werben mit dem Label „Business Intelligence“
(und seit einiger Zeit auch mit „Big Data“)

■ Klar unterscheidbar

- ❑ Spezialanbieter
 - Lösungen für ausgewählte Teilbereiche der BI-Referenzarchitektur
- ❑ Ganzheitliche BI-Lösungen
 - Abdeckung des gesamten Spektrums von BI-Aufgaben
 - Datensammlung → Datenspeicherung → Datenbereitstellung
→ Datenanalyse
 - Beispiele: SAP, IBM...

■ Prinzipielles Vorgehen:

Eigene Lösungen und Zukäufe = (mehr oder weniger) integriertes BI Angebot

- ❑ SAPs OLAP/Reporting Lösung basiert z.T. auf Zukauf von Business Objects in 2008
- ❑ IBM BI Plattform basiert weitgehend auf Zukäufen

■ Weitere Anbieter ganzheitlicher BI-Lösungen

- ❑ Microsoft
- ❑ Oracle
- ❑ SAS
- ❑ ...

■ Spezialisierte Anbieter und Mischformen...

- Zahlreiche SW-Werkzeuge bietet Funktionalität, die auch im Rahmen von BI genutzt werden kann
- Beispiele
 - Crystal Reports, Zoho Reports, ...
 - MS Excel, ...
 - IBM SPSS Statistics, R, ...
 - Kennzahlen-Systeme, Controlling-Werkzeuge, ...
 - CRM Systeme, ...
- Solche SW wird gern als BI-System vermarktet
- Exakte Abgrenzung im Einzelfall schwierig
 - z.B. ODBC (Open Database Connectivity)-Funktionalität wird als ETL-Lösung deklariert
 - Welches BI-Verständnis (vgl. BI-Begriff) wird zugrunde gelegt?

■ Führungsaufgaben im Unternehmen

- ❑ Managementebenen
- ❑ Managementprozess
- ❑ Anforderung an IS zur Entscheidungsunterstützung

■ Klassische Konzepte von MUS

■ Business Intelligence

- ❑ Gegenstand und Begriff
- ❑ Nutzergruppen
- ❑ Architekturkonzepte
- ❑ Reifegradmodelle
- ❑ Systeme



„Begriffe ohne
Anschauungen sind leer,
Anschauungen ohne
Begriffe sind blind.“
(Kant)