

# Datenstrategie

IHK Webinar, 24.01.2024

Dr. Felix Böhmer,  
Head of AI & Data Analytics  
iteratec GmbH

WE MAKE **AI&DATA** WORK.



A smiling man and woman are shown in a purple-tinted setting. The man is on the right, wearing a dark t-shirt, and the woman is on the left, wearing glasses and a dark t-shirt. They appear to be looking at something together. The background is a solid purple color. There are several colorful geometric shapes: a yellow square in the top left, a white square and a cyan square below it, another cyan square to the right of the white one, a yellow square on the right side, a cyan square below it, and a black square and a white square at the bottom right.

iteratec

# iteratec auf einen Blick

500

Kolleg:innen

100

Studierende

27

Jahre Erfahrung

64

Mio. Euro Umsatz

7

Standorte



# DEVELOPING DIGITAL CHAMPIONS

# Breit aufgestellt

**Security**

**Web3**

**IoT**

**APM**

**AI&Data**

**27 Jahre Engineering Excellence**

# Agenda

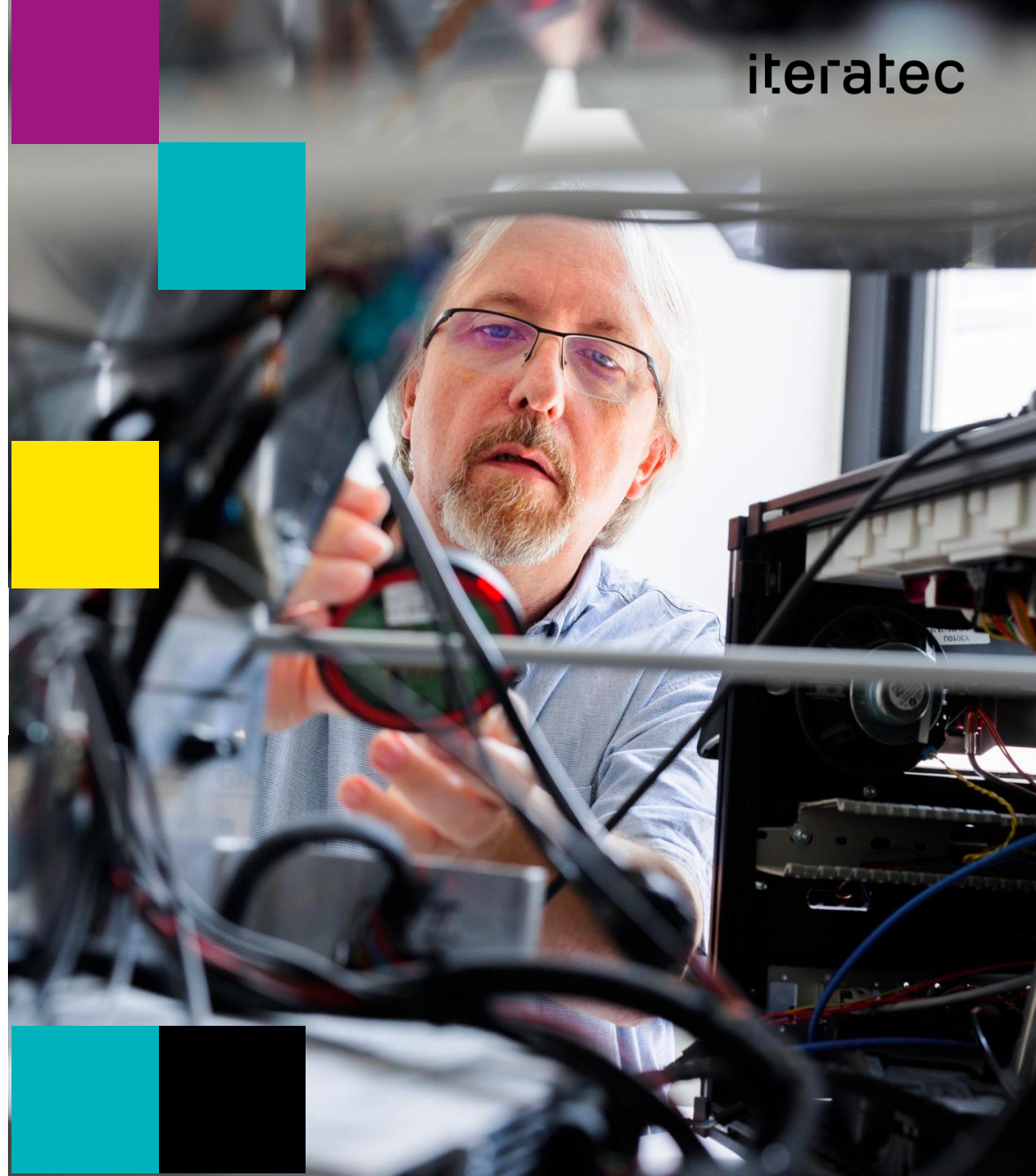
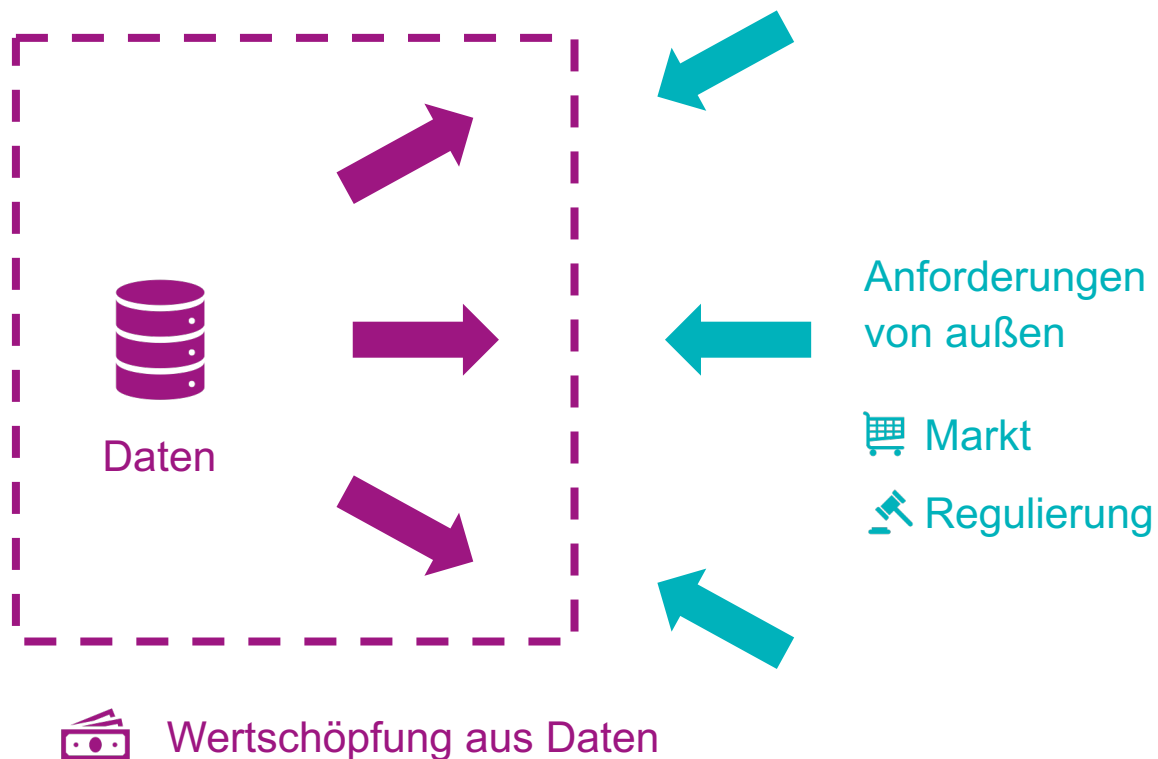
- 01** Motivation
- 02** Inhalte einer Datenstrategie
- 03** Datenhaltung
- 04** Wertschöpfung (AI & Analytics)
- 05** Top Down & Bottom Up
- 06** Disruptionen



# Datenstrategie Motivation



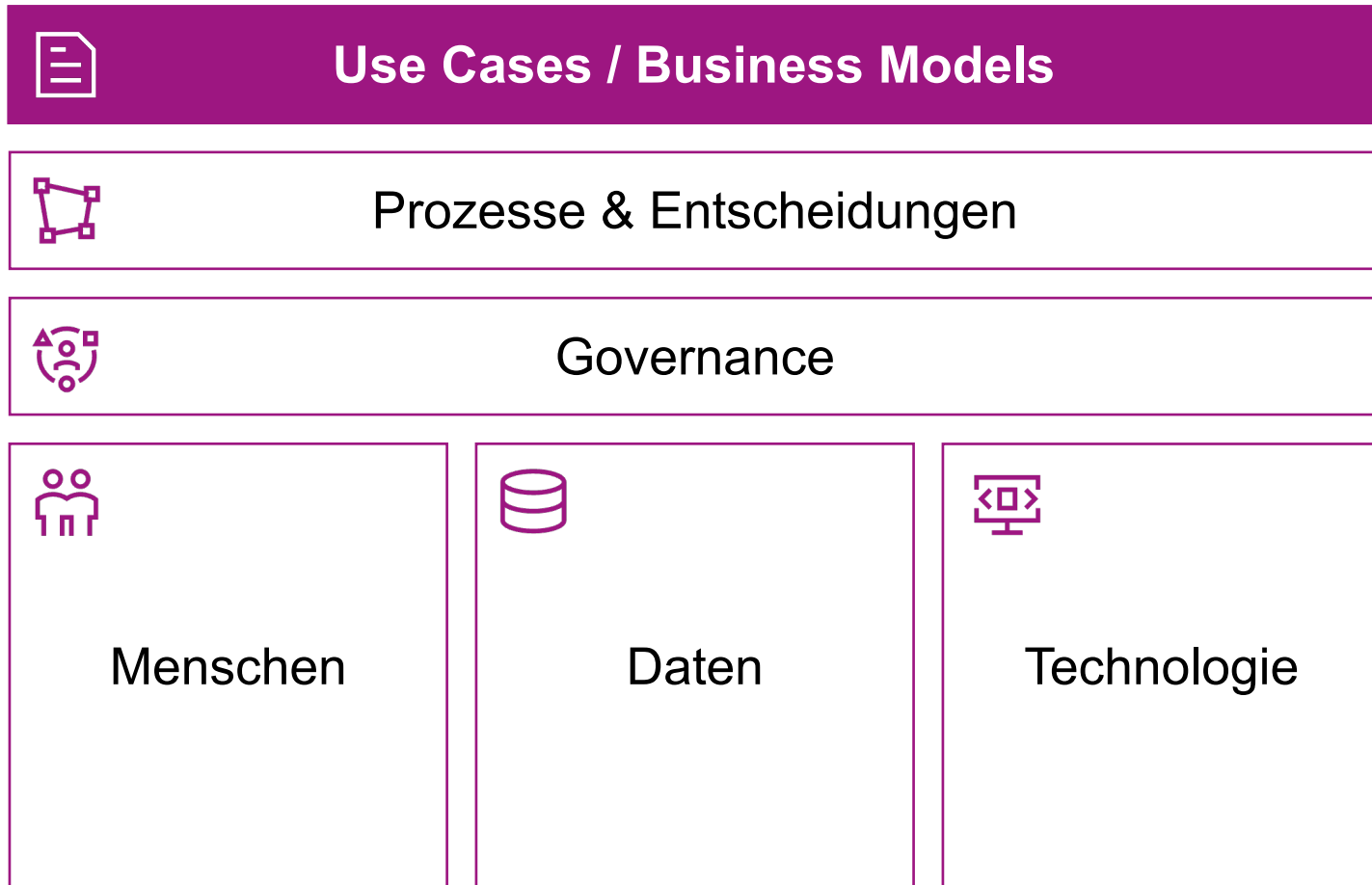
# Wozu brauche ich eine Datenstrategie?



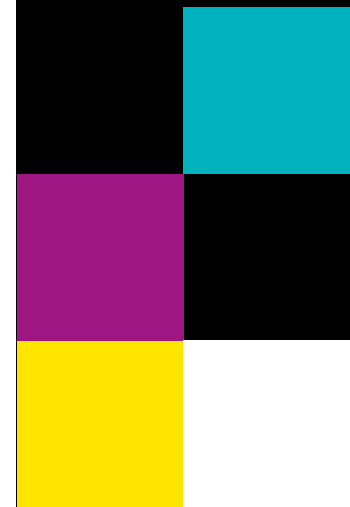


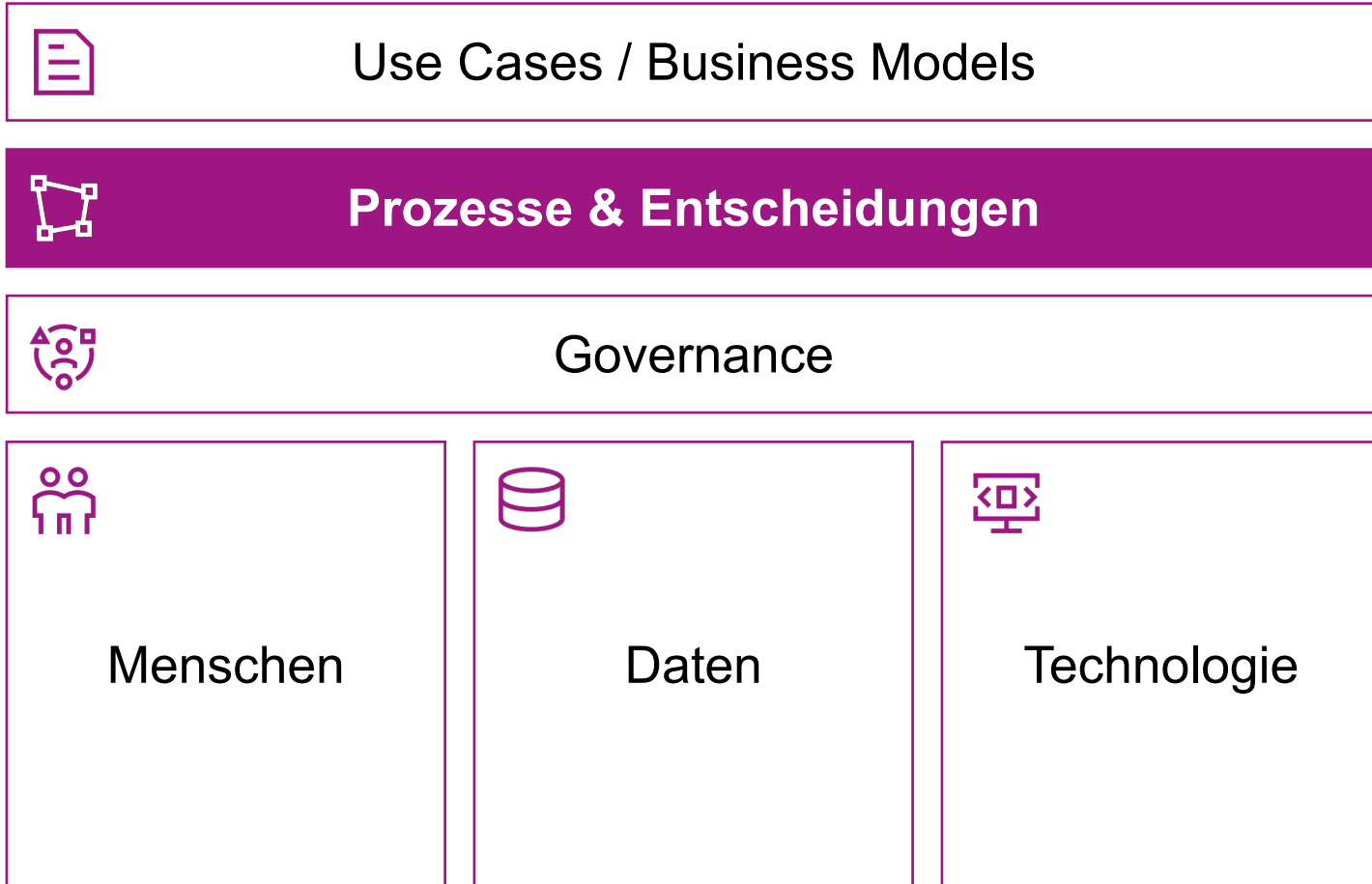
# Datenstrategie Inhalte





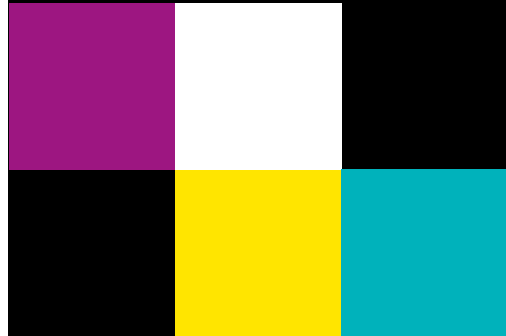
Ihr Business  
Datengetriebene  
Geschäftsmodelle

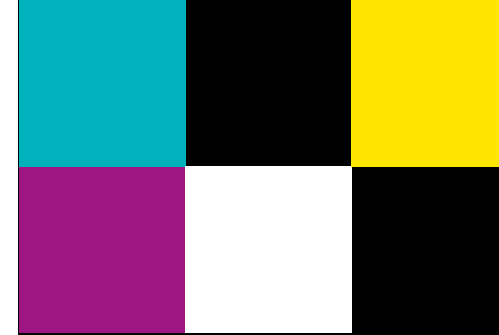




Effizienzsteigerung & Optimierung

Automatisierung von Entscheidungen





 Use Cases / Business Models

 Prozesse & Entscheidungen

 **Governance**

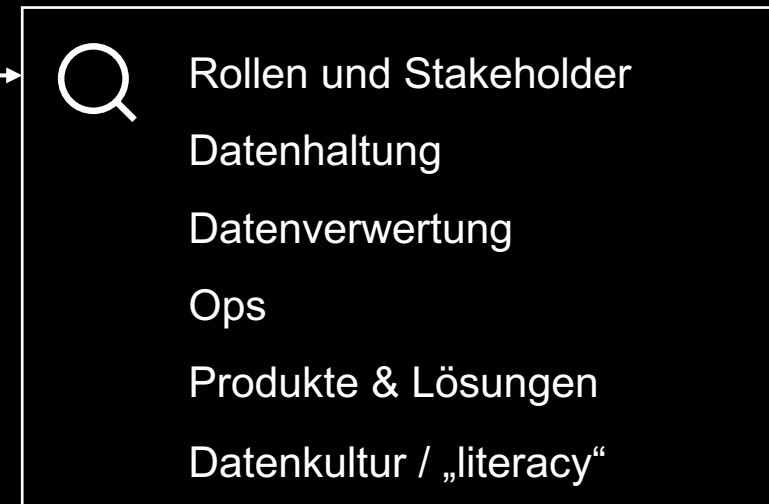
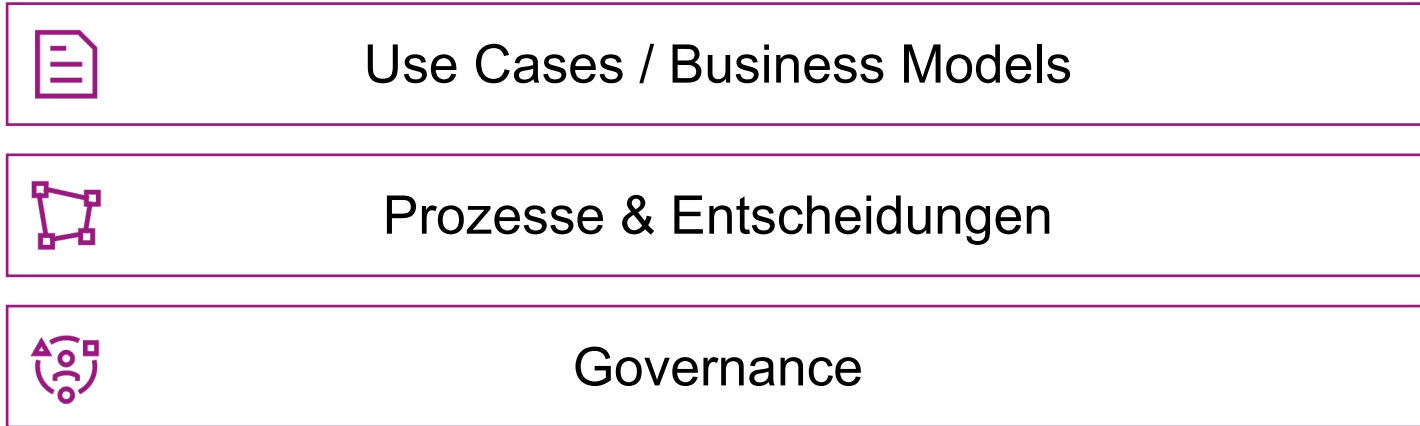
  
Menschen

  
Daten

  
Technologie

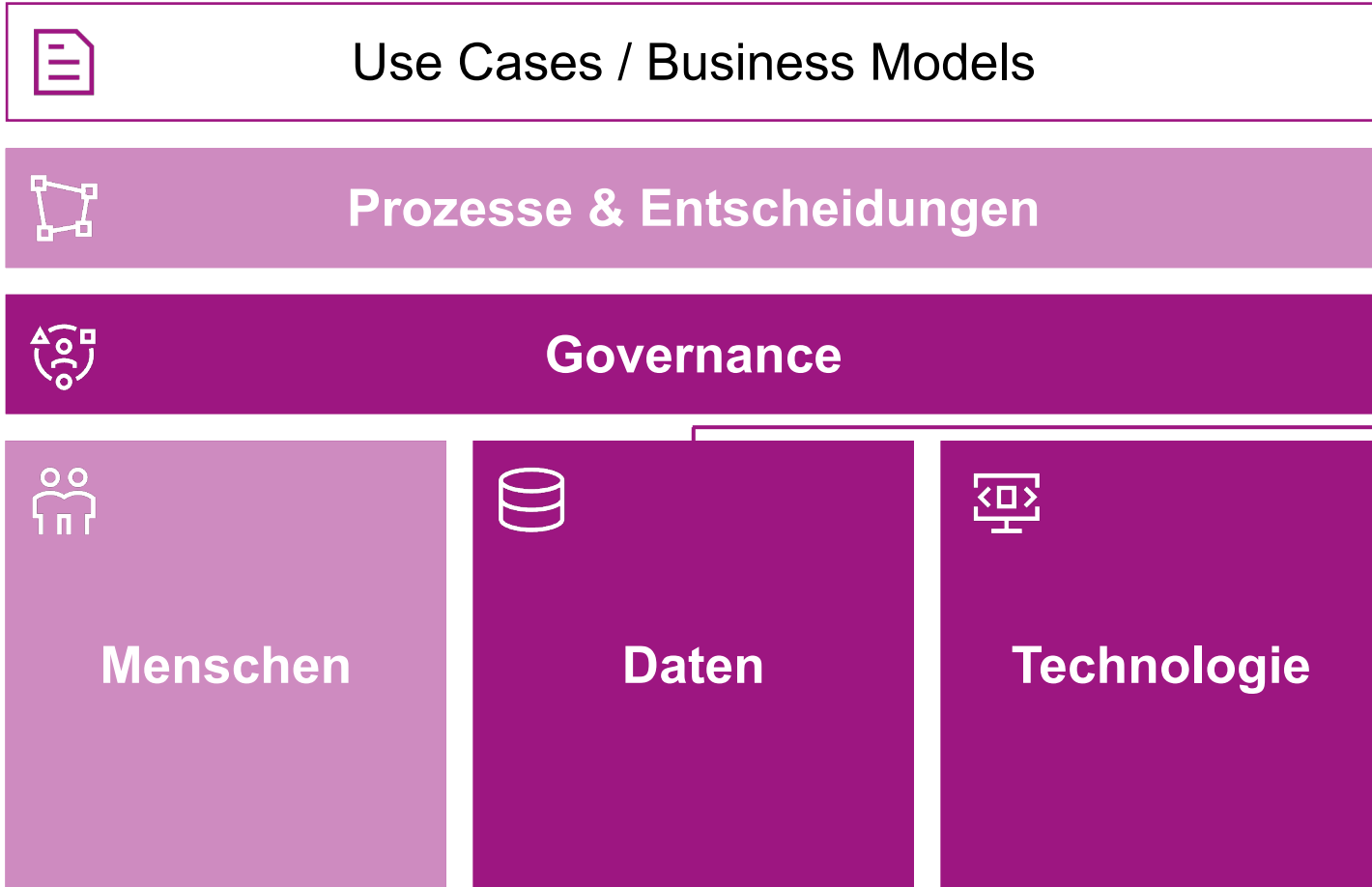
 Standards & Richtlinien für Daten und Zugriffe  
Regulatorische und gesetzliche Vorgaben





# Datenstrategie Datenhaltung

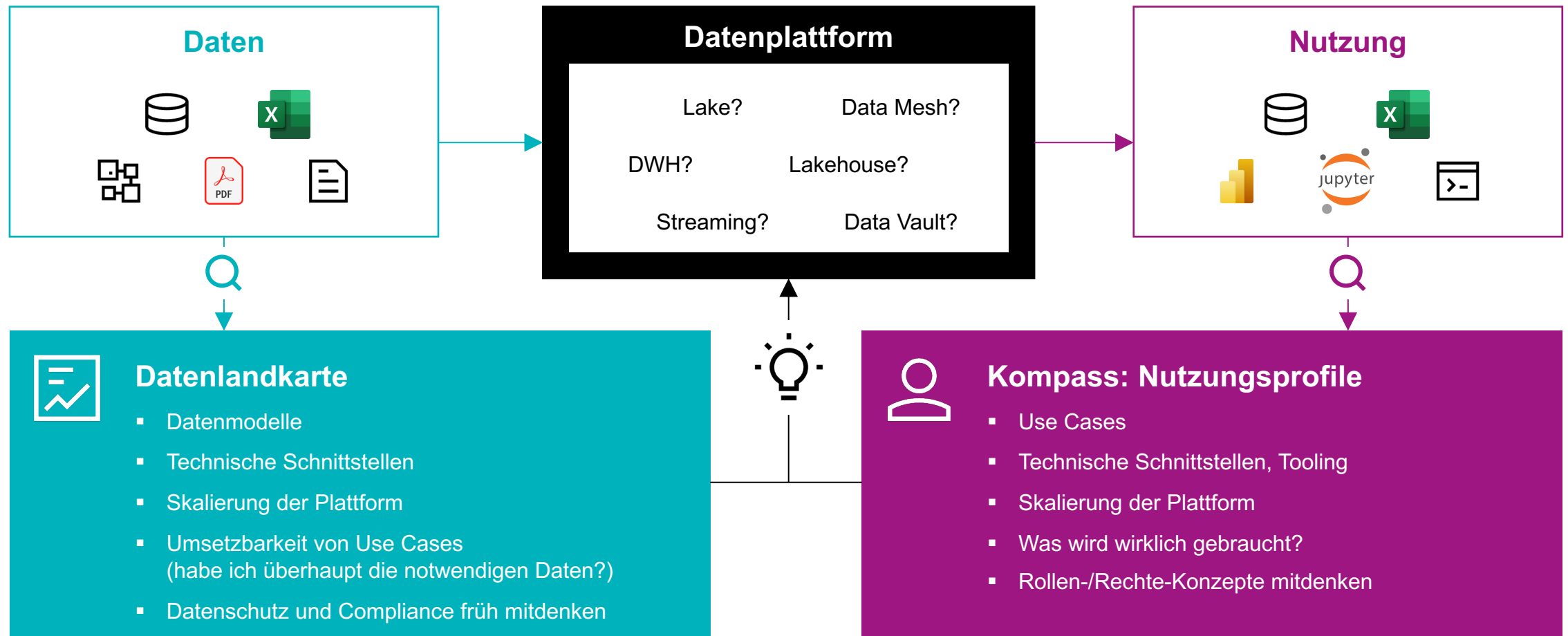




- Q Wie und unter welchen Voraussetzungen werden Daten gespeichert?
- Q Welche Daten und in welcher Form? Fachliche Datenmodelle?
- Q Welche Werkzeuge?

## DATENPLATTFORM

Wo fange ich an?







DALL-E 3 / ChatGPT

# Die Datenlandkarte

Was ich über meine Daten wissen sollte

### Datenformate und -Schnittstellen

- Datenbanktabelle, CSV, Mediendaten,...
- Dateiexport, Datenbank, Streaming-Daten,...

### Datenmengen

- Datenfrequenz: Einmal am Tag, tausend Events pro Sekunde,...
- Datenvolumen pro Datensatz oder pro Datentransfer

### Datenqualität

- Befüllungsgrad
- Konsistenz, auch in Bezug zu anderen Datenquellen
- Sonstige Kriterien abhängig von der Fachlichkeit

### Governance und Compliance

- DSGVO, Data Act, AI Act, ...
- Anonymisierung / Pseudonymisierung?
- Eigene Compliance-Regeln

# Nutzungsprofile

Wer will was mit meinen Daten anfangen?

Ein täglicher  
Kennzahlen-Export in  
Excel

Eigene Dashboards  
erzeugen und  
veröffentlichen

Real-Time-Monitoring  
von Vertriebsdaten

Explorative Analyse  
der Rohdaten im  
Notebook

...

Fortgeschrittene KI-  
Anwendungen





## Design nach dem Baukasten-Prinzip

- Für die Architektur und das Tooling setzen wir auf Standard-Bausteine entsprechend den Rahmenbedingungen und Vorgaben beim Kunden (bspw. Cloud oder On-Premise).
- Spezialisierte Anforderungen, die mit Standard-Tooling nicht abgebildet werden können, realisieren wir durch individuelle Entwicklungen.



### Storage

- Data Lake / Cloud Storage
- Data Warehouse (Data Vault, Star Schema,...)
- Lakehouse
- Streaming-Plattform



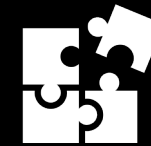
### Aufbereitung

- Low-Code / No-Code
- Scripting
- Compute-Cluster wie Spark



### Analytics & Usability

- BI-Tools
- Notebook-Systeme
- Datenkataloge



### Individuelle Erweiterung

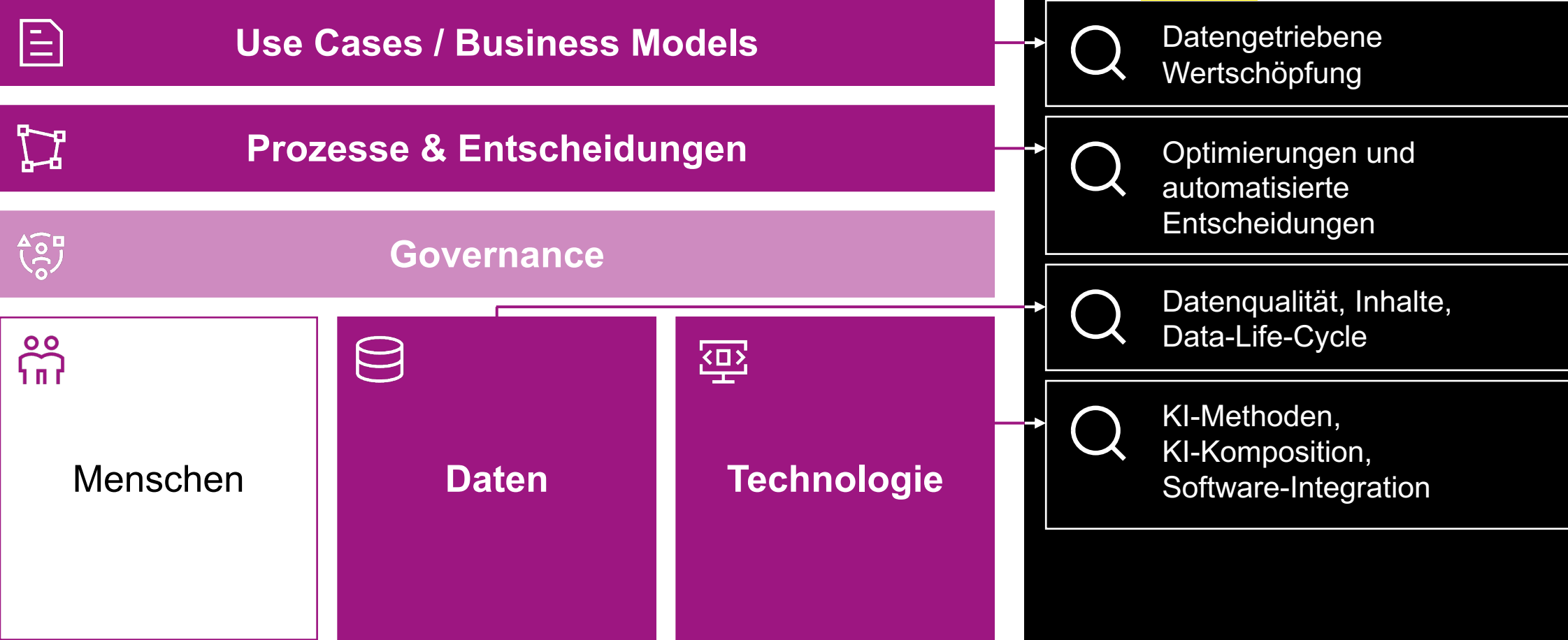
- Individuelle Compliance-Anforderungen
- Spezialisierte Technische Adapter
- ...

Datenstrategie  
**Wertschöpfung**  
**(AI & Analytics)**

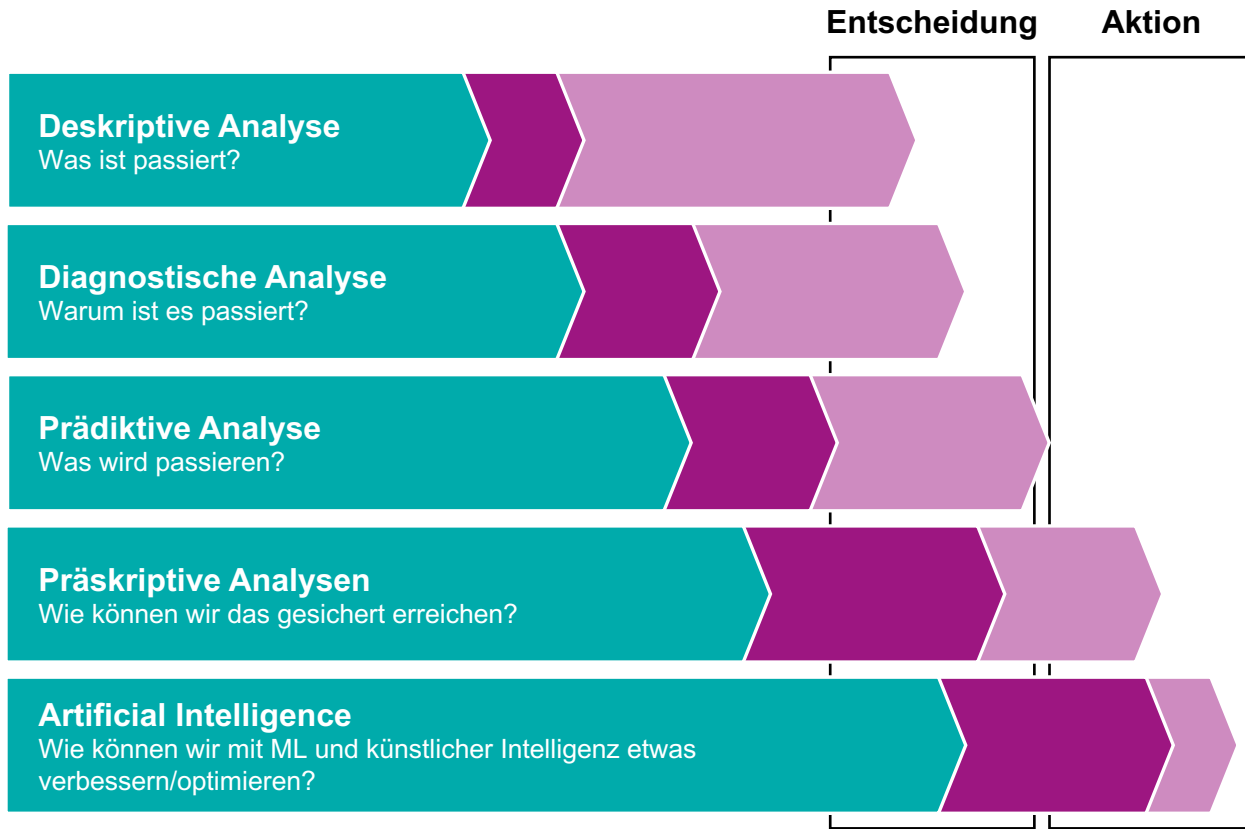




# Datenstrategie



# KI-Reifegrade („AI Maturity“)



■ Technologische Fähigkeiten   
 ■ Menschlicher Input Modellbildung   
 ■ Menschlicher Input Ergebnisinterpretation

## Beispiele

### Beschreibend

- Kundensegmentierung & Präferenzen
- Verkaufsmuster und Trends

### Diagnostisch

- Margen & Kosten
- Vorfälle und Chancen
- Betrug und Sicherheitsverstöße

### Prädiktiv

- Planung und Vorhersage
- Risiko und Stresstests
- Kundenverhalten

### Präskriptiv

- Verkaufs- & Marketingstrategien
- Operative Strategien
- Risiko & Compliance Strategien

### Artificial Intelligence

- Nächste beste Maßnahme in Echtzeit
- Beschwerde Text Mining
- Fraud Detection

# Zur Orientierung: Typische Use Cases



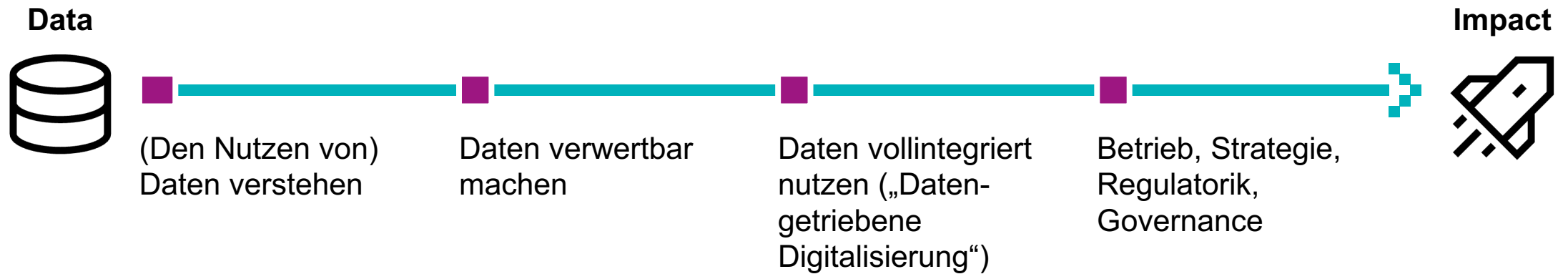
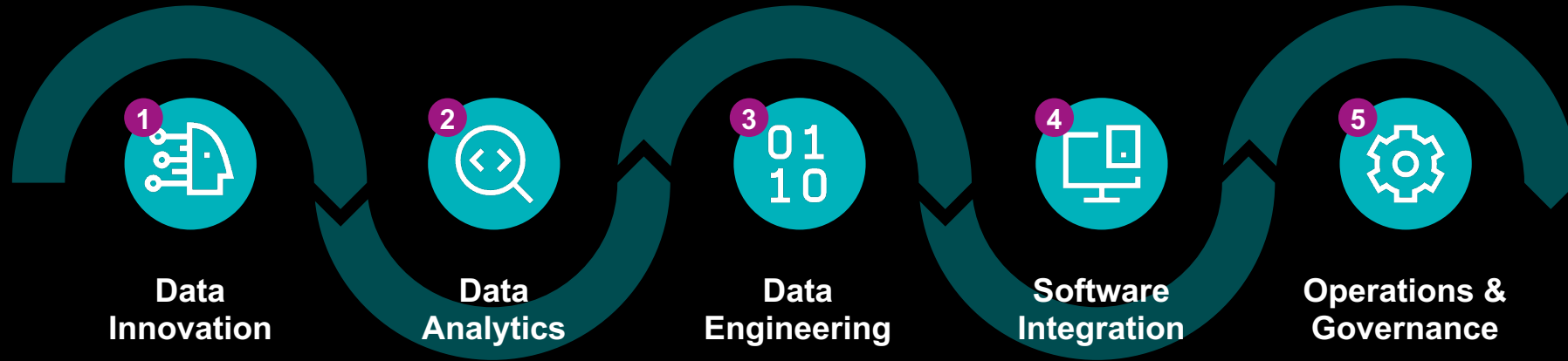
Forecasting

Anomaly Detection

Customer Support

Knowledge  
Management

## Elemente datengetriebener Wertschöpfung





# Datenstrategie Top down



## Data Lake für die Vertriebsplattform

Konzeption, Implementierung und produktive Nutzung eines Data Lakes einer neuen Vertriebsplattform eines Großkunden

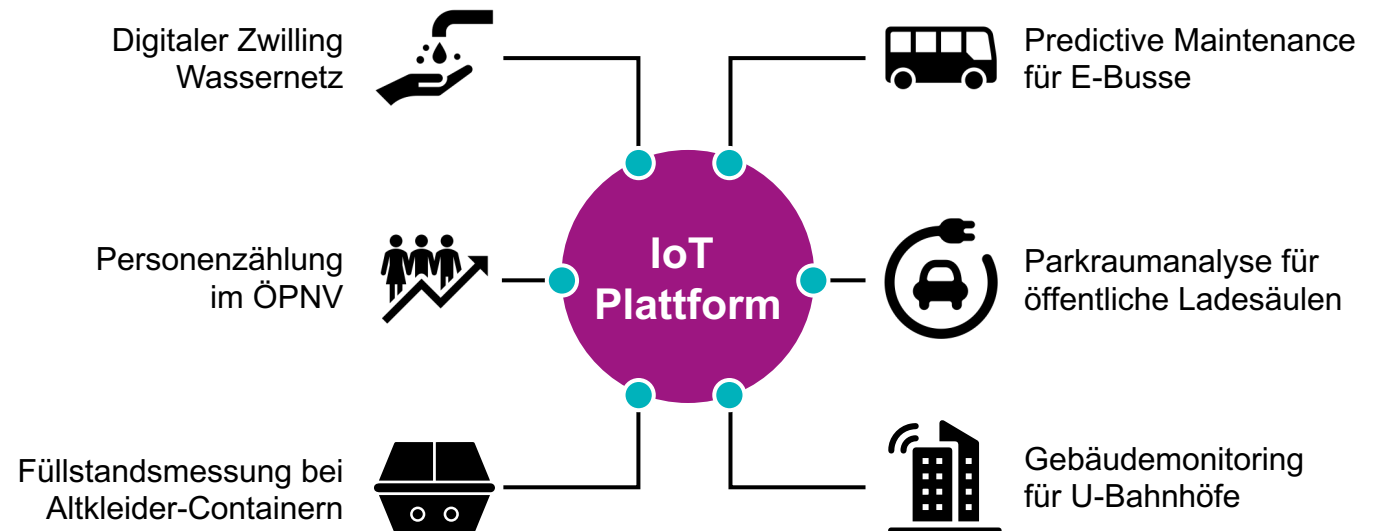
### Strategische Ansätze

- Data First: Bei der Entwicklung neuer Services und Funktionen in der operativen Vertriebsplattform ist die Anbindung an den Data Lake mit qualitativ guten Daten ein Feature erster Klasse
- Privacy By Design: Die Vereinbarkeit von Datenschutz und Datenmehrwert wird früh mitgedacht und durch fachliche Prozesse und spezialisierte technische Lösungen im Kern der Plattform verankert.



# Schaltzentrale für die Stadt der Zukunft

- Als End-to-End Partner für die digitale Transformation begleitet iteratec die IoT-Strategie großer Versorger
- Technische und organisatorische Reife helfen massiv dabei, a) auf regulatorische Anforderungen zu reagieren und b) schnell Use Cases aus den Fachabteilungen in einem Data-First-Mindset umzusetzen





## Namenloser Kunde

Gegenbeispiel: Death by Strategy

### Problem

- Jahrelanges Sammeln von möglichen datengetriebenen Use Cases
- „FOMO-driven design“ (Fear Of Missing Out), keine Priorisierung
- Nie den Schritt vom High-Level-Konzept zu konkreten Zielen oder Umsetzung gewagt

### Effekt

- Viele Use Cases und Rahmenbedingungen sind schon wieder veraltet
- Kein konkreter Mehrwert wird realisiert





iteratec

# Datenstrategie Bottom Up





## Datenanalyse für Früherkennung von Pflanzenschäden

Exploration experimenteller Messdaten in enger Zusammenarbeit mit fachlichen und technischen Stakeholdern.

### Der Mehrwert

- Data Science schafft Datenverständnis:
  - Exploration, Visualisierung und Ergebnispräsentation
  - Quantitative Belastbarkeitsprüfung
- Datenverständnis ermöglicht
  - Ableitung von Handlungsempfehlungen für die weitere Produktentwicklung
  - Grundsatzentscheidungen zum Aufbau organisatorischer Rollen und technologischer Leitplanken







## Eine KI-Anwendung zur Reduktion des CO<sub>2</sub>-Ausstoßes bei der Zementherstellung

Entwicklung eines Web-Anwendungs-MVPs für die **KI-basierte** Vorhersage neuer Zementrezepturen

### Der Mehrwert

- Einfacher Zugang für ein breites Publikum aus Wissenschaft und Industrie erzeugt Traktion und liefert Anforderungen für die Datenstrategie
- Basis für weitere Schritte in Richtung KI-gesteuerte, autonome Forschungslabore





## Wissensmanagement im juristischen Bereich

Sichere Datenräume für GenAI

PoC zur Auswertung komplexer juristischer Dokumente in der Bankenbranche mit GenAI

### Der Mehrwert

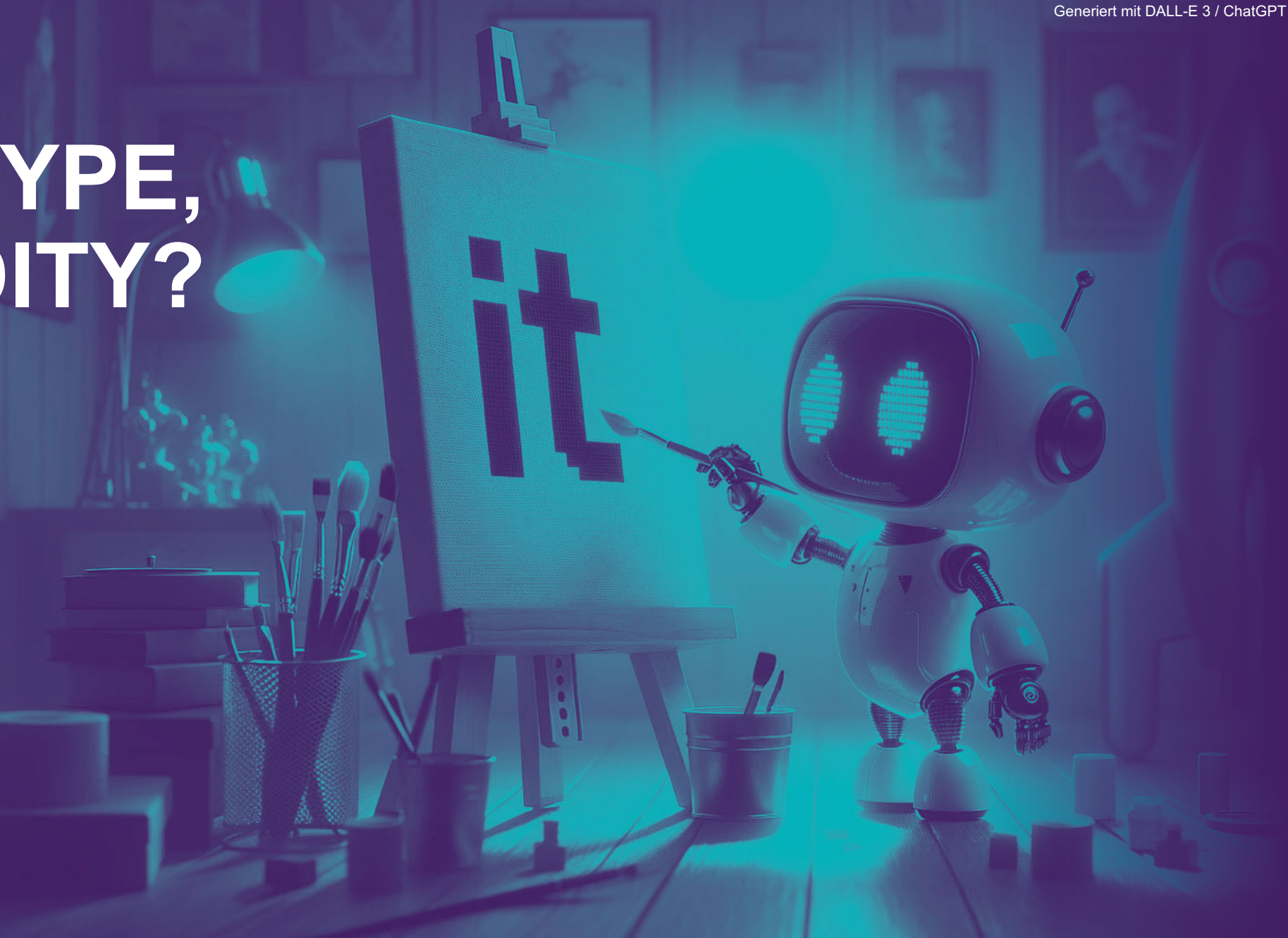
- Leitplanken und Erkenntnisse über technische Machbarkeit
- Input für strategische Unternehmensentscheidungen bzgl. der Verwendung von Daten im Kontext GenAI
- Traktion über Nachweis von Nützlichkeit
- Abbau von Ängsten



# Datenstrategie Disruptionen



# GENAI MAGIE, HYPE, COMMODITY?



# Rasante Entwicklungssprünge innerhalb weniger Monate

Eine verkürzte Chronik der Entwicklung von generativer KI (GenAI)

Entwicklung von Generative Adversarial Networks (GANs), die plausible Bilder mit niedriger Auflösung erzeugen können

2014

2017

Veröffentlichung des Transformer-Papers „Attention Is All You Need“, das die Performance der Textübersetzung verbessert

ChatGPT-4: Großes Sprachmodell mit weit über 1,8 Billionen Parametern

Text 8k

03/2023

11/2022

Text 4k

ChatGPT-3.5: Riesiges Modell mit 175 Milliarden Parametern, die verblüffend menschenähnlichen Text erzeugen können

05/2023

Code 16k

Text 32k

OpenAI Data Analytics mit impliziter automatischer Python-Code-Generierung

Publication

Anthropic/CLAUDE 2: Großes Sprachmodell mit extrem großem Kontext

Text 100k

07/2023

Agenten-

Systeme und KI-Plattform inkl. GPTs, 2M-Entwickler

Text 128k

11/2023

10/2023

Image-2-Text

Text-2-Image

Speech-2-Text

Text-2-Speech

Multimodalität ist

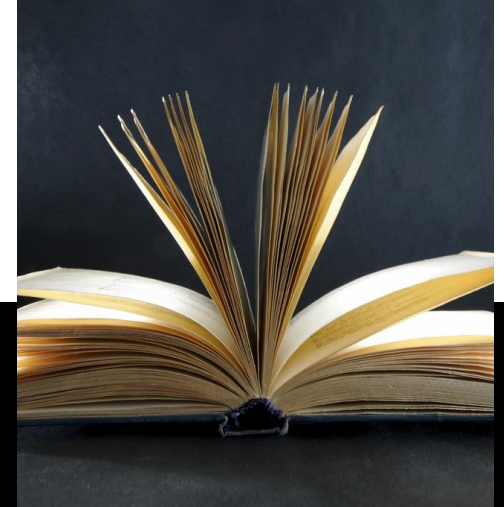
sichtbar:

GPT-4V, DALL-E-3, Mobile-ChatGPT kann sehen, hören und sprechen

12 Monate

Nein, GenAI wird nicht wieder verschwinden ...





Forecasting

Anomaly Detection

Customer Support

Knowledge Management

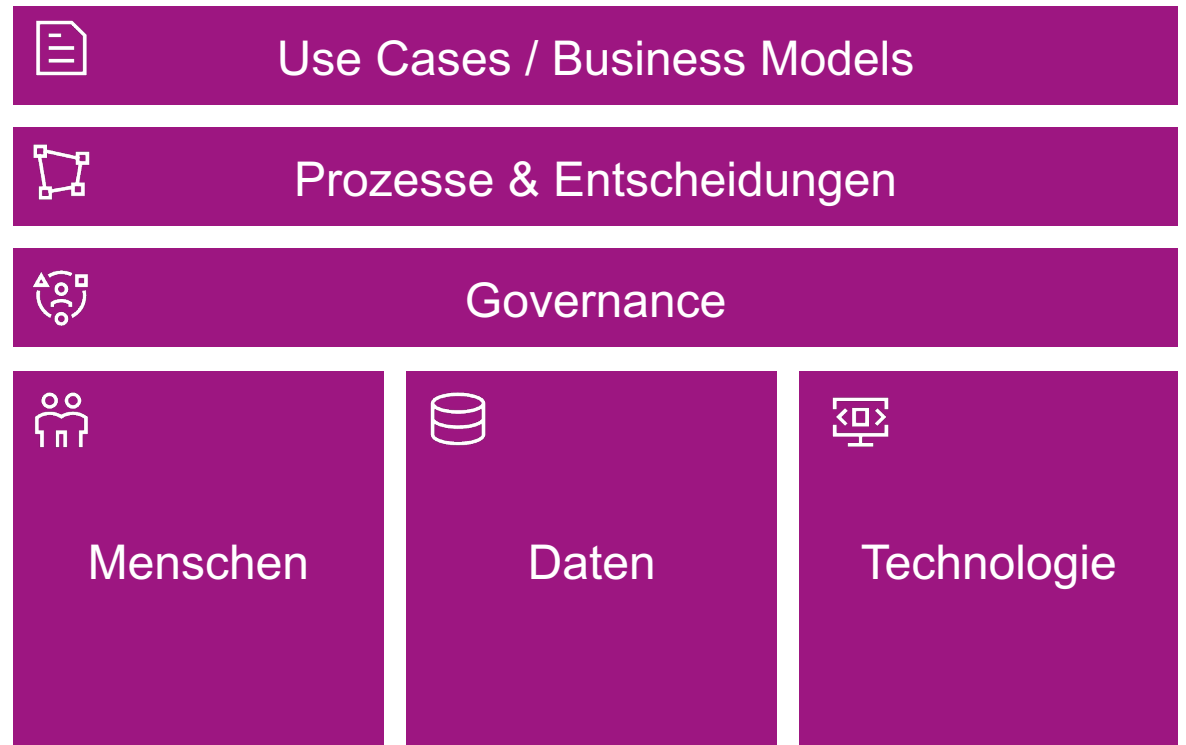


# Der EU Data Act

Herausforderung! Chance?



Eine im Unternehmen verankerte Datenstrategie hilft, die weitreichenden Herausforderungen zu beherrschen



## Aufstieg und Hauptbereiche der Veränderung



### Wirtschaftliche Veränderungen

- Automatisierung von Routineaufgaben → *Effizienzsteigerung und Kostenreduktion*
- Neue Geschäftsmodelle und Dienstleistungen → *KI als Treiber für Innovation*
- Marktdynamik und Wettbewerbslandschaft → *Schnellere Anpassungsfähigkeit gefragt*



### Unternehmensspezifische Disruptionen:

- Entscheidungsfindung und Analytik → *Datenbasierte und präzisere Entscheidungen*
- Kundeninteraktion und Personalisierung → *Verbesserte Kundenerfahrung*
- Talentmanagement und Arbeitsplatzgestaltung → *Neue Anforderungen und Kompetenzen*



### Spezialthema Generative AI als Katalysator:

- Innovative Content-Erstellung → *Texte, Bilder, Videos automatisiert generieren*
- Produktentwicklung und Design → *Beschleunigte Kreativprozesse*
- Ethik und Regulierung → *Neue Herausforderungen und Rahmenbedingungen*
- Mensch-Maschine-Interface in der Digitalisierung → *Bessere Alltags-Integration*

**Adapt or die out**