

## Topic 2: Strategische Perspektive

Block 1: Grundlagen Data Science

## Häufige Irrtümer

- *“Garbage in, garbage out”* – schlechte Daten liefern schlechte Ergebnisse
- *“Mehr Daten lösen alle Probleme”* – falsch!
- KI ersetzt menschliche Expertise vollständig
- Automatisierung ist immer die beste Lösung

## Was wirklich funktioniert

- Klare, präzise Fragestellung definieren
- Gute Datenqualität sicherstellen
- Passende Methode für das Problem wählen
- Iteration und kontinuierliche Validierung
- Domain-Wissen einbinden

---

Realistische Erwartungen und methodisches Vorgehen sind entscheidend für den Erfolg

## Wertschöpfung durch Daten

- **Wettbewerbsvorteil** durch datenbasierte Entscheidungen
- **Automatisierung** von Prozessen und Workflows
- **Personalisierung** von Produkten und Services
- **Neue Geschäftsmodelle** ermöglichen

## Data Governance

- DSGVO-Compliance und Datenschutz
- Klare Verantwortlichkeiten
- Dokumentation und Nachvollziehbarkeit

## Herausforderungen

- **Datensilos** – fragmentierte Systeme verhindern ganzheitliche Sicht
- **Datenqualität** – inkonsistente, unvollständige oder veraltete Daten
- **Talent-Mangel** – Data Scientists und Engineers sind schwer zu finden
- **Technologie-Stack** – richtige Tools und Infrastruktur wählen
- **Kulturwandel** – von Intuition zu datengetriebenen Entscheidungen

---

Daten sind wertvoll, aber nur mit richtiger Governance und Infrastruktur nutzbar

## Zeitallokation in realen Data Science Projekten

**40–60%**

### **Datenbereinigung**

- Fehlende Werte behandeln
- Duplikate entfernen
- Datenformate vereinheitlichen
- Outliers identifizieren

**20–30%**

### **Feature Engineering**

- Relevante Variablen erstellen
- Transformationen anwenden
- Domain-Wissen einbringen

**10–20%**

### **Modellierung & Tuning**

- Modell auswählen und trainieren
- Hyperparameter optimieren

---

Die meiste Zeit geht für Datenvorbereitung drauf – nicht fürs Modellieren!

## Erwartungsmanagement und Kommunikation

### Realistische Timelines

- Proof-of-Concept: 4–8 Wochen
- MVP (Minimum Viable Product): 3–6 Monate
- Production-Ready System: 6–12 Monate
- Iterationen einplanen
- Unerwartete Probleme (Datenqualität!) puffern

### Ihre Kernaufgaben

- **Übersetzer** zwischen Fachbereich und Data Scientists
- Geschäftsziele in technische Anforderungen übersetzen
- Prioritäten setzen und kommunizieren
- Erwartungen realistisch halten
- Ressourcen und Budget sichern
- Erfolge sichtbar machen

---

Gutes Management bedeutet: klare Kommunikation, realistische Erwartungen, Brücke zwischen Business und Technik