



Fakultät für  
**Mathematik und  
Informatik**

# Aufbau und Struktur

Einführung in Data Science

## Inhaltsverzeichnis

- Allgemeine Ziele
- Überblick über die einzelnen Lektionen
- Übungsbetrieb
- Terminübersicht
- Modulabschlussprüfung

## Über diese Lehrveranstaltung

- Diese Lehrveranstaltung bietet eine konzeptionelle Einführung in das Fach Data Science
- Vorbereitung für weitere Module im Curriculum:
  - Data Engineering für Data Science (Prof. Störl)
  - Mathematische Grundlagen von Data Science (Prof. Riedel)
  - Einführung in maschinelles Lernen (Prof. Thimm)
- Bei dieser Lehrveranstaltung handelt es sich nicht um einen reinen Programmierkurs (obwohl wir zentrale Programmiersprachen kennen lernen werden)

## Allgemeine Lernziele

- Nach Abschluss dieser Lehrveranstaltung können Sie ...
  - erläutern, was Data Science bedeutet, und exemplarische Anwendungsgebiete beschreiben.
  - die wichtigsten Fragestellungen des Gebiets Data Science benennen.
  - einen Data Science Life Cycle beschreiben, die damit verbundenen Schritte und Methoden anwenden sowie eigene Fragestellungen entwickeln und beantworten.
  - unterschiedliche Analysearten für Ihre Daten unterscheiden und anwenden.
  - verschiedene Aspekte der Datensicherheit und Datenintegrität kritisch reflektieren und beurteilen.
  - ethische Fragestellungen theoriebasiert auf dem Data Science Life Cycle anwenden und diskutieren.
  - die grundlegenden Werte- und Normensysteme ethischer und rechtlicher Grundlagen erläutern.

## Die einzelnen Lektionen im Überblick

1. Was ist Data Science?
2. Datenethik
3. Data Science Life Cycle und Vorgehensmodelle
4. Grundlegende Konzepte und Methoden
5. Methoden, Tools und Anwendungen
6. Datenrecht
7. Visualisierung und Kommunikation



## Lektion 1: Was ist Data Science?

- *Was ist Data Science und was lerne ich in diesem Modul?*
- Data Science ist mehr als die Wissenschaft über Daten
- Von zentraler Bedeutung für die Entwicklung von intelligenten Systemen (vgl. Künstliche Intelligenz)

### Technologien und Infrastruktur

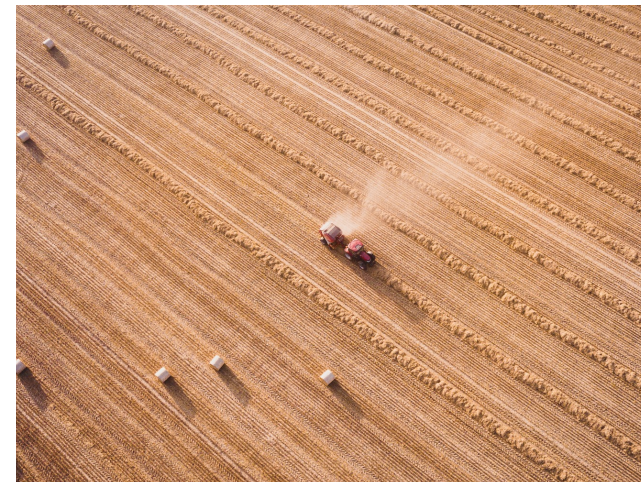
(Database and Information Systems, Tools, Libraries,  
Big Data Frameworks, Cloud Computing)

### Methoden und Algorithmen

(Data Mining, Machine Learning, Information Retrieval,  
Visual Analytics, Data Engineering)

### Grundlagen und Theorie

(Mathematik, Statistik, Optimierung, Algorithmen)



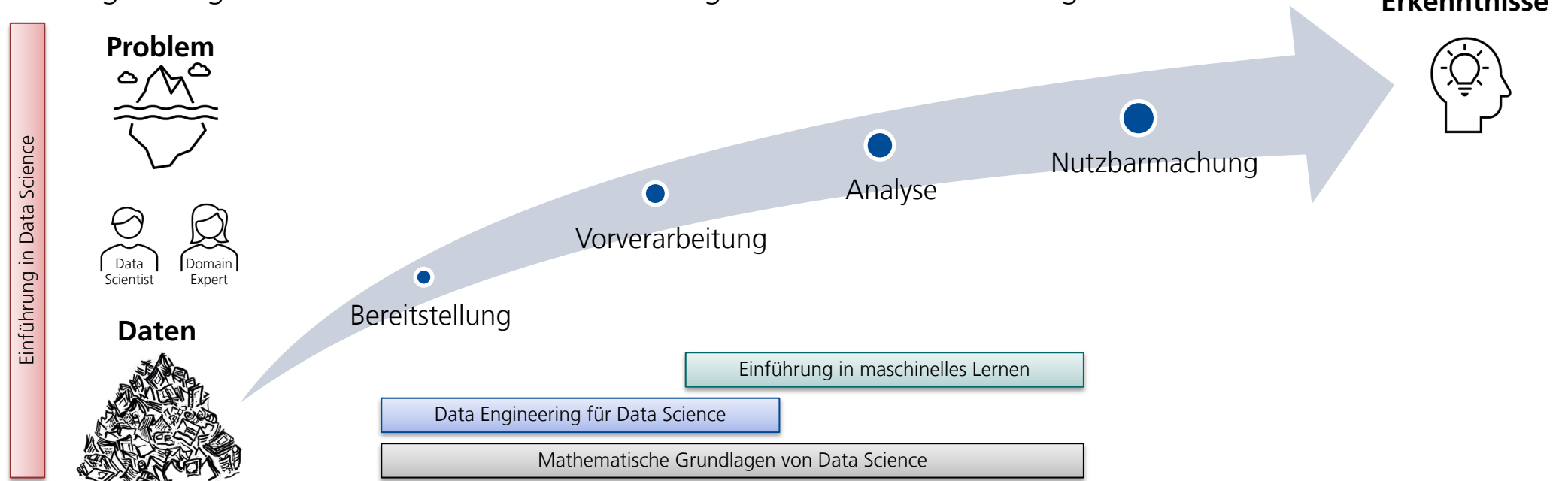
## Lektion 2: Datenethik

- Warum sollte sich ein Data Scientist auch mit ethischen Fragen auseinandersetzen?
  - „What does it take to build a good data product or service? Not just a product or service that’s useful, or one that’s commercially viable, **but one that uses data ethically and responsibly.**“ (Loukides, 2018)
- Sensibilisierung für einen Systementwurf, der „Ethics by Design“ einschließt



## Lektion 3: Data Science Life Cycle und Vorgehensmodelle

- *Wie können wir Data Science strukturiert anwenden?*
- Strukturiertes Vorgehen fördert Akzeptanz, Nachvollziehbarkeit und Übertragbarkeit
- Ein langer Weg von Daten über Nutzbarmachung bis hin zum Erkenntnisgewinn:





## Lektion 4: Grundlegende Konzepte und Methoden

- *Welche Konzepte sollten wir kennen und welche Methoden verbergen sich dahinter?*
- Datenanalytisches Denken auf einer konzeptionellen Ebene hilft uns komplexe Probleme zu verstehen, Lösungen zu entwickeln und zu bewerten
- Konzepte können mit einer Vielzahl von Methoden umgesetzt werden

### Strukturiertes Vorgehen

*KDD, CRISP-DM, DASC-PM*

### Überwachtes Lernen

*Lineare Regression, Nächste  
Nachbarn, Entscheidungs-  
bäume, SVM, Bayes*

## Lektion 5: Methoden, Tools und Anwendungen

- *Müssen wir alle Methoden selber implementieren?*
- Programmiersprachen wie Python oder R bieten vielfältige Möglichkeiten
- Analyseplattformen und Cloudbasierte Lösungen ermöglichen Skalierung



## Lektion 6: Datenrecht

- Die Diskrepanz zwischen dem erlaubten und dem wünschenswerten Umgang mit Daten – kann diese aufgelöst werden?
  - Gesetze sind keine ethischen Leitlinien
  - Aber es gibt Leitlinien zum ethischen Umgang mit Daten
- Forschungsdatenmanagement – welche Auswirkungen hat dies auf mich, meine Daten und meine Forschung?



## Lektion 7: Visualisierung und Kommunikation

- *Wie präsentieren und kommunizieren wir Daten und Erkenntnisse?*
- *Es gibt viele Visualisierungstechniken, doch welche grundlegenden Regeln sollten wir beachten?*
- Storytelling mit Daten als Prozess der effektiven Kommunikation



angepasste Version von <https://www.storytellingwithdata.com>

## Übungsbetrieb

- Insgesamt gibt es drei fakultativ zu bearbeitende Einsendeaufgaben
- Wir empfehlen Ihnen die Teilnahme, ...
  - um sich intensiver mit den Themen auseinanderzusetzen
  - um Ihren individuellen Lernstand zu reflektieren
  - um ein individuelles Feedback von uns zu erhalten
  - um sich auf die abschließende Modulabschlussprüfung vorzubereiten

## Terminübersicht

Lektion	Bearbeitungsbeginn	Einsendeaufgabe	Einsendeschluss
1	02.10.2023		
2	16.10.2023	1	19.11.2023
3	30.10.2023		
4	13.11.2023	2	17.12.2023
5	27.11.2023		
6	11.12.2023	3	21.01.2024
7	08.01.2024		



## Modulabschlussprüfung

- Für diese Lehrveranstaltung wird eine **mündliche Prüfung** angeboten
- Informationen zur Anmeldung und den Terminen werden wir zu einem späteren Zeitpunkt bekanntgeben
- Weiterführende Informationen:
  - Informationen zur Prüfungsordnung finden Sie unter <https://www.fernuni-hagen.de/mi/studium/pdf/pomscds.pdf>
  - Informationen zur Prüfungsanmeldung finden Sie in der Rubrik Prüfungsinformationen unter <https://www.fernuni-hagen.de/mi/download/studium.shtml>
  - Die Prüfungsanmeldung selbst finden Sie unter <https://pos.fernuni-hagen.de/>



Fakultät für  
**Mathematik und  
Informatik**

# Vielen Dank für Ihre Aufmerksamkeit!

## Einführung in Data Science