



Fakultät für
**Mathematik und
Informatik**

Willkommen beim Projektpraktikum „Data Mining Cup“!

Auftaktveranstaltung am 08.04.2024

Fakultät für **Mathematik und Informatik**

Lehrgebiet Data Science / Prof. Dr. Christian Beecks
Sabine Folz-Weinstein

Agenda

1. Einführung
2. Organisatorisches
3. Timeline
4. DASC-PM
5. Aufgabenstellung
6. Diskussion

Vorstellungsrunde

Kursbetreuung FernUni:

- Sabine Folz-Weinstein
- Max Pernklau
- Prof. Dr. Christian Beecks

Teilnehmer:

- Johannes Winkler
- Dominik Lewin
- Mario Teßmann

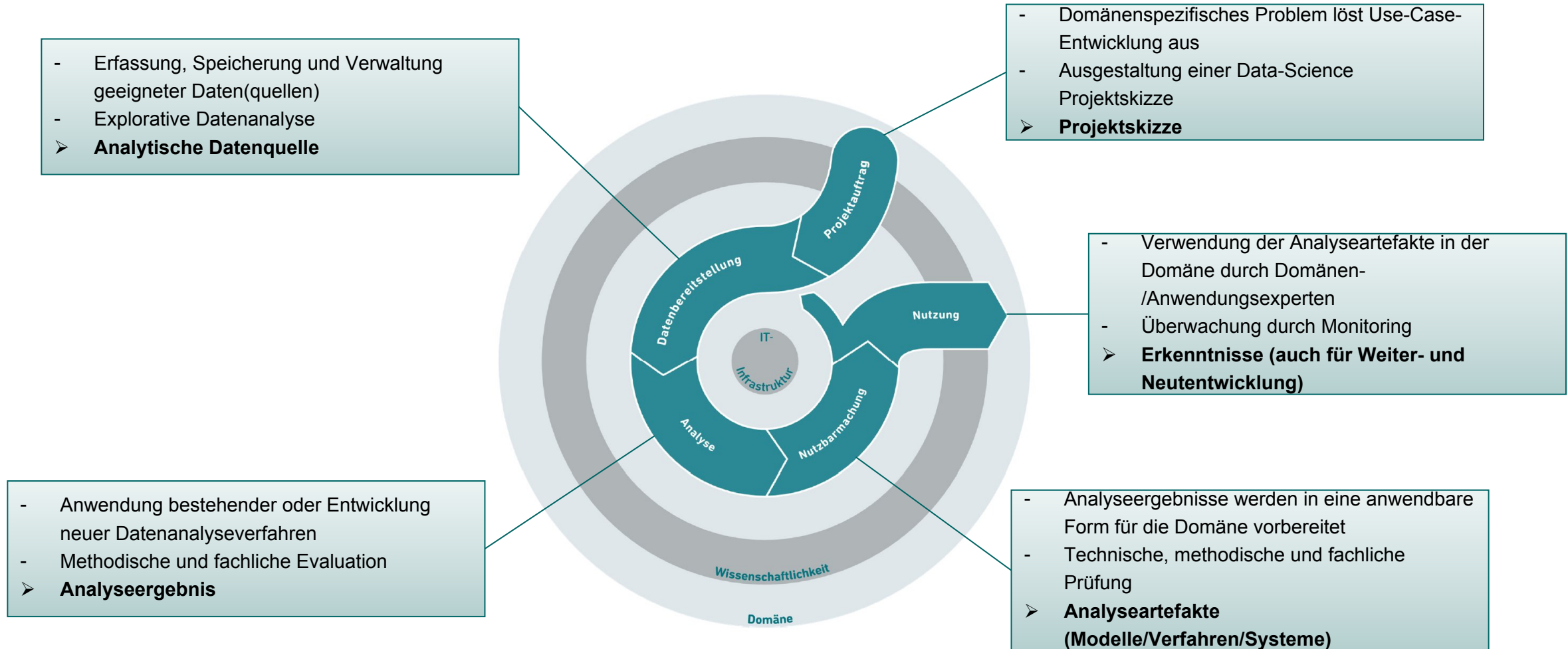


Organisatorisches

- Data Mining Cup wird 2024 nicht ausgerichtet
- Aufgabe inspiriert durch frühere DMC Aufgabe aus dem Bereich Retail Commerce
- Teamarbeit
- Projekt nach DASC-PM
- Bewertung: 3 Säulen (Technisch-fachlich, Dokumentation, Präsentation)

Timeline

08.04.2024	Kick-off
1-2 wöchentlich (ca. 30min)	Regelmäßiger Statusupdate / Zwischenpräsentationen Präsentation von Artefakten je nach Phase Selbständige Projektphasen- und Meilensteinplanung gemäß DASC-PM !
04.07.2024	Abschlusspräsentation



M. Schulz et al.: DASC-PM v1.1 – Ein Vorgehensmodell für Data-Science-Projekte. NORDAKADEMIE gAG Hochschule der Wirtschaft. 2022.

DASC-PM

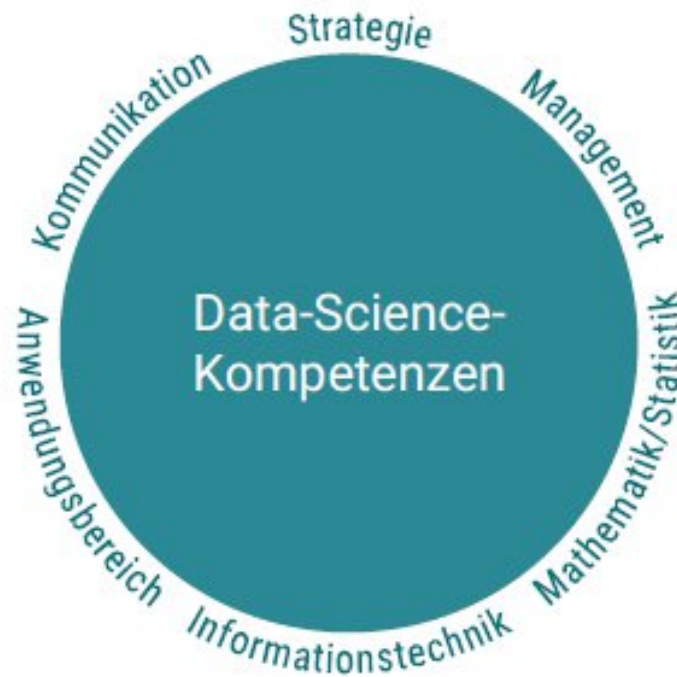


Abbildung 3: Notwendige Kompetenzen in einem Data-Science-Projekt

M.Schulz et al.: DASC-PM v1.1 - Ein Vorgehensmodell für Data-Science-Projekte. NORDAKADEMIE gAG Hochschule der Wirtschaft. 2022.

DASC-PM

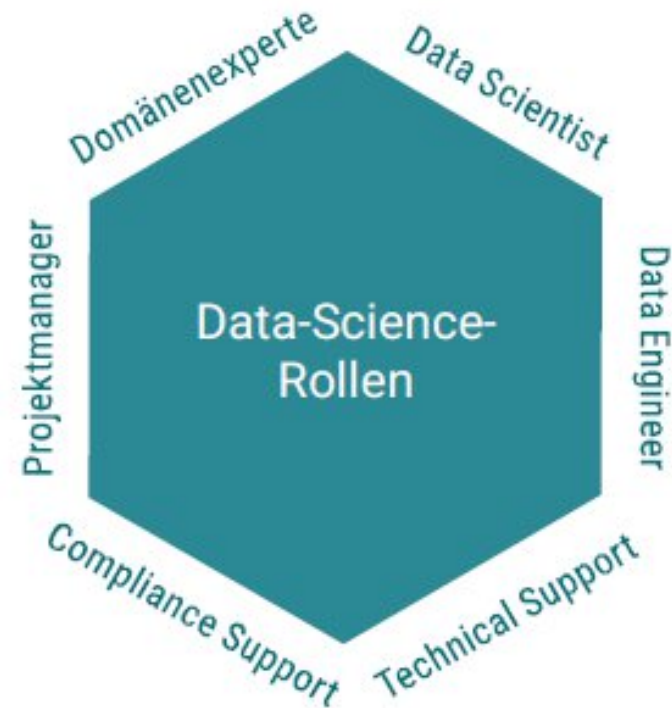


Abbildung 4: Rollen in einem Data-Science-Projekt

M.Schulz et al.: DASC-PM v1.1 - Ein Vorgehensmodell für Data-Science-Projekte. NORDAKADEMIE gAG Hochschule der Wirtschaft. 2022.

Aufgabenstellung

Betrug an Self Checkout-Kassen

Studierende: Datenanalyse-StartUp

Kursbetreuung: Einzelhandelsgeschäft



Bildquelle: https://www.glory-global.com/de-de/blogs/de_de/2019/self-checkout-kassen/

Aufgabenstellung

- Selbst-Scan-Kassen mit mobilen Handscannern
- Risiko: nicht alle Waren werden von den Kunden gescannt
- Gründe: Absicht, Versehen, technische Probleme
- Empirische Untersuchungen: ca. 5% der Scans sind inkorrekt („fraudulent“)
- zielgerichtete Nachkontrollen bei verdächtigen Scanvorgängen
- Checks so gering wie möglich (Kosten, Kundenverärgerung) <-> mögl. viele Falschscans aufdecken
- Trainingsdaten für Modell verfügbar (Moodle)

Diskussion

Ihre Fragen ?

