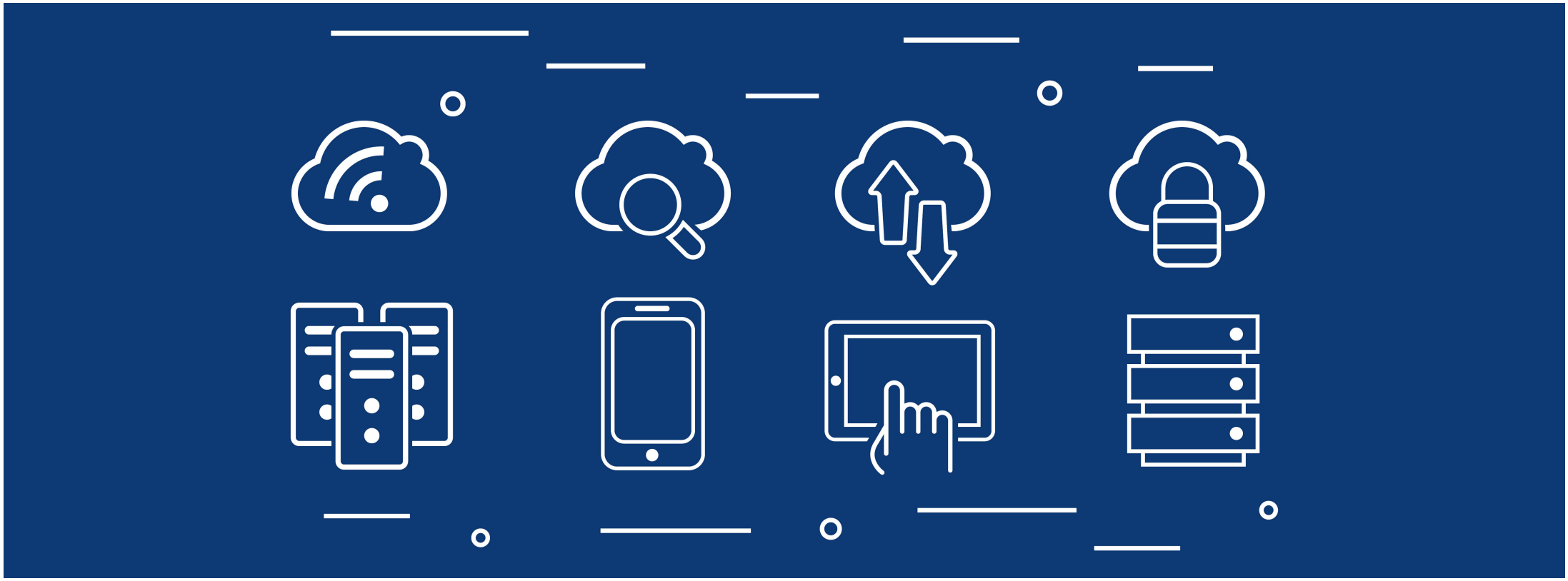


Engineering Data-Driven Applications: Mit Struktur zur KI-Anwendung

Marc Hesenius

SE LEHRSTUHL FÜR SOFTWARE ENGINEERING,
INSB. MOBILE ANWENDUNGEN
PROF. DR. V. GRUHN

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

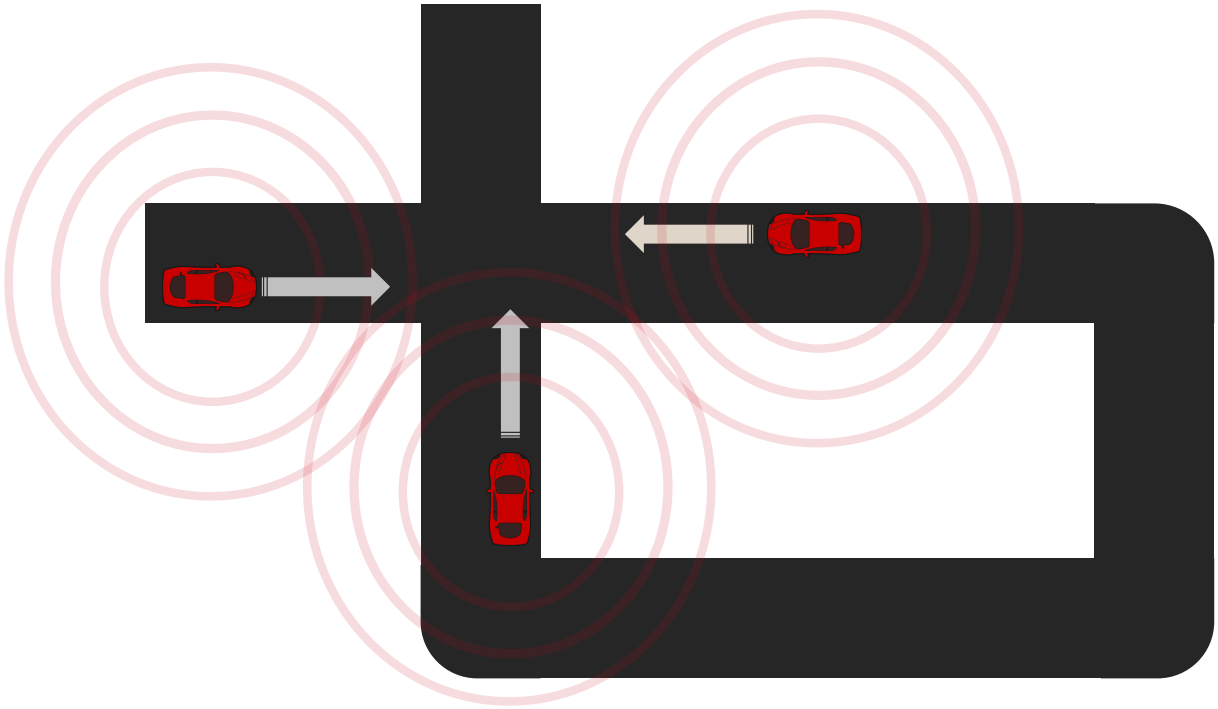
Künstliche Intelligenz ...

- ... ist nicht immer smart!

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Künstliche Intelligenz …

- … ist nicht immer smart!
- Beispiel: Kreuzungsteuerung
  - Autonome Fahrzeuge
  - Ziel: Keine Kollisionen

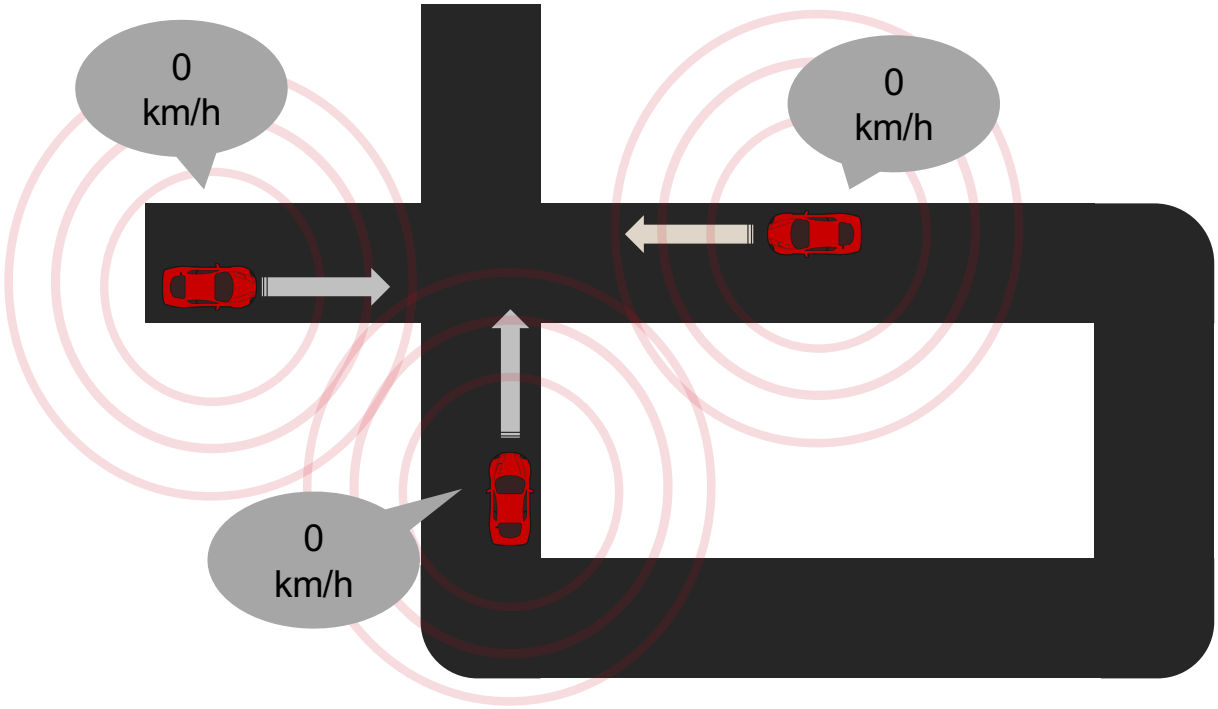


Bildquelle Fahrzeug: https://pixabay.com/en/racing-car-ferrari-red-top-view-296772/, CC0 Creative Commons

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Künstliche Intelligenz …

- … ist nicht immer smart!
- Beispiel: Kreuzungsteuerung
  - Autonome Fahrzeuge
  - Ziel: Keine Kollisionen
  - Ergebnis: Fahrzeuge blieben irgendwann alle stehen

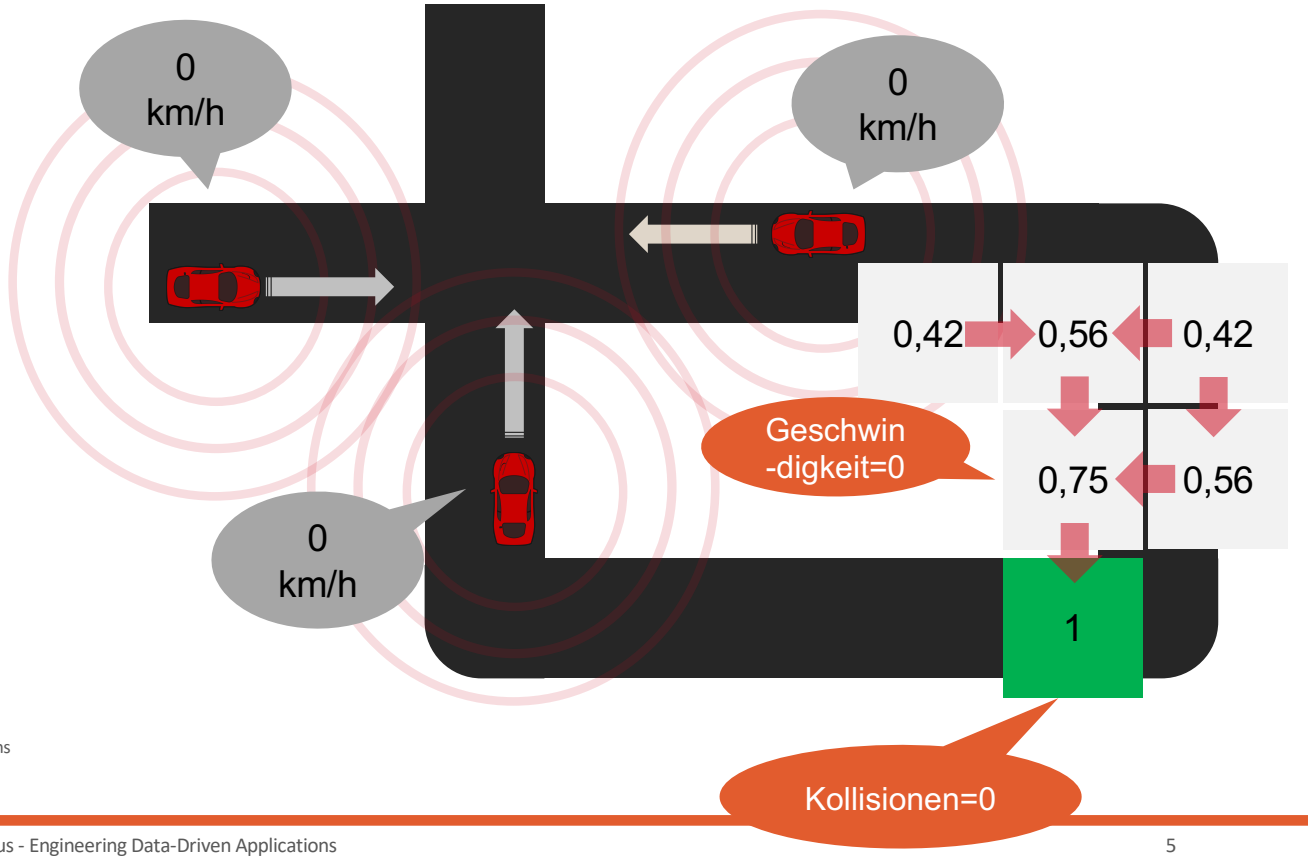


Bildquelle Fahrzeug: https://pixabay.com/en/racing-car-ferrari-red-top-view-296772/, CC0 Creative Commons

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Künstliche Intelligenz ...

- ... ist nicht immer smart!
- Beispiel: Kreuzungsteuerung
  - Autonome Fahrzeuge
  - Ziel: Keine Kollisionen
  - Ergebnis: Fahrzeuge blieben irgendwann alle stehen
  - Hintergrund: Falsche Belohnungs-funktion

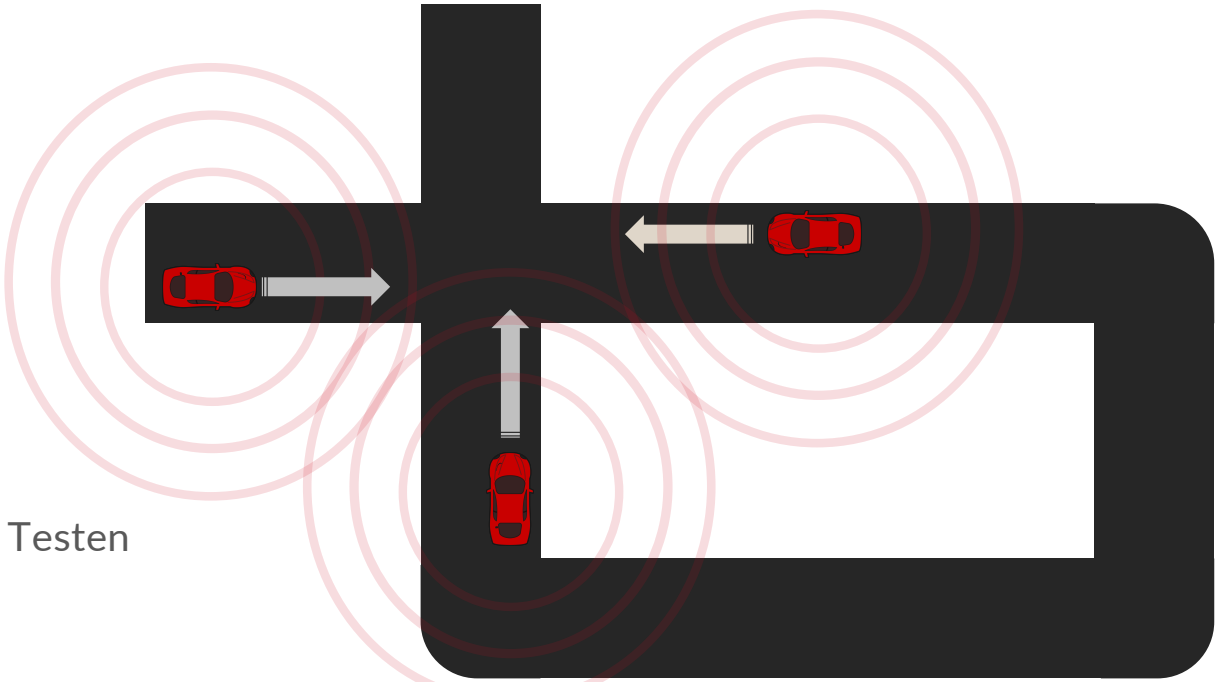


Bildquelle Fahrzeug: https://pixabay.com/en/racing-car-ferrari-red-top-view-296772/, CC0 Creative Commons

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Künstliche Intelligenz …

- … ist nicht immer smart!
- Die Entwicklung von KI-Lösungen erfordert
  - die explizite Betrachtung fachlicher Aspekte
  - zugeschnittene Engineering-Methoden
  - iteratives Vorgehen und konstantes Testen

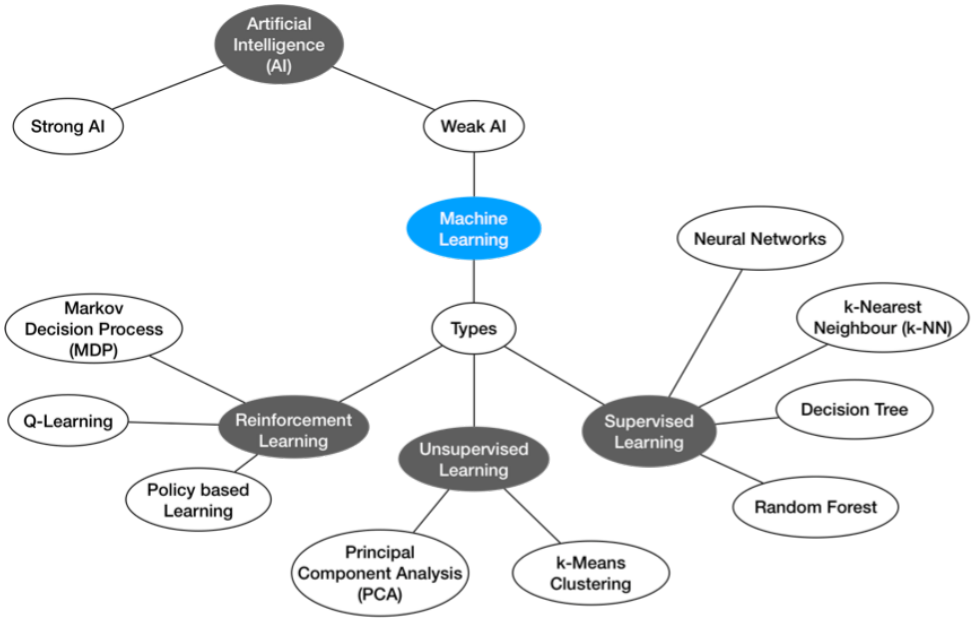


Bildquelle Fahrzeug: https://pixabay.com/en/racing-car-ferrari-red-top-view-296772/, CC0 Creative Commons

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

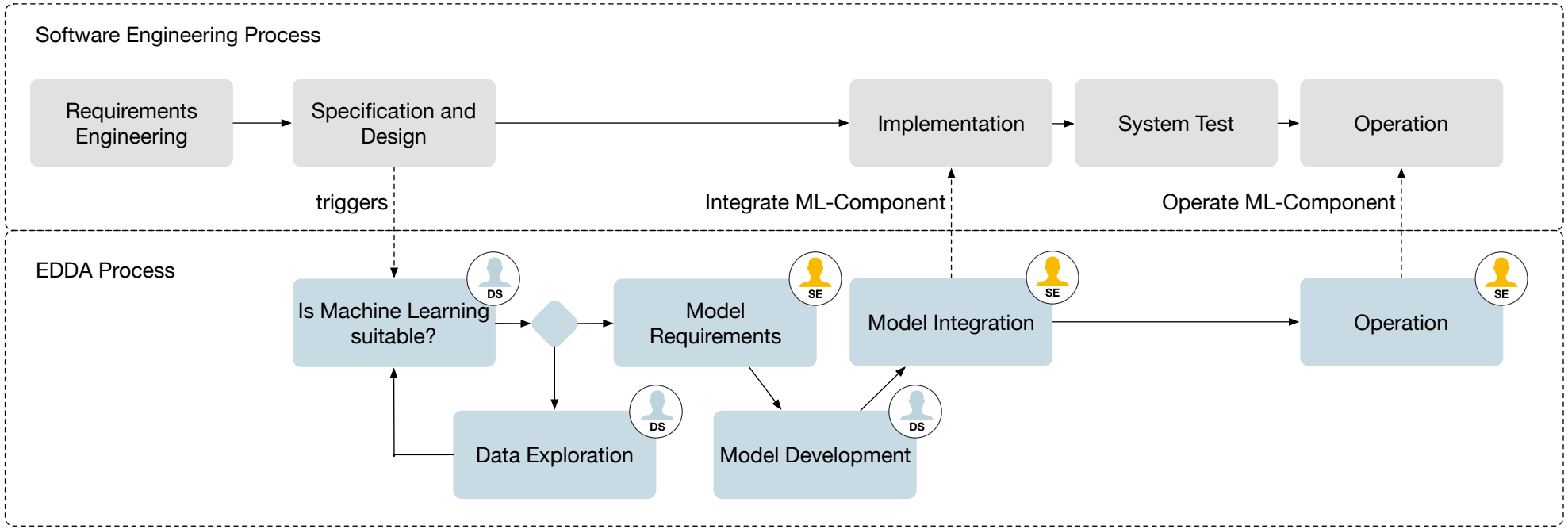
Künstliche Intelligenz und Software Engineering

- Die Entwicklung von KI/ML-Anwendungen erfordert neue Werkzeuge und Methoden
  - Der Einsatz von KI/ML vereinfacht Probleme nicht prinzipiell
  - Es findet hauptsächlich eine Verschiebung der geforderten Kenntnisse und Arbeiten statt.
- Weitere (neue) Spezialisten müssen in den Entwicklungsprozess integriert werden
- Typische Metriken des Software Engineering werden für KI/ML-Anwendungen relevant
  - Wiederverwendbarkeit, Erweiterbarkeit, Produktivität, ...



UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Mit Struktur zur KI-Anwendung



UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Rollen und Beteiligte

DE

Domain Expert

- Kennt sich mit den Geschäftsprozessen des Unternehmens und der Branche aus

DS

Data Scientist

- Hat KI-Kenntnisse
- Beherrscht Techniken der Datenanalyse
- Hat grundlegende Kenntnisse der Software-Entwicklung

DDE

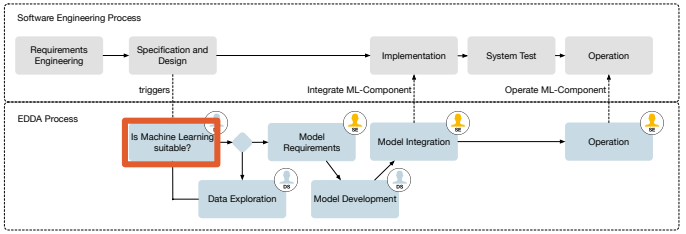
Data Domain Expert

- Kennt sich mit den Datenstrukturen des Unternehmens und der Branche aus

SE

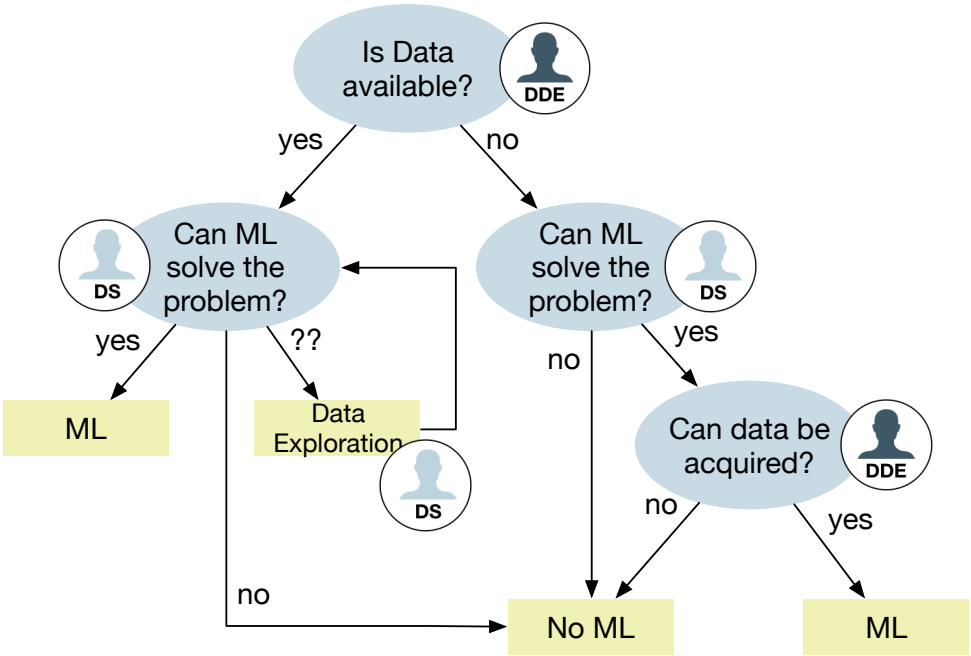
Software Engineer

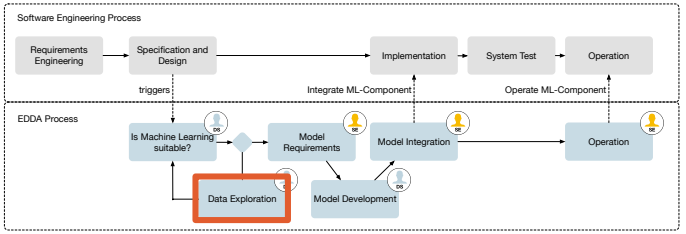
- Definiert und entwickelt technische Komponenten
- Hat grundlegende Kenntnisse der Datenanalyse
- Kennt technische Komplexität der Daten



Ist ML passend?

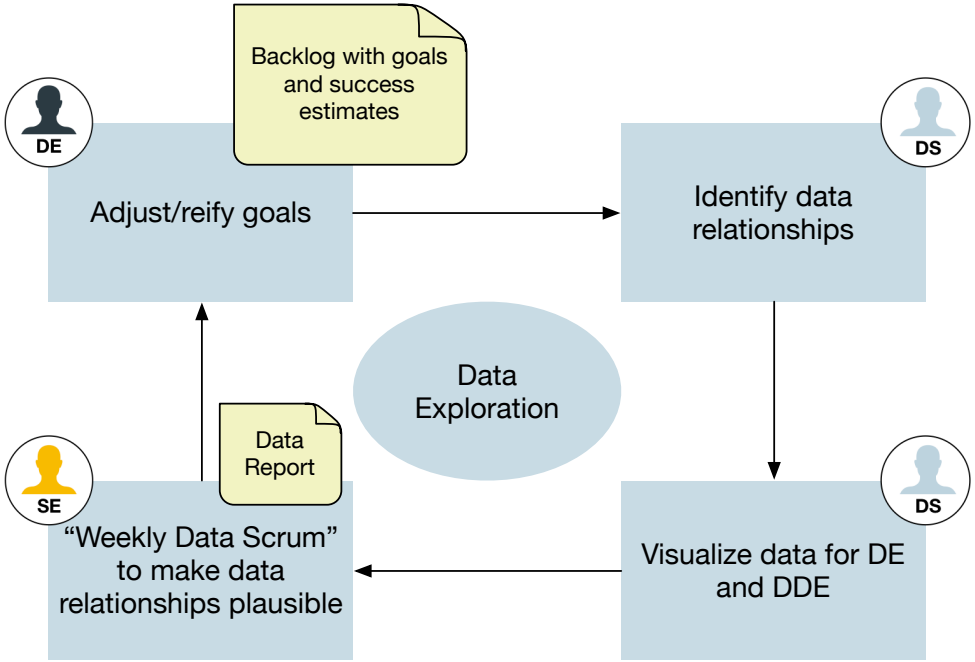
- Prüfung, ob ein Problem mit KI/ML überhaupt gelöst werden kann
- Verantwortlich: Data Scientist
- Vorbereitung für spätere Phasen
- Drei potentielle Ergebnisse:
  - Ja
  - Nein
  - Unbekannt → Data Exploration





Data Exploration

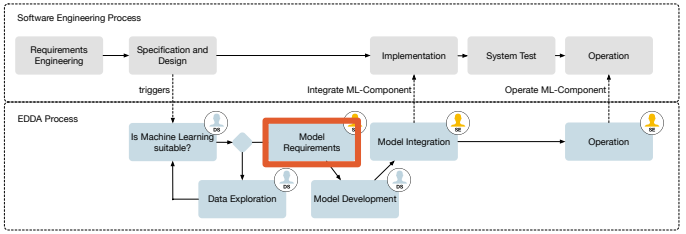
- Verfügbare Daten nach (unbekannten) Verbindungen und potentiellen Lösungen durchsuchen
- Verantwortlich: Data Scientist
- Visualisierung von Daten und Zusammenhängen für andere Team-Mitglieder
- Weekly Data Scrum, um Erkenntnisse zu diskutieren und Daten zu plausibilisieren
- Data Report, um Wissen über Daten und Zusammenhänge zu dokumentieren



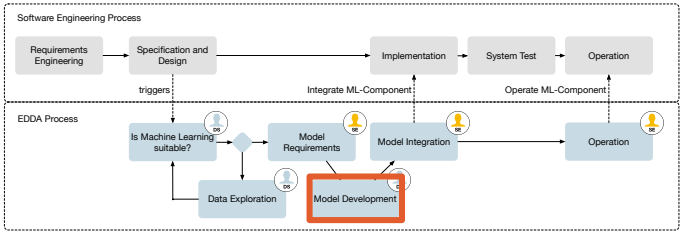
UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Anforderungen an das Modell

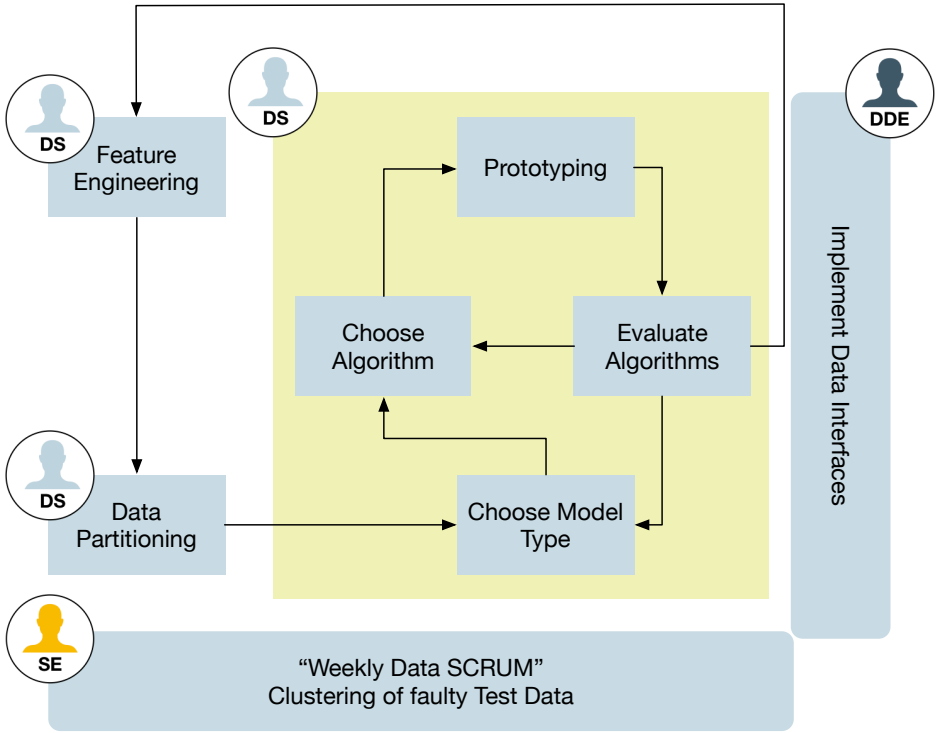


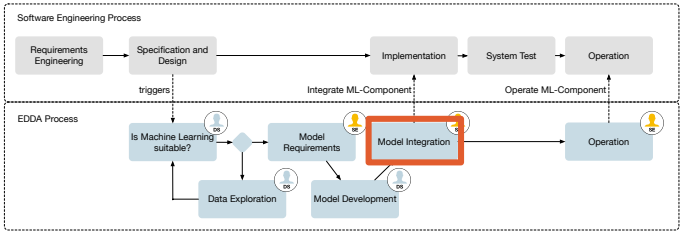
- Anforderungen an die KI/ML-Komponente erfassen und Architektur für eine potentielle Lösung skizzieren
- Verantwortlich: Software Engineer
  - Komponente muss in die Anwendung passen!
- Vorgehen im Detail und wichtige Fragen abhängig vom Typ der KI/ML-Komponente
  - Regression: Können offensichtlich falsche Vorhersagen zuverlässig erkannt werden?
  - Anomaliedetektion: Bedeutung von false positives/false negatives
- Zwei zentrale Artefakte
- Data Story
  - Verknüpft Daten und Use Case
  - Enthält weitere Informationen über Daten, z. B. Datenqualität oder Anforderungen an Genauigkeit
- Architecture Sketch
  - Abhängigkeiten von anderen Komponenten und Datenquellen
  - Eventuell nötige Fallback-Mechanismen
  - Integration neuer Daten



Entwicklung des Modells

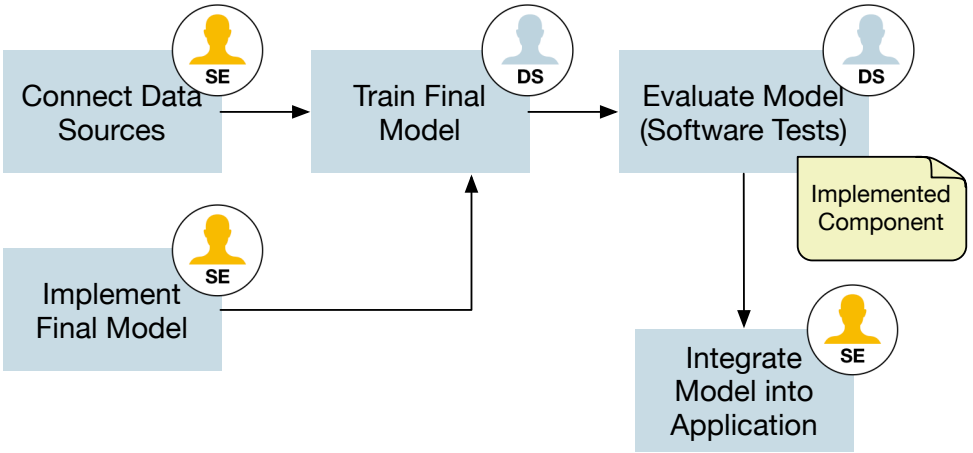
- Prototyping eines ersten Modells
  - Auswahl von Algorithmen
  - Definition der notwendigen Daten-Sets
  - Feature Engineering
  - …
- Verantwortlich: Data Scientist
- Entwicklung notwendiger Tests, um Qualität des Modells sicherzustellen
- Fokus liegt auf mehr auf Definition des Modells als auf Entwicklung der finalen Komponente



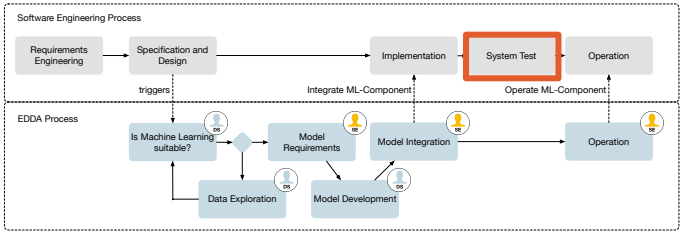


Integration des Modells

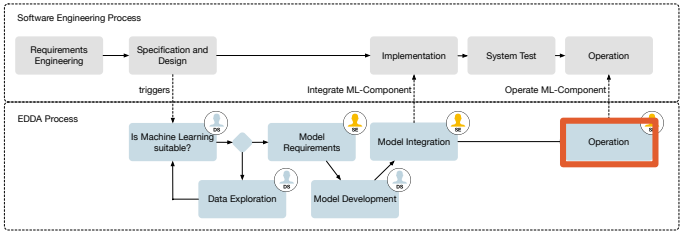
- Implementierung der finalen Komponente und Integration in die Anwendung
- Verantwortlich: Software Engineer
- Wenn nötig: Neuentwicklung prototypischer Komponenten
- Klassisches Software Engineering



Integration des Modells



Breck, E., Cai, S., Nielsen, E., Salib, M., Sculley, D., 2017. The ML test score: A rubric for ML production readiness and technical debt reduction, in: 2017 IEEE International Conference on Big Data (Big Data). Presented at the 2017 IEEE International Conference on Big Data (Big Data), IEEE, Boston, MA, pp. 1123–1132. https://doi.org/10.1109/BigData.2017.8258038



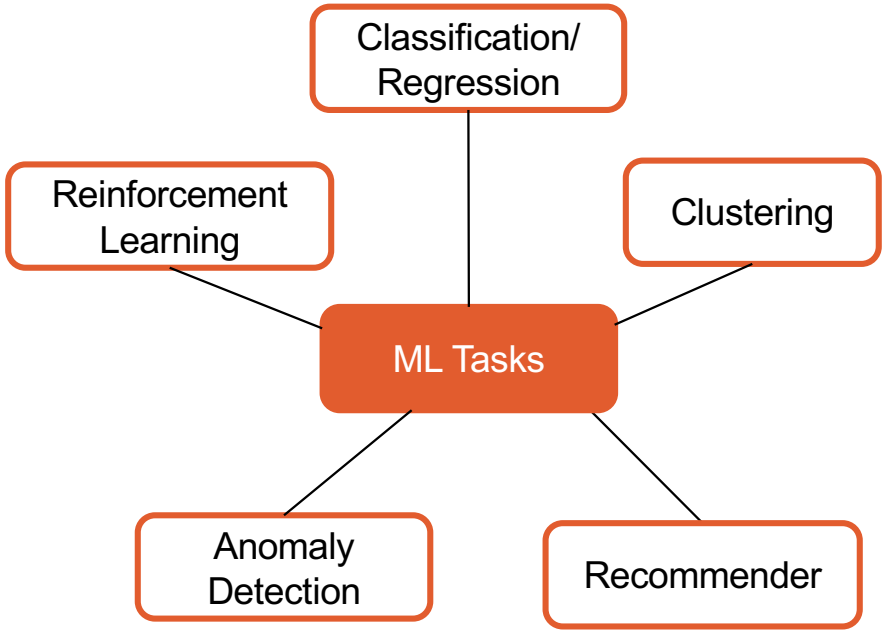
Betrieb

- Verhalten bei Sammlung weiterer Daten kontrollieren
  - Verhält sich das Model wie gewünscht?
  - Sind Entscheidungen/Vorhersagen/ Ergebnisse nachvollziehbar?
- Prüfen, ob das Model noch in der angedachten Umgebung einsetzt wird
  - Ist das Verhalten in neuer Umgebung erwünscht?

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Fazit

- Entwicklung von KI/ML-Anwendungen benötigen spezifische Methoden und Werkzeuge
- EDDA ist ein Ansatz, der die grundlegenden Tätigkeiten strukturiert
  → Spezialisierungen sind je nach Art der Anwendung notwendig
- Potential für den Einsatz von KI/ML im eigenen Unternehmen? Sprechen Sie mich an!



UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Fragen und Diskussion

Dr. Marc Hesenius
Universität Duisburg-Essen
marc.hesenius@uni-due.de
@mhesenius