

# WIE KOMMEN DIE DATEN IN DIE KI?

Theoretische Einführung und Best Practices für die Nutzung von KI auf Basis von Unternehmensdaten

Referenten: Jörg Schließer, Steven Lehmann

17.4.2024



# MENSCHLICH INTELLIGENT AGENDA

- Einführung
- Wie kommen Daten in die KI?
- Datenbeschaffung
- Praktische Ansätze
- Anwendungsbeispiele
- Gemeinsames Brainstorming

Jörg Schließer, Steven Lehmann 2 17.4.24



# MOTIVATION - BEST PRACTICE KI







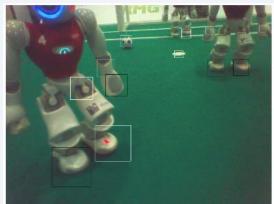


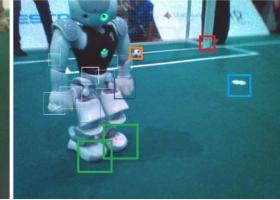


Kaggle A-Z





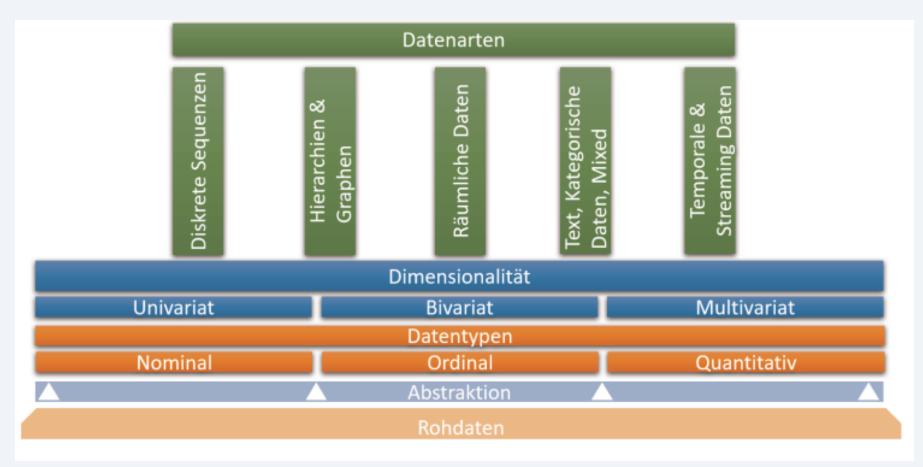




3 17.4.24 Jörg Schließer, Steven Lehmann



# EINFÜHRUNG - DATENTYPEN



Nazemi, K., Kaupp, L., Burkhardt, D., & Below, N. (2021). 5.4 Datenvisualisierung. *De Gruyter Praxishandbuch Ser, Walter de Gruyter GmbH: Berlin, Germany*, 477-502.

Jörg Schließer, Steven Lehmann 4 17.4.24



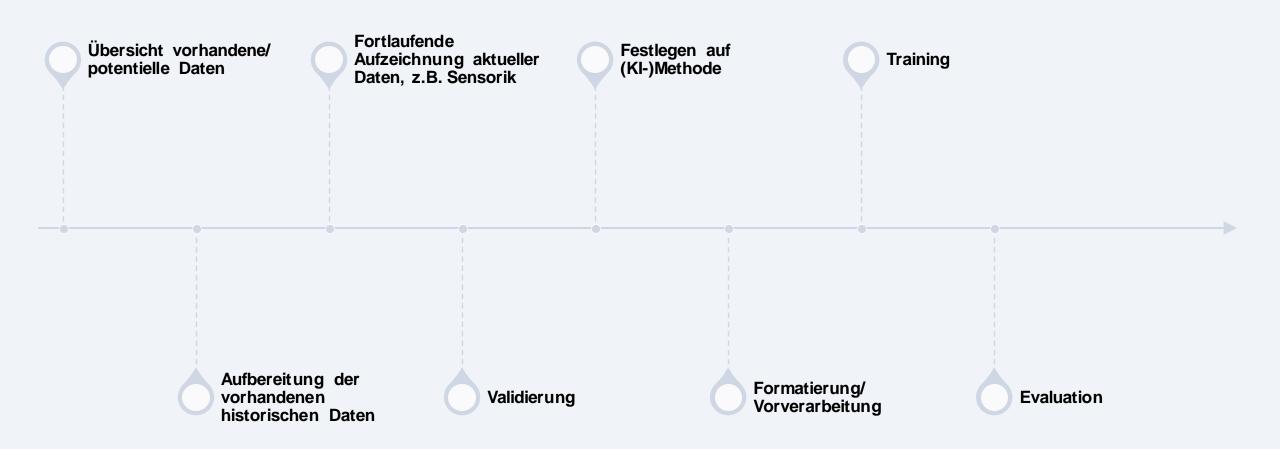
# MENSCHLICH INTELLIGENT EINFÜHRUNG



Jörg Schließer, Steven Lehmann 5 17.4.24



# PIPELINE DATEN-KI



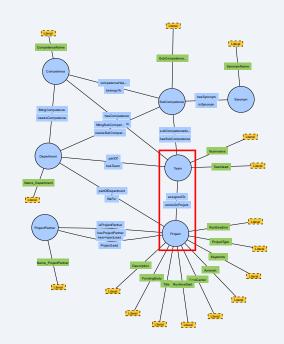
Jörg Schließer, Steven Lehmann 6 17.4.24

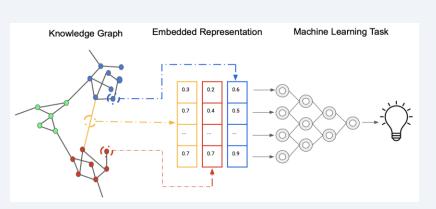


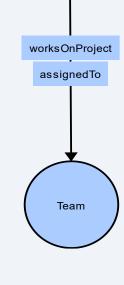
# PIPELINE - BEISPIEL

Kostenstelle 🗸	Kostenträger ~	Akronym	Titel	Projektinhalt ~	Industrie/FuE ~	DBI-Gruppe ~
6100	6821	H2-Transformation	H2-Transformation		Industrie	GUT
6100			H2Readyness		Industrie	GUT
6100			H2-Transformation		Industrie	GUT
6100			H2-Transformation		Industrie	GUT

#### Annotations: Projekt H2Readyness Annotations 🕕 rdfs:label Projekt H2Readyness bearbeitetVon localhost:8000/rdf/ontology#Team10 bearbeitetVon localhost:8000/rdf/ontology#Team8 hatProjektpartner localhost:8000/rdf/ontology#Projektpartner73 Kostenstelle 6100 Projektart GUT teilFG localhost:8000/rdf/ontology#Fachgebiet61 Projekt







Project

EdoardoRamalli. https://commons.wikimedia.org/wiki/File:KnowledgeGraphEmbedding.png

Jörg Schließer, Steven Lehmann 7 17.4.24



# DATEN IN UNTERNEHMEN



- Interne Dokumente
- Kundendaten
- Mitarbeiterdaten
- Maschinendaten

https://www.pipedrive.com/de/blog/big-data



### MENSCHLICH INTELLIGENT WOHER KOMMEN DIE DATEN?

#### Intern

- Digitalisierung alter Dokumente
- Aufzeichnen aktueller Prozesse

#### Extern

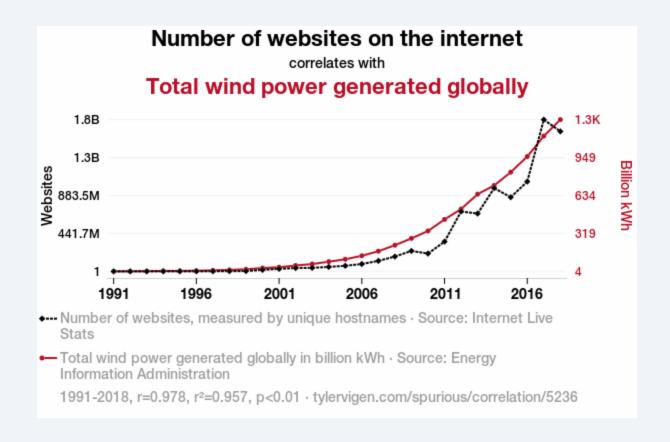
- o Kauf
- Kooperation mit Unternehmen, die ähnliche Daten / Interessen besitzen
- Öffentlich zugängliche Daten

Jörg Schließer, Steven Lehmann 9 | 17.4.24



# EINFÜHRUNG - SPAß MIT DATEN

- Korrelation vs. Kausalität beachten!
- Große Datenmengen bringen neue Herausforderungen



Jörg Schließer, Steven Lehmann



# LABELN

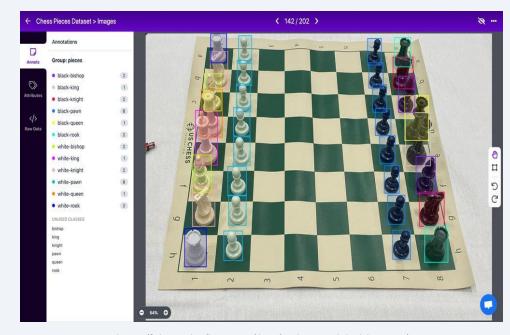
### Taskabhängig:

#### Klassifizierung:



X0	X1	X2	Xn	у
0	0	0	0	1
0	1	1	0	1
0	1	0	1	0
1	1	1	0	0

### Objekterkennung:

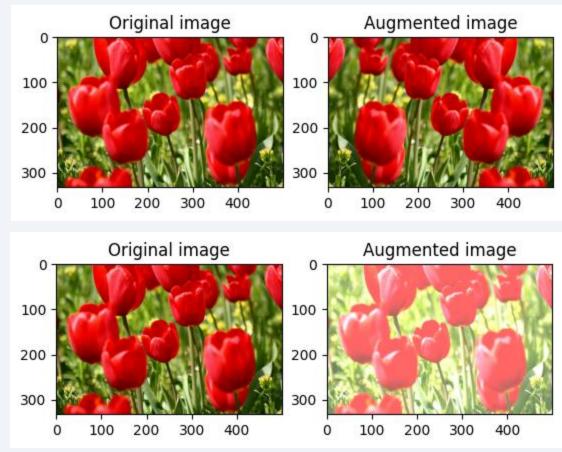


https://blog.roboflow.com/tips-for-how-to-label-images/

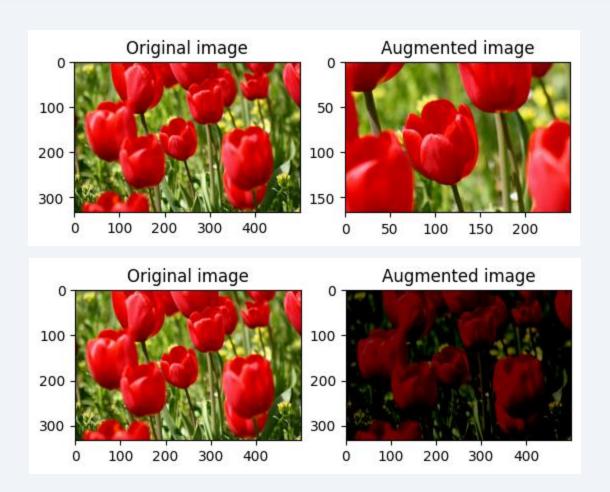
Jörg Schließer, Steven Lehmann 11 17.4.24



# AUGMENTIEREN



https://www.tensorflow.org/tutorials/images/data\_augmentation



Jörg Schließer, Steven Lehmann 12 17.4.24



# MENSCHLICH INTELLIGENT PRAKTISCHE ANSÄTZE



Jörg Schließer, Steven Lehmann





# **VIELEN DANK!**

#### KONTAKT

Tel:+49 341 9733430

Email: kmi-project@infai.org

www.kmi-leipzig.de

www.kmi-netzwerk.org



Das Forschungs- und Entwicklungsprojekt KMI wird im Rahmen der Fördermaßnahme "Zukunft der Arbeit: Regionale Kompetenzzentren der Arbeitsforschung – Künstliche Intelligenz" im Programm "Innovationen für die Produktion, Dienstleistung und Arbeit von morgen" des Bundesministeriums für Bildung und Forschung (BMBF) gefördert und vom Projektträger Karlsruhe (PTKA) betreut.



GEFÖRDERT VOM



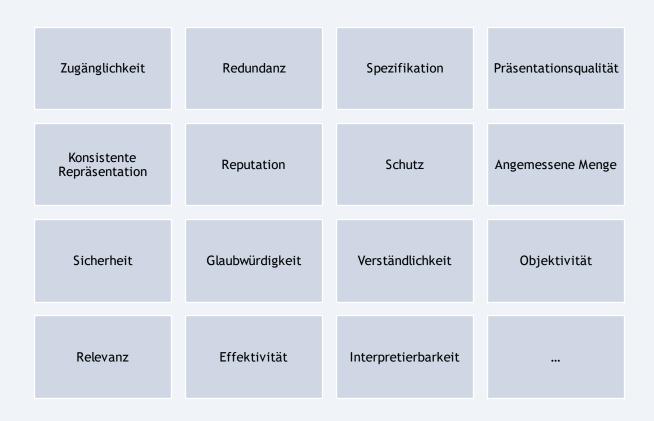


# BACKUP: DATENQUALITÄT

Aktualität Währung/Wert

Konsistenz Genauigkeit

Vollständigkeit

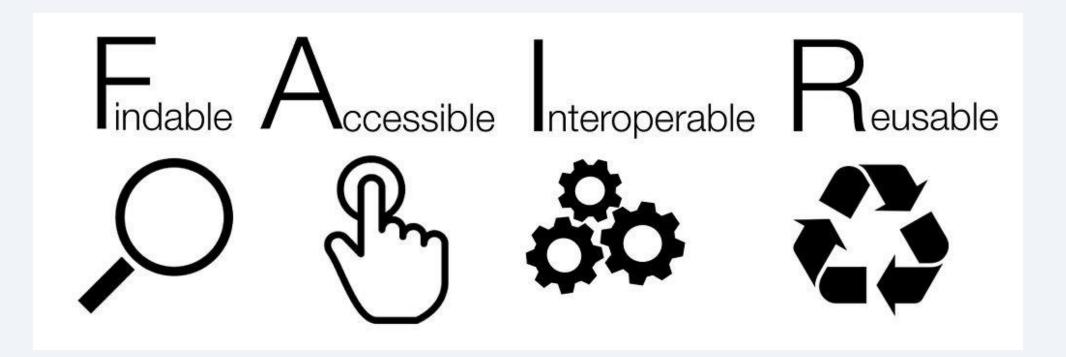


Sidi, F., Panahy, P. H. S., Affendey, L. S., Jabar, M. A., Ibrahim, H., & Mustapha, A. (2012, March). Data quality: A survey of data quality dimensions. In 2012 International Conference on Information Retrieval & Knowledge Management (pp. 300-304). IEEE.

Jörg Schließer, Steven Lehmann 16 | 17.4.24



# BACKUP: FAIR PRINZIPIEN



Jörg Schließer, Steven Lehmann 17 17.4.24