HERAUSFORDERUNGEN BEI HUMAN-CENTERED AI UND ETHISCHE ASPEKTE

Vortrag im Rahmen des Workshops AI for GOOD – AI Anwendungen in der Qualitätssicherung

Dr. Bernhard Moser
Software Competence Center Hagenberg
Foto: persönlich

Der Trend in der Industrie geht immer mehr Richtung Individualisierung statt Massenproduktion von Produkten und der Green-Tech Aspekt spielt mittlerweile ebenfalls eine große Rolle. Zudem soll der ökologische Fußabdruck verbessert werden. Für die Produktion bedeutet es häufigeres Umrüsten von Prozessen, Rekalibrieren von Prozessmodellen, insbesondere von Qualitätsmodellen. Dies hat auch einen Ressourcenaspekt, denn bessere Qualitätsmodelle führen zu weniger Ausschuss und dies führt zu mehr Nachhaltigkeit. Der Schlüssel für diese Lösung sind die Daten bzw. datengetriebene Modelle für Fehleranalysen, was durch maschinelles Lernen erreicht werden kann. Dafür werden nutzbare Daten gebraucht.

Des Weiteren werden Menschen in diese Prozesse der Umrüstung integriert, wobei man aber vorsichtig sein und die ethischen Aspekte (Schutz der Privatsphäre, Integrität, Unterstützung des Menschen bei seiner Arbeit) beachten muss.

ETHISCHE ASPEKTE

Human Centered AI ist Künstliche Intelligenz plus Ethik. Diese muss die Grundrechte beachten und damit einem ethischen Zweck dienen. Zudem muss das System robust und zuverlässig sein und der Mensch muss die Kontrolle über das System haben. Weitere Aspekte sind Privacy und Data Governance, Accountability (Verantwortlichkeit) und Transparenz, wo es um die Nachvollziehbarkeit von Entscheidungen geht.

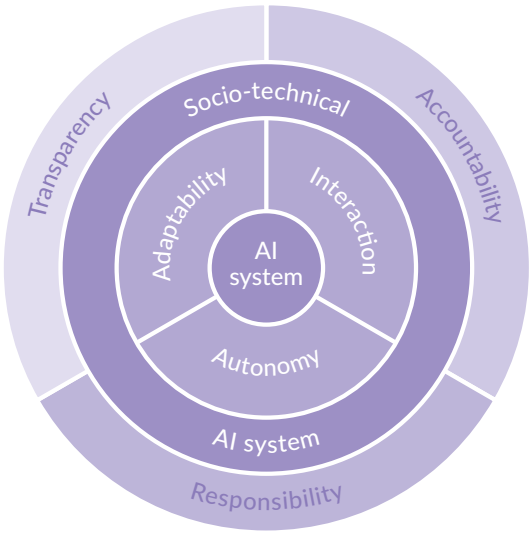

Abb. 1: die ART Aspekte von AI[1]

[1] Quelle: https://www.humane-ai.eu/wp-content/uploads/2019/11/D13-HumaneAI-framework-report.pdf

DATENPRIVACY SCHUTZ

Bei der Datenverarbeitung kann man mit AI Methoden wertvolle Informationen ableiten, jedoch soll verhindert werden, dass DSVGO relevante Daten abgeleitet werden können. Dafür müssen Methoden entwickelt werden, die dies verhindern (Privacy Preserving Machine Learning). Wie in der unteren Grafik dargestellt, besteht ein Verhältnis zwischen der Notwendigkeit an Datenschutz und der Notwendigkeit an Transferierbarkeit von Daten. Gerade in der Produktion von kleineren Losgrößen reichen Methoden wie Federated Learning nicht aus. Das Problem liegt in der Unterschiedlichkeit der Daten. Um diesen Mangel zu kompensieren braucht es flexiblere Methoden des Transfer Learning.

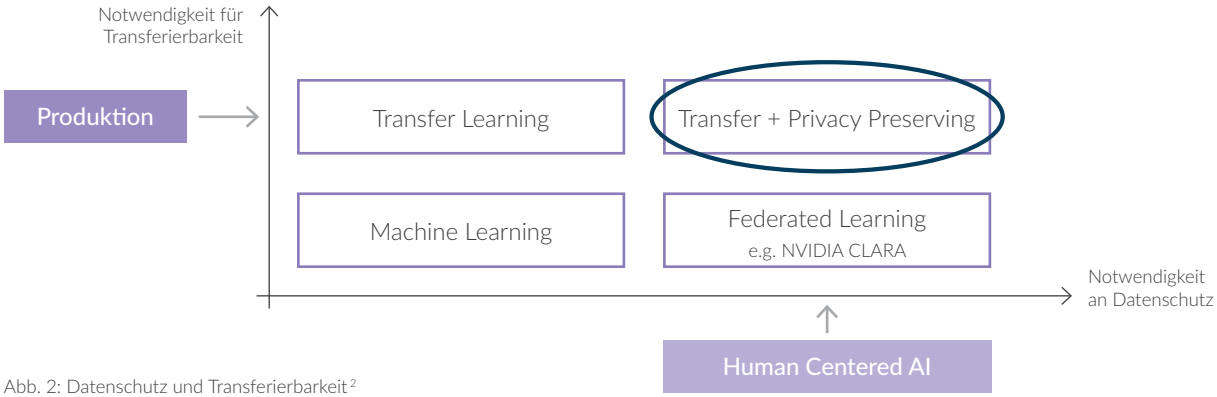

Abb. 2: Datenschutz und Transferierbarkeit[2]

BEDINGUNGEN FÜR MASCHINELLES LERNEN

Die Aufbereitung der Daten kann sehr aufwändig und kostspielig sein. Das Problem ist, dass generell in Österreich in der Produktion Open Data Sets nicht vorhanden sind. Jedes Unternehmen besitzt selbst ihre Datensilos aus Sicherheitsgründen. Die Datensätze haben zudem oft unterschiedlichen Charakteristika, was zu Hindernissen führt. Es stellt sich die Frage, wie man dieses Hindernis abmildern kann und ähnliche Innovationen wie im Bereich Autonomous Driving (Abb. 3), wahrnehmen kann.

	KITTI	Cityscapes	ApolloScape	Mapillary	BDD100K
# Sequences	22	~50	4	N/A	100,000
# Images	14,999	5000 (+2000)	143,906	25,000	120,000,000
Multiple Cities	No	Yes	No	Yes	Yes
Multiple Weathers	No	No	No	Yes	Yes
Multiple Times of Day	No	No	No	Yes	Yes
Multiple Scene types	Yes	No	No	Yes	Yes

Abb. 3: angereicherte Trainingsdaten im Bereich Autonomous Driving[3]

EXKURS TRANSFER LEARNING

In der Produktion müssen für jedes neue Teil die Trainingsdaten neu aufgesetzt werden und die Aufgabe neu definiert werden. Der Grundgedanke von Transfer Learning ist, dass bei neuen Teilen, die ein wenig anders aussehen, aber gewisse Aspekte das System bereits gelernt hat (z. B. Oberfläche, Ausprägungen), die Daten wiederverwendet werden können. Diese Methode hat ein großes Potential für Anwendungen in der Produktion, ist aber noch in der Entwicklungsphase. Dabei müssen zusätzlich die Privacy Aspekte berücksichtigt werden.

[2] Quelle: Bernhard Moser, SCCH, 2020
[3] Quelle: https://bdd-data.berkeley.edu/