

Künstliche Intelligenz: KI Karten - eine Möglichkeit zur Strukturierung

Roland Konopac & Dr. Ulli Waltinger

Siemens AG

© 2021 Engineering und IT Tagung 2021 - Chemnitz

siemens.com





*Die datengetriebene
Gesellschaft: Zum ersten Mal
in der Geschichte können wir
genug über uns selbst
sehen, um ...*

*... Orte und Produkte bauen,
die besser funktionieren als
die, die wir (möglicherweise)
schon immer hatten*



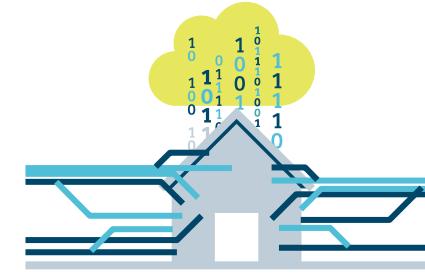
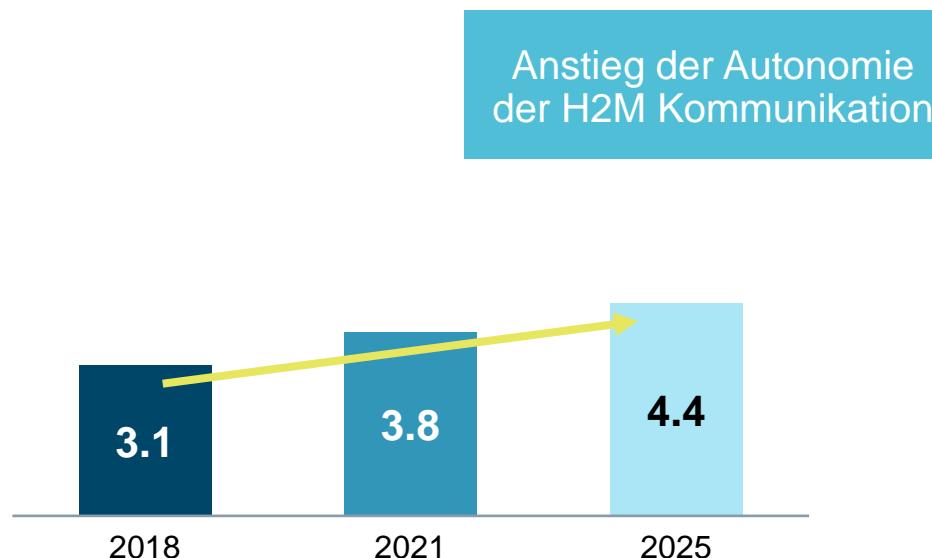
Artwork: Chris Labrooy, Darty, Gadget Portrait

SIEMENS

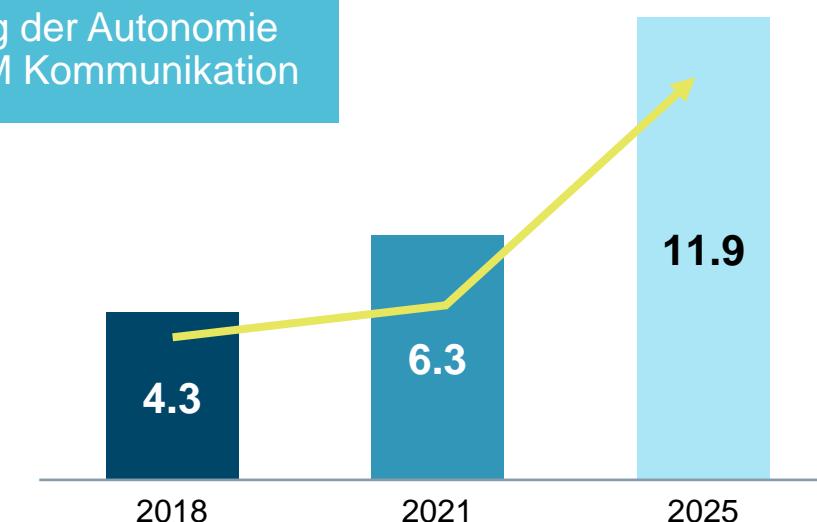
Den digitalen Wandel gestalten - Intelligente Vernetzung vorantreiben



Wachstum vernetzter Menschen



Wachstum vernetzter Maschinen



*Zahlen in Mrd

Auf welche Werte hin optimieren wir technische oder Softwaresysteme?

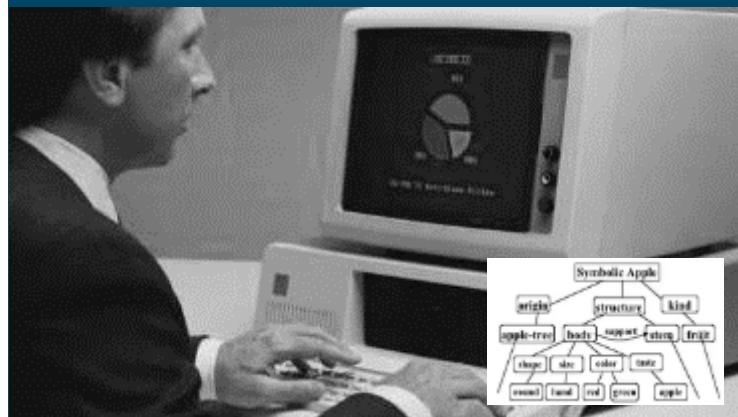
$$C(\mathbf{y}, \mathbf{o}) = \frac{1}{N} \sum_{i=1}^N (y_i - o_i)^2$$



Die datengetriebene Gesellschaft vom Bauchgefühl zur Selbstkontrolle

Human-driven

- Systeme tun nur das, was der Mensch explizit definiert
- ... daher auf den menschlichen Zustand angewiesen



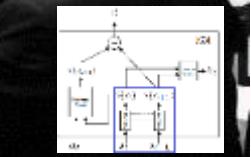
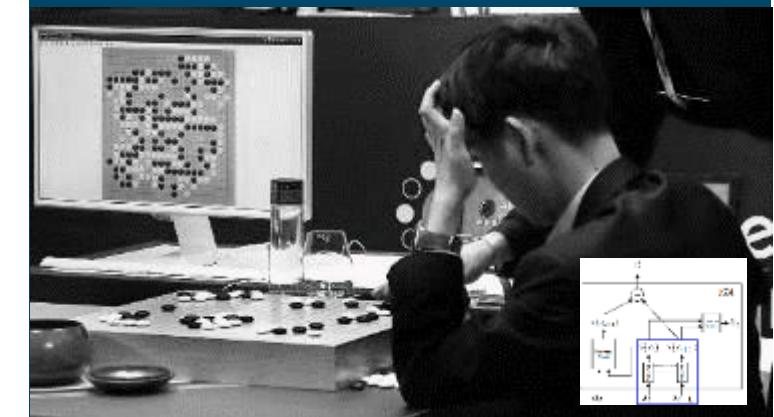
Data-driven

- Systeme tun nur das, was sie aus Daten lernen
- ... daher auf den Datenzustand angewiesen



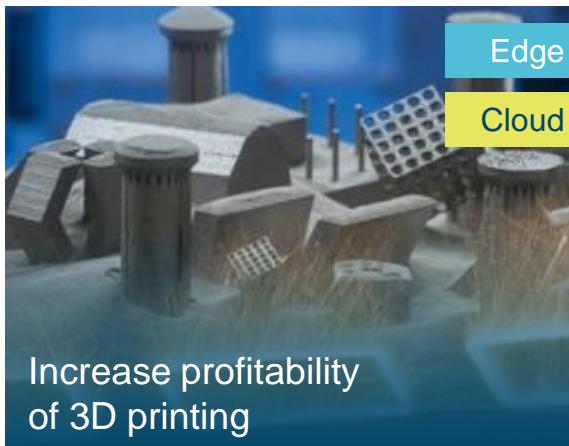
System-driven

- Systeme erkunden ihre Umgebung autonom – nutzen ihre Neugierde aus
- ... daher auf Selbstüberwachung und Systemzustand angewiesen



SIEMENS

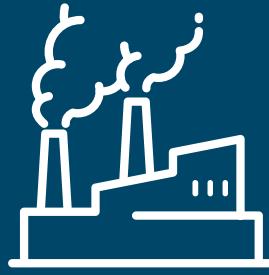
Umfassender Einsatz von Künstlicher Intelligenz entlang aller Anwendungsbereiche



Zur Konvergenz von Wertschöpfungsketten in AoT



+



=



Enterprise AI

nutzt Künstliche Intelligenz und Automatisierung, um die **Digitale Transformation** von **Geschäftsprozessen** voranzutreiben

Industrial AI

nutzt Künstliche Intelligenz und Steuerung, um die **Digitale Transformation** von **Industriellen Anwendungen** voranzutreiben

Ermöglicht neue Geschäftsmodelle, die über die gesamte Wertschöpfungskette des Unternehmens skalieren



AoT wird die Wertschöpfung auf Geräte- und Prozessebene verstärken und kombinieren, um Effizienz zu steigern, Lieferketten optimieren, Qualität zu verbessern und Kosten zu senken

KI-Lösungen basieren auf leistungsstarken Technologien und bieten bei verantwortungsvollem Einsatz mehrere Vorteile



**AI-driven
coaching?**

oder

**Mitarbeiter
überwachen?**



**Helfen Sie uns,
bessere
Entscheidungen
zu treffen?**

oder

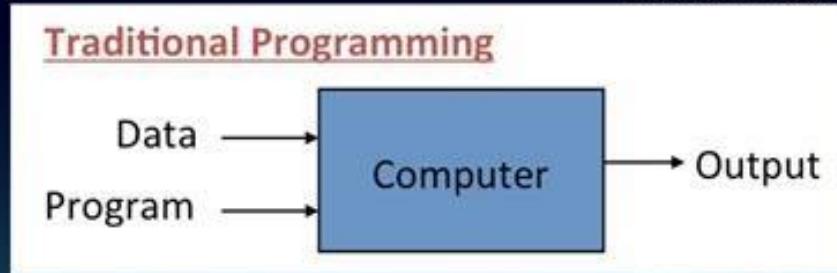
**Überwältigen
Sie uns mit
ihren
Fähigkeiten?**



Artificial Intelligence

“Software is eating the world...”

Marc Andreessen



“....but AI Is Going to Eat Software”

Jensen Huang (CEO Nvidia)



Why now?

- Big Data
- Powerful and cheap hardware
- Deep learning algorithms

“If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future.”

Andrew Ng

SIEMENS



KI verändert Euer Verhalten

Blackbox

- Ihr gebt ein Stück freien Willen ab
- Ihr verzichtet auf Transparenz, Vergleichbarkeit und Vielfalt
- Ihr akzeptiert Beurteilungsfehler
- Ihr akzeptiert Optimierung als obersten Sinn des Handelns
- Ihr habt weniger Empathie

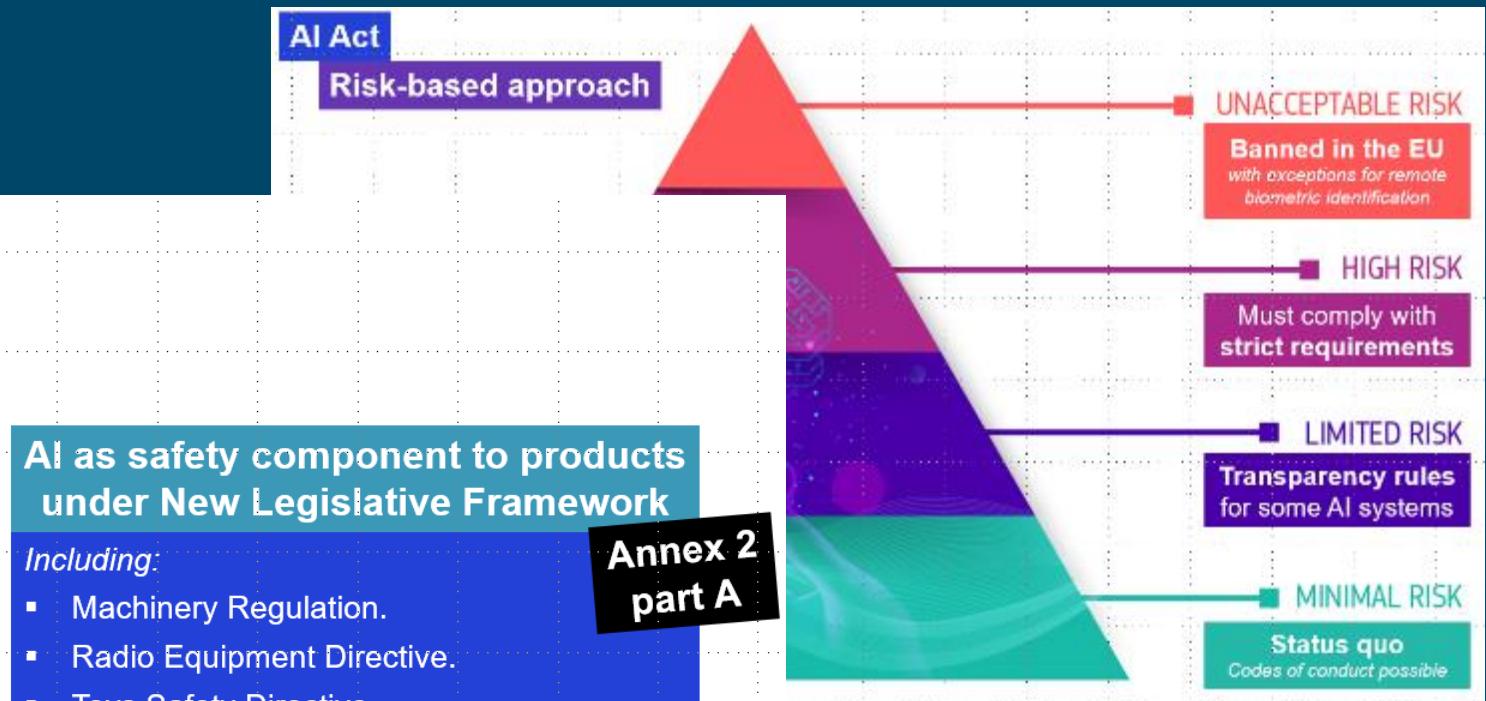
Freiwillige Selbstbeschränkungen der entwickelnden Firmen
(AI Ethics Teams)

„Since early last year, Google has also blocked new AI features analyzing emotions, fearing cultural insensitivity, while Microsoft restricted software mimicking voices and IBM rejected a client request for an advanced facial-recognition system.“



Digitales Europa

AI Act	Scope – High-risk AI Article 6	Annex 3
	8 categories of intended use of AI 1. Biometric identification of persons. 2. Critical infrastructure management. 3. Education (admission tests & selection processes). 4. Employment (HR/recruitment processes). 5. Essential services (access to social benefits or credit, emergency first response). 6. Crime prediction & response. 7. Migration, asylum & border control. 8. Justice.	
	Update possible within 8 categories if risk for health & safety or impact on fundamental rights, based on criteria of art. 7(2)	Article 7



Vertrauenswürdige KI

1 Vorrang menschlichen Handelns und menschliche Aufsicht

z.B. Grundrechte, Vorrang menschlichen Handelns und menschliche Aufsicht

2 Technische Robustheit und Sicherheit

z.B. Widerstandsfähigkeit gegen Angriffe und Sicherheitsverletzungen, Auffangplan und allgemeine Sicherheit, Präzision, Zuverlässigkeit und Reproduzierbarkeit

3 Schutz der Privatsphäre und Datenqualitätsmanagement

z.B. Achtung der Privatsphäre, Qualität und Integrität der Daten sowie Datenzugriff

4 Transparenz

z.B. Nachverfolgbarkeit, Erklärbarkeit und Kommunikation

5 Vielfalt, Nichtdiskriminierung und Fairness

z.B. Vermeidung unfairer Verzerrungen, Zugänglichkeit und universeller Entwurf sowie Beteiligung der Interessenträger

6 Gesellschaftliches und ökologisches Wohlergehen

z.B. Nachhaltigkeit und Umweltschutz, soziale Auswirkungen, Gesellschaft und Demokratie

7 Rechenschaftspflicht

z.B. Nachprüfbarkeit, Minimierung und Meldung von negativen Auswirkungen, Kompromisse und Rechtsbehelfe



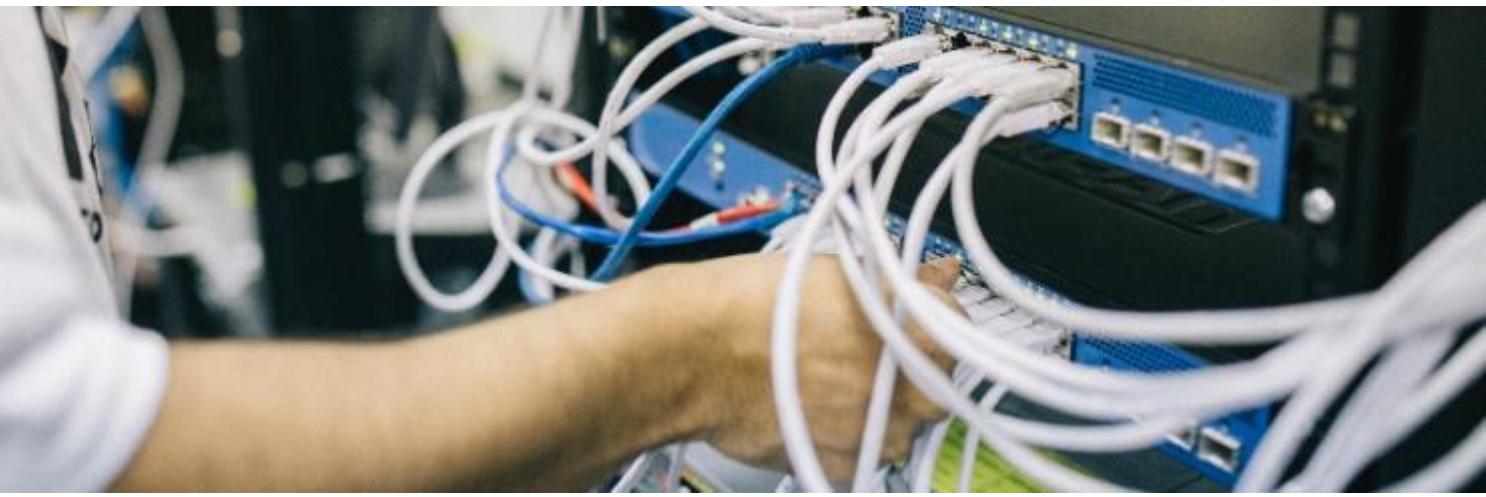
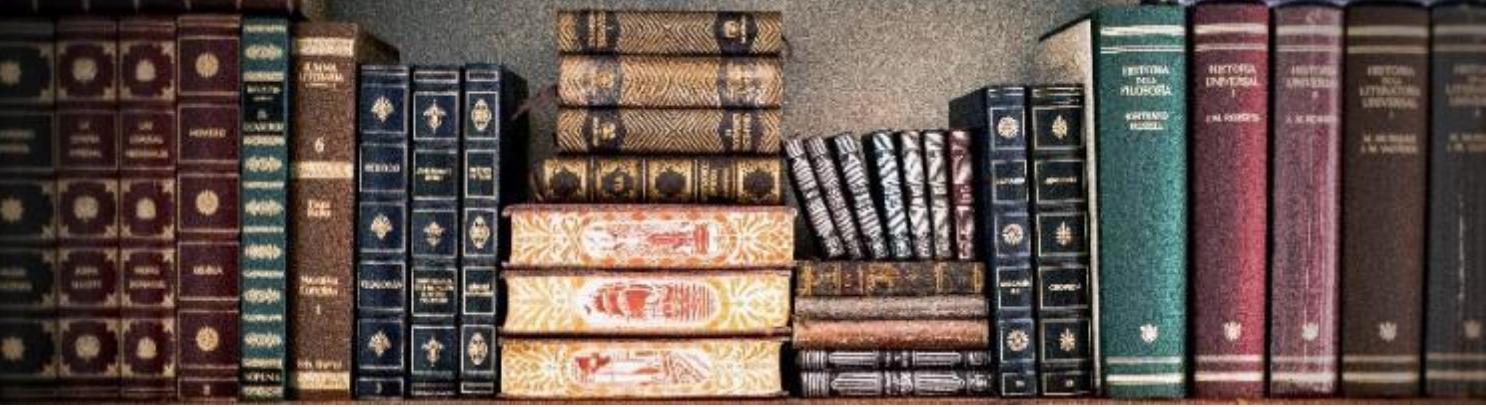
Möglichkeiten für den Betriebsrat

- Mitbestimmung der DV Ausschüsse u.a. bei Leistungs- und Verhaltenskontrolle gemäß § 87 Abs. 1 Ziff. 6 BetrVG,
- Informations- und Beratungsrechte bei der Planung von Arbeitsverfahren und Arbeitsabläufen, einschließlich des Einsatzes von Künstlicher Intelligenz § 90 Abs. 1, Ziff. 3 BetrVG und
- Überwachung der DSGVO, Need to know
- Performance der Künstlichen Intelligenz beeinflussbar durch:
 - Änderungen am Algorithmus selbst
 - Bewertung der Daten
 - Die zur Verfügung stehenden Daten
- Es gilt die Hinzuziehung eines Sachverständige als erforderlich § 80 Abs. 3 BetrVG
- Wildwest der KI Anbieter / Startups



AI Cards in der Siemens AG

- Ausfüllen der AI Cards durch die Bedarfsträger / Verantwortlichen
 - Selbstreflektion
 - Dokument für Entscheider und den Betriebsrat
- Inhouse Consulting, incl. des Monitorings der Künstlichen Intelligenz
- Eskalationsinstanz zur Künstlichen Intelligenz (Ethikrat und Unternehmenskultur)

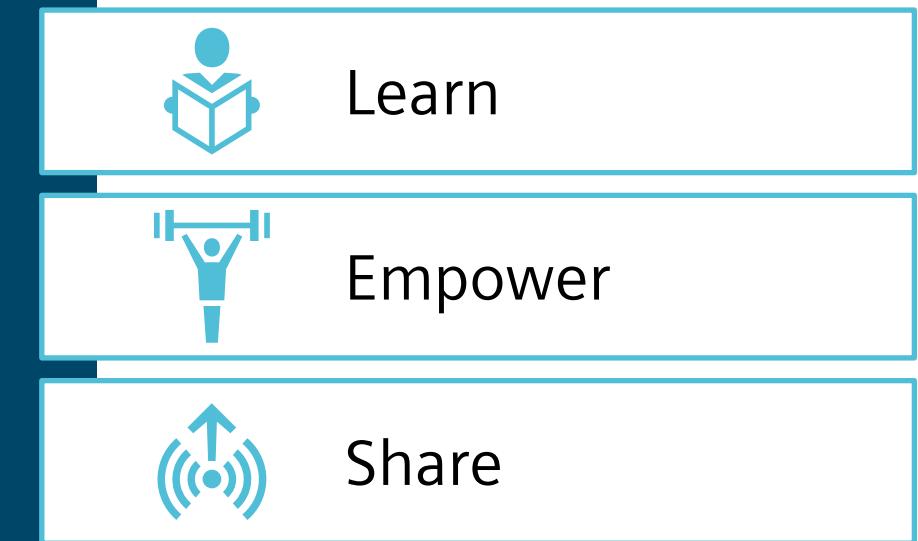


Was sind „KI Karten“ (AI Cards)

„AI Application Cards“ sind ein Ansatz, welches ermöglichen soll, die wesentlichen Fakten, Nutzen und möglicher Risiken, von Anwendungen aus dem Bereich der Künstlichen Intelligenz & Machine Learning in strukturierter Weise zu dokumentieren und für den gemeinsamen Diskurs zu kontextualisieren.

Ansatz:

- KI-Lösungen basieren auf leistungsstarken Technologien und bieten bei verantwortungsvollem Einsatz vielfältige Vorteile
- Gemeinsam lernen von dem was funktioniert und was mit Risiken behaftet ist
- Gemeinsam Fähigkeiten aber auch Grenzen von angewandten KI Technologien einzuschätzen lernen
- Teilen von Erfahrungen und Best Practices um Aspekte der Unsicherheit und Bias zu reduzieren, und gleichzeitig Innovations- und Explorations-Mechanismen zu erhöhen



AI Card

XXXXX

1 | Anwendung

- Kurze Beschreibung zur Anwendung, Modell oder Service?
- Was ermöglicht die Anwendung?

2 | Organisation

- Welche Organisation/Firma stellt die Software bereit?
- Wie etabliert ist die Firma mit welchen Referenzen?

3 | Service

- Welche Services, Modelle oder Algorithmen werden aus dem Angebot verwendet?
- Welche Features/Merkmale haben die jeweiligen Services?
- Um welche Software / welche Komponenten geht es?

4 | Algorithmen & Daten

- Welche konkreten Algorithmen werden verwendet?
- Auf welchen Daten und wie wurden die Algorithmen trainiert?
- Werden weitere Daten zur Verfügung gestellt?
- Gibt es Referenz-Implementierungen (open source et al)?
- Welche Aussage trifft die Software und mit welchem Wahrheitsgehalt?
- Nach welchen Kriterien entscheidet die Software?
- Wie wurde das ML-Verfahren getestet?
- Wie ist die Qualität der Implementierung (des Programmcodes) sichergestellt?

AI Card

XXXXX

5 | KI Aufgabe

- | | | |
|--|--|---|
| <input type="checkbox"/> Automatisierung | <input type="checkbox"/> Prognose | <input type="checkbox"/> Visualisierung |
| <input type="checkbox"/> Outlier Detection | <input type="checkbox"/> Klassifikation | <input type="checkbox"/> Generation |
| <input type="checkbox"/> Optimierung | <input type="checkbox"/> Information Retrieval | |

6 | Verwendete Daten

- | | | | |
|-------------------------------------|--------------------------------|---------------------------------------|--|
| <input type="checkbox"/> Bilder | <input type="checkbox"/> Text | <input type="checkbox"/> Tabellarisch | <input type="checkbox"/> Mitarbeiter-Daten |
| <input type="checkbox"/> Zeitreihen | <input type="checkbox"/> Graph | <input type="checkbox"/> Transaktion | <input type="checkbox"/> Kunden-Daten |

- Welche Daten/Merkmale werden für die Evaluierung verwendet?
- Auf welche Daten hat die Software Zugriff?
- Welche Kriterien für die Daten-Auswahl wurden herangezogen?

7 | Verwendungszweck

Primärer Verwendungszweck

- Welche menschliche Entscheidung soll durch die Anwendung unterstützt werden?
- Für welchen Zweck wird die Anwendung verwendet?
- In welche Prozess wird die Software integriert

Primäre beabsichtigte Benutzer

- Welche Personen/Profile verwenden die Anwendung?
- Wie verwenden die Anwender die Software?
- Welche Fähigkeiten und Kenntnisse werden auf Seiten der AnwenderInnen der Software benötigt?

AI Card

XXXXX

8 | Reifegrad

- Demo
- Pilot
- PoC
- Roll-Out

9 | Anwendung

- Product
- Service
- Solution

10 | Leistungsmetriken

- Wie werden die Ergebnisse ausgewertet?
- Welche KPIs werden für die Evaluierung herangezogen?
- Wer legt die Kennzahlen fest, anhand derer in der Software Ziele definiert werden?
- Wie transparent ist der Entscheidungsweg?

11 | Ethische- & Risikoüberlegungen

- Wurde eine Risikoabschätzung vorgenommen?
- Wie wird Bias in den Daten transparent gemacht?
- Wie bewerten wir das Missbrauchspotenzial?
- Identifiziert der Datensatz irgendwelche Teilstichprobengruppen (z. B. nach Alter, Geschlecht)?
- Wie werden Mitarbeiter geschult um Algorithmen zu interpretieren?
- Werden mögliche subtile Beeinflussungen durch die Gestaltung der Softwareoberfläche ausgeschlossen?

12 | Vorsichtsmaßnahmen und Empfehlungen

- Welche Risiken entstehen, wenn die Anwendung in die Skalierung kommt?
- Wofür sollte die Anwendung nicht eingesetzt werden?
- Können automatische Entscheidungen korrigiert werden/wie?

13 | Verantwortlicher Kontakt

Name: Domain Expert

Name: IT Application

SIEMENS

KI-Lösungen basieren
auf leistungsstarken
Technologien und
bieten bei
verantwortungsvollem
Einsatz mehrere
Vorteile



A photograph of a man and a woman in a natural setting. The man, in the foreground, is seen from the waist up, wearing a white long-sleeved shirt, yellow pants, and a brown leather belt with a multi-colored striped suspender. He is holding a bouquet of wildflowers. The woman, in the background, is wearing a pink floral dress and is slightly out of focus. The background is a lush, green, overgrown hillside.

This is a **call**

for engagement

Vielen Dank!

