

ProKI - Dresden

KI-Demonstrations- und Transferzentrum Umformtechnik

16. InfoPoint

Wie arbeiten Mensch und Computer intelligent zusammen?

Demonstrations- und Transfernetzwerk KI in der Produktion

Struktur des Netzwerks

ProKI-Netz = Kostenloses Angebot für Unternehmen
KI-Systeme: Neue Werkzeuge für die Produktion

Projektlaufzeit
 01.10.2023 bis 31.12.2024

ProKI-Hannover
 ProKI-Berlin
 ProKI-Karlsruhe
 ProKI-Aachen

TRENNEN

ProKI-Aachen

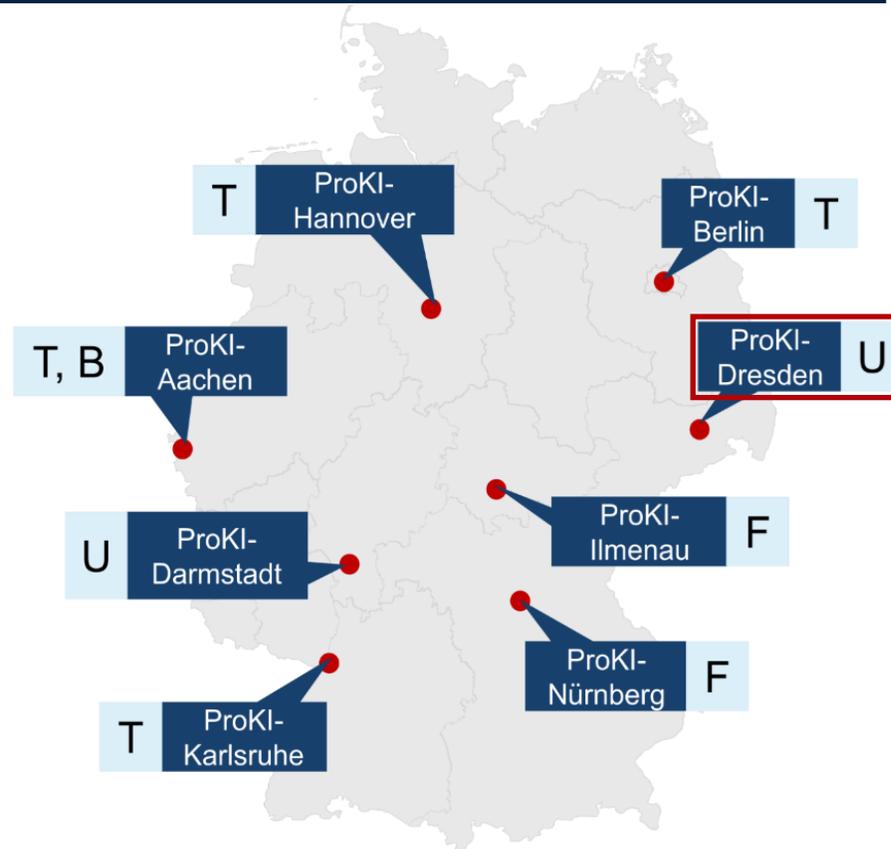
BESCHICHTEN

ProKI-Ilmenau
 ProKI-Nürnberg

FÜGEN

ProKI-Darmstadt
 ProKI-Dresden

UMFORMEN



GEFÖRDERT VOM



Bundesministerium
 für Bildung
 und Forschung

PROKI
 D R E S D E N

16. ProKI-InfoPoint

Agenda

Carolin Böhme

CIMTT Zentrum für Produktionstechnik und Organisation, TU Dresden

KI und Mensch – wo bleibt der humane Faktor?

Dr. Felix Gräßer

Linguwerk GmbH – alphaspeech

Intuitive und berührungslose Benutzerinteraktion durch intelligente Sprachinterfaces

Edgar Scherstjanoi

Professur für Arbeitswissenschaft, TU Dresden

Labelling – Das Menschliche in der künstlichen Intelligenz

Carolin Böhme

KI und Mensch – wo bleibt der humane Faktor?

Wie menschenzentrierte Entwicklung, Transparenz und Ethik den Weg für den Erfolg ebnen



M. Sc. Carolin Böhme

CIMTT – Zentrum für Produktionstechnik und Innovation der TU Dresden

KI und Mensch – wo bleibt der humane Faktor?

Wie menschenzentrierte Entwicklung, Transparenz und Ethik den Weg für den Erfolg ebnen

ProKI Infopoint

Dresden, 14.03.2024

Agenda

1. Einleitung: Die Bedeutung von KI für Unternehmen



2. Block I: Menschzentrierte und partizipative KI-Entwicklung



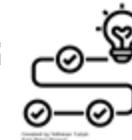
3. Block II: Erklärbare KI – Vertrauen durch Transparenz



4. Block III: Ethische Hintergründe zum KI-Einsatz

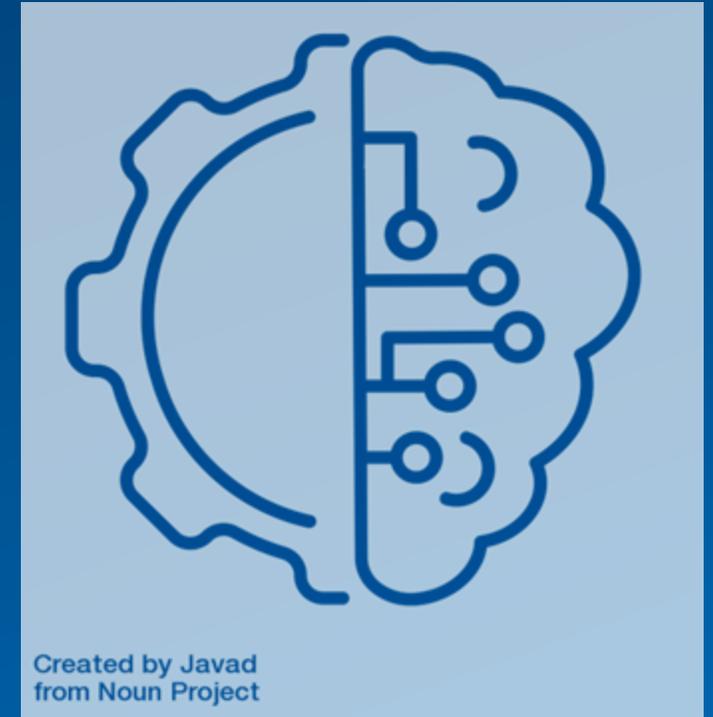


5. Schlussfolgerung: Die Zukunft einer verantwortungsvollen KI-Nutzung



Einleitung

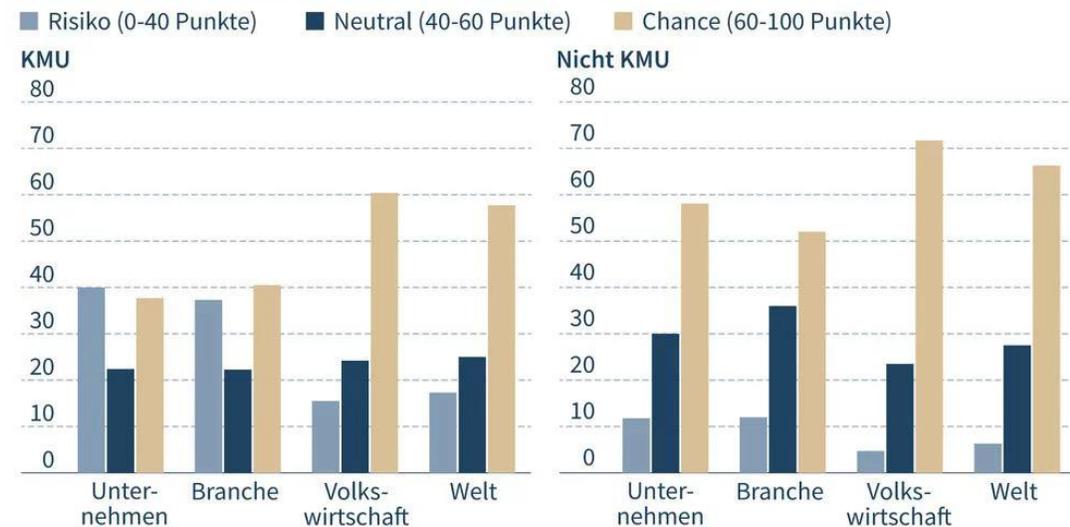
Die Bedeutung von KI für Unternehmen



Die Bedeutung von KI für Unternehmen

Verteilung der Einschätzung der Bedeutung von KI

Anteil der Unternehmen an den Bewertungen in Prozent



1) Frage: Wie bewerten Sie die Bedeutung von Künstlicher Intelligenz? Bewertung auf einer Skala von 0 Punkten (maximales Risiko) bis 100 Punkten (maximale Chance). 1. Quintil (0-20 Punkte) und 2. Quintil (20-40 Punkte) bedeuten Risiko; 4. Quintil (60-80 Punkte) und 5. Quintil (80-100 Punkte) bedeuten Chance.
Quelle: IW-Zukunftspanel, 33. Welle, 2019

- Rasante Entwicklung
- Mittelstandsskepsis
- Spannungsfeld: KI als Chance vs. KI als Risiko
- 9% der KMU nutzen Methoden der KI (16,5% GU)
- Herausforderung: KI im unternehmerischen Kontext noch nicht greifbar

Block I

Menschzentrierte und partizipative KI-Entwicklung



Menschzentrierte und partizipative KI-Entwicklung



Menschzentrierte Entwicklung bedeutet, dass die Technik am Menschen orientiert wird und nicht nur um der Technik willen umgesetzt wird.



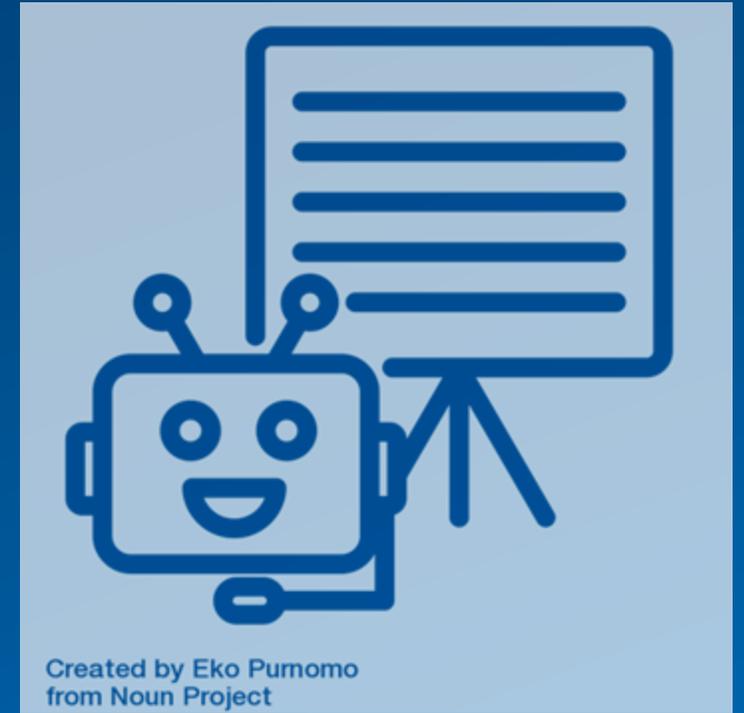
Partizipation bedeutet, dass die Mitarbeitenden (ggf. auch weitere Stakeholder) von Beginn an in die Entwicklung der Technologien und relevante Entscheidungen einbezogen werden, um Bedenken und Vorbehalte frühzeitig erkennen und beheben zu können.



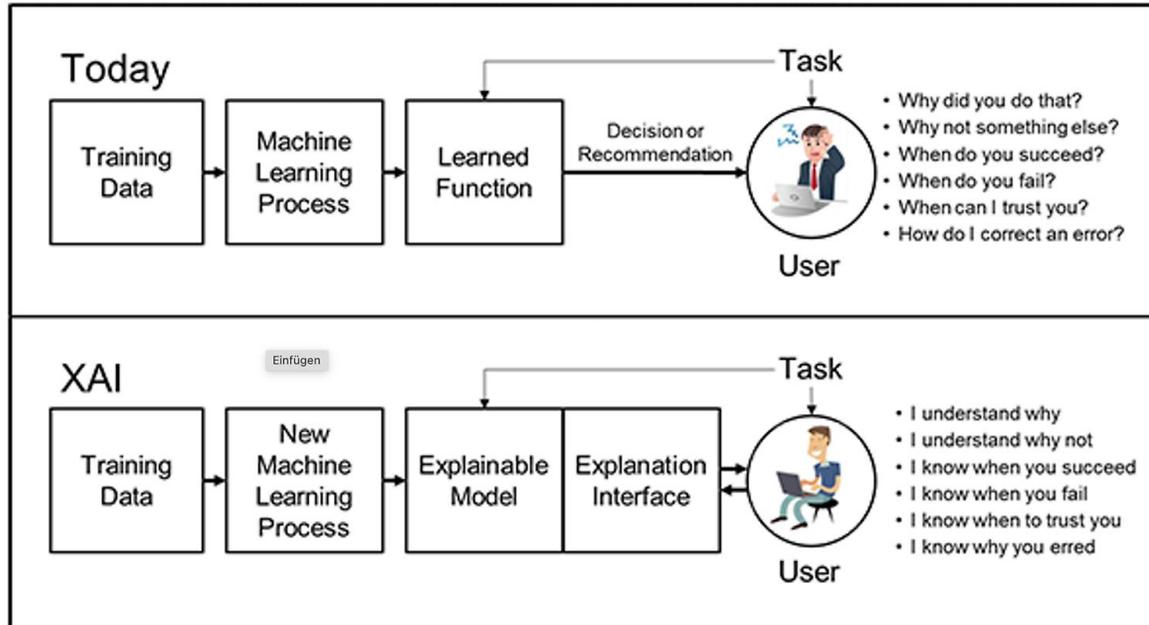
Partizipative Technologieentwicklung im Projekt ProKI Dresden

Block II

Erklärbare KI – Vertrauen durch Transparenz

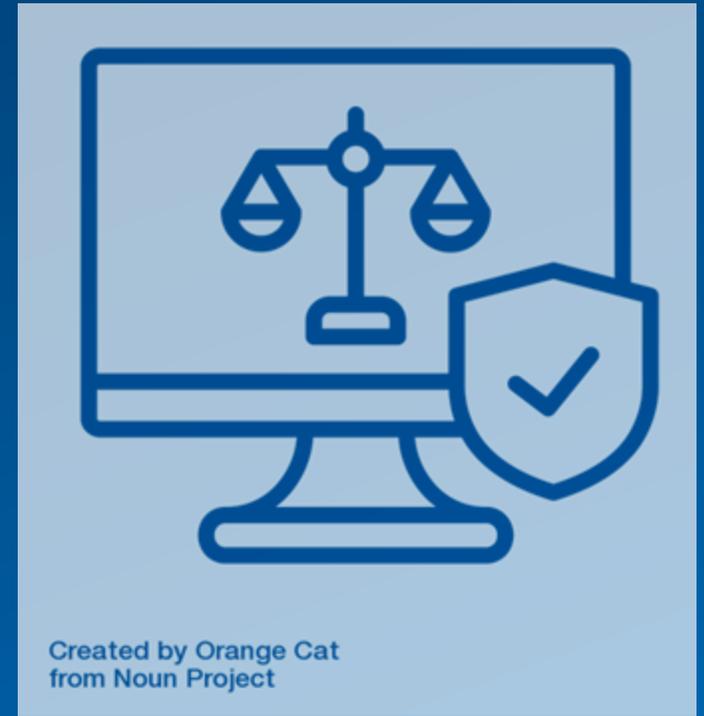


Erklärbare KI – Vertrauen durch Transparenz



Warum erklärbare KI?

1. Vertrauen und Akzeptanz
2. Fehlererkennung und Fehlerbehebung
3. Ethik und Rechenschaftspflicht
4. Vermeidung von Bias und Diskriminierung



Block III

Ethische Hintergründe zum KI-Einsatz

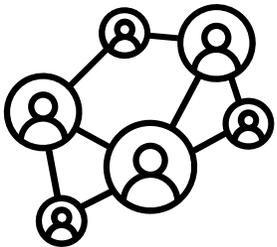
Ethische Hintergründe zum KI-Einsatz



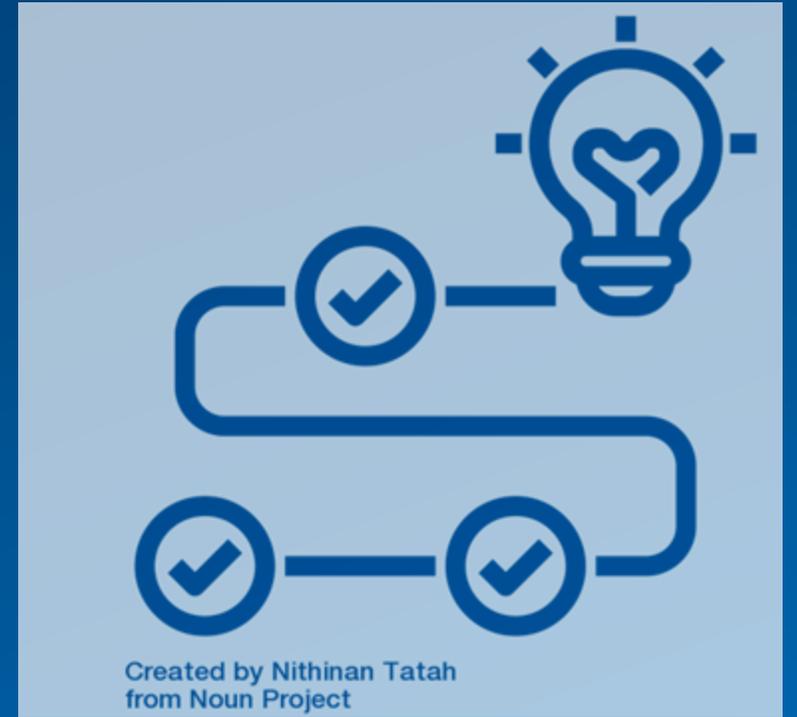
Fairness und Gerechtigkeit



Privatsphäre und Datenschutz



Verantwortung und Rechenschaftspflicht



Schlussfolgerung

Die Zukunft einer verantwortungsvollen KI-Nutzung

Die Zukunft einer verantwortungsvollen KI-Nutzung

Menschzentrierte & partizipative Entwicklung

Einbindung verschiedener Interessengruppen + Berücksichtigung von Bedürfnissen und Interessen der Nutzer
= erfolgreiche technische Entwicklung



Created by Phonlaphat Thongsiphong
from Noun Project

Erklärbare KI

Transparenz +
Nachvollziehbarkeit
= verstärktes Nutzervertrauen und Akzeptanz



Created by Eko Purnomo
from Noun Project

Ethische Ansätze

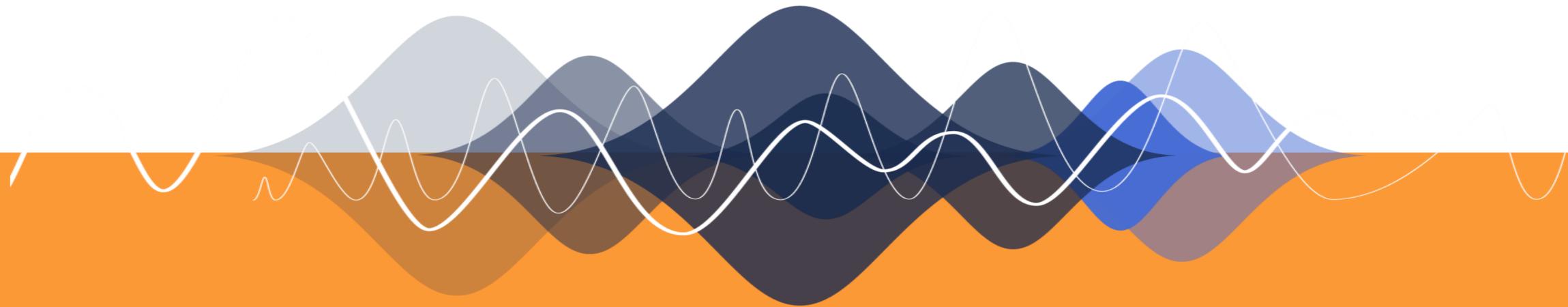
Fairness +
Datensicherheit
= höhere Anwendungswahrscheinlichkeiten und Akzeptanz



Created by Orange Cat
from Noun Project

Dr. Felix Gräßer
Intuitive und berührungslose Benutzerinteraktion durch intelligente Sprachinterfaces

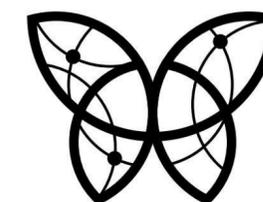
alphaspeech



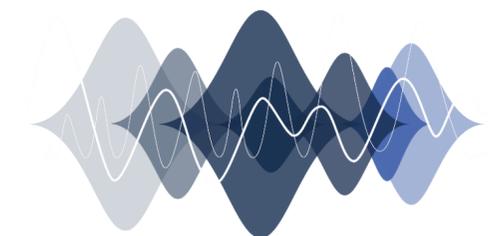
**Intuitive und berührungslose
Benutzerinteraktion durch
intelligente Sprachinterfaces**

Who we are

- Trademark and Business Unit of **Linguwerk GmbH**
- Spin-off the **Institute for Acoustics and Speech Communication, TU Dresden (2011)**
- International team of 30 software engineers, data scientists, computer linguists and economists
- Based in **Dresden, Leipzig and Kobe (Japan)**



EMBEDDED AI



alphaspeech

Spracherkennung. Made in Germany.

Wie arbeiten Mensch und Computer intelligent zusammen?



Steuerung von Maschinen
und Anlagen

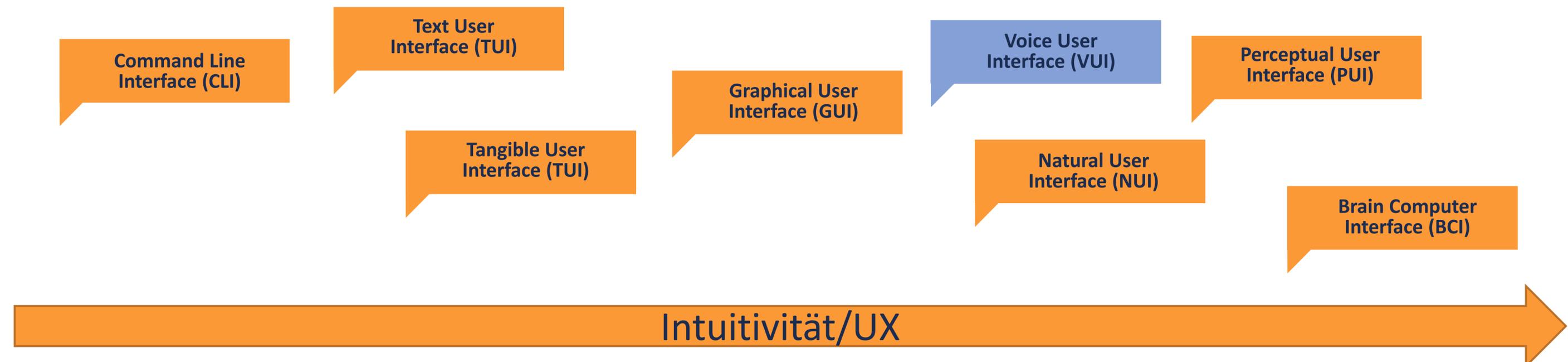
Eingabe von Daten und
Parametern

Abfrage und Darstellung
von Zuständen

Was macht ein User Interface (UI)?

Erfüllt hauptsächlich zwei Funktionen:

- Annahme von **Eingaben** über Geräte wie Tastatur, Maus, Touchscreen, Kamera oder Mikrofon
- Darstellung der **Ausgabe** über Geräte wie Monitore, Drucker oder Lautsprecher

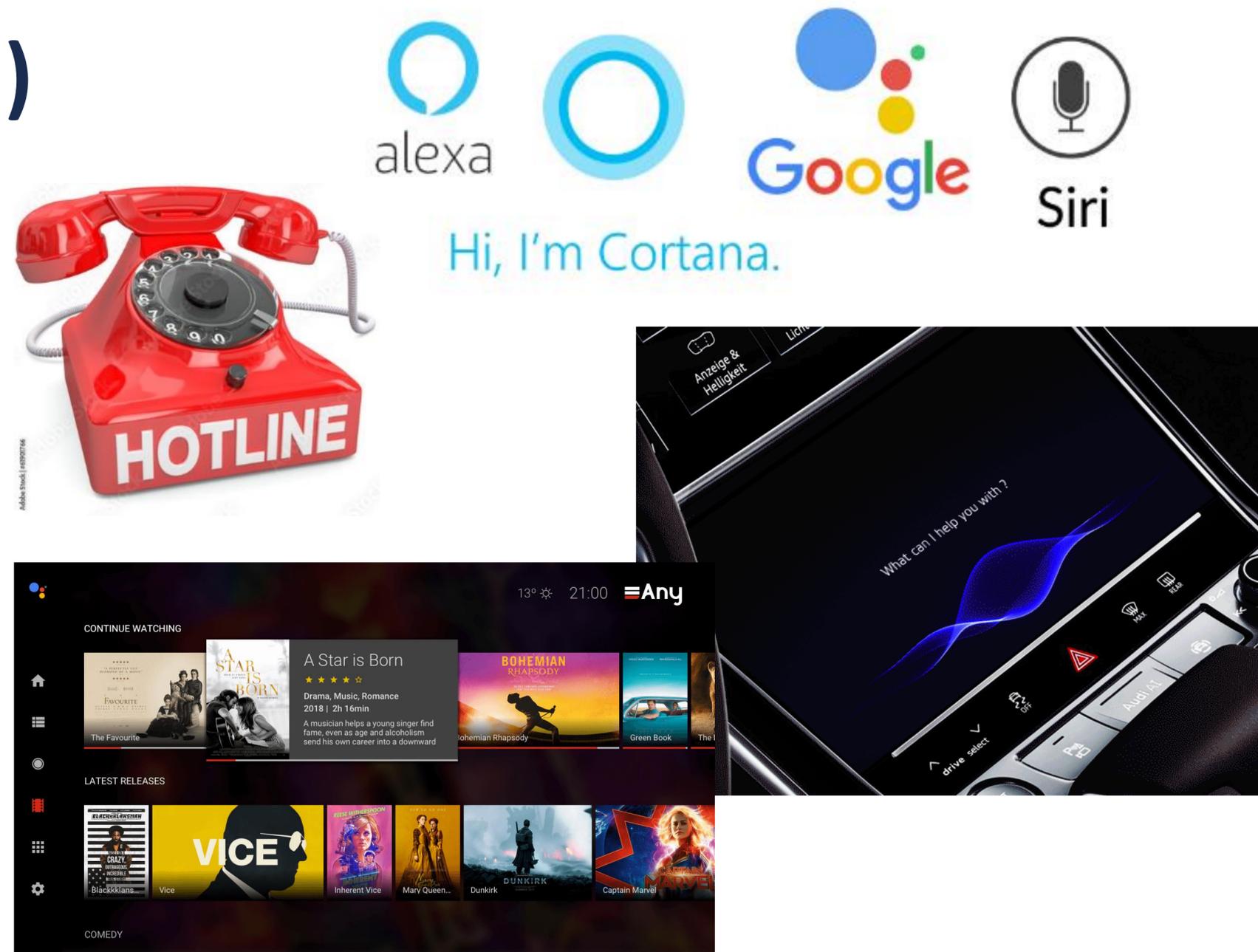


Voice User Interfaces (VUI)

Interaktion, d.h. Ein- (und Ausgabe) per gesprochener Sprache

- Natürlich und intuitiv
- Effizient und effektiv
- Berührungslos und hands-free
- Verbesserte Zugänglichkeit
- Erhöhte Sicherheit

Sprache: „Medium des Denkens und der Weltauffassung“



Short History of VUI

Frühe Anfänge (1950er bis 1960er Jahre):

- Grundlagen der Spracherkennungstechnologie
- Einfache Sprachbefehle (Command & Control)

Weiterentwicklung (1970er bis 1990er Jahre):

- Fortschritte in digitaler Signalverarbeitung und maschinellem Lernen
- Ersten kommerziellen ASR-Systeme für Telefonie/IVR und Automobil
- Meist Grammatik-ASR

Aufstieg der Sprachassistenten (2000er Jahre):

- Internet, Datenverfügbarkeit und zunehmenden Rechenleistung
- Sprachassistenten wie Siri (2011), Amazon Alexa (2014) und Google Assistant (2016)
- zunehmend LVCSR

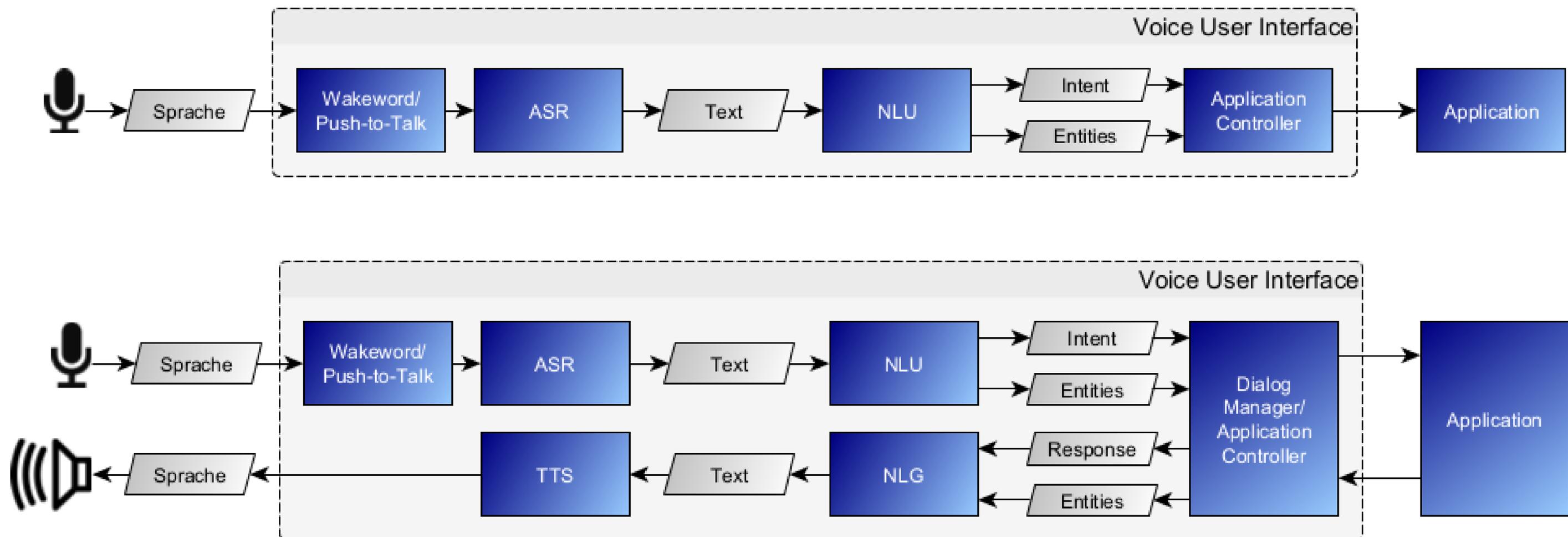
Verbreitung (heute):

- Integration in Smartphones, Smart-Home-Geräte, Automotive, ...
- LVCSR und Interaktionen per natürlicher Sprache

Intelligent?

Intuitivität/UX

The Setup of an „intelligent“ VUI



The Challenges of VUI

- **Wakeword/Push-to-talk:** Fehlauslösungen, Umgebungsgeräusche
- **ASR (Spracherkennung):** Variabilität der Aussprache, Fachsprache und Domäne, Umgebungsgeräusche, Sprecher-Mikrofon-Abstand, Raumhall)
- **NLU (Sprachverstehen):** Sprachvariation und Ambiguität, Kontextverständnis
- **NLG (Spracherzeugung):** Natürlichkeit der Sprache, Beherrschung von Sprache und Vokabular
- **TTS (Sprachsynthese):** Natürlichkeit der Stimme, Aussprache Fachbegriffe und Namen
- **DMS (Dialogmanagement):** Intuitiver und konsistenter Dialogfluss



Fazit

- VUIs in Consumer-Produkten zunehmende Popularität
- möglicher Nutzen in industriellen Umgebungen noch nicht ausgeschöpft
- VUI haben Potenzial Effizienz, Produktivität, Sicherheit und Zugänglichkeit zu verbessern
- Herausforderungen wie Geräuschinterferenzen, Fehlinterpretationen und Nutzerakzeptanz müssen gelöst werden

Our Products

Scalable Speech Recognition and Speech Processing



alphaspeech **pro**

Die professionelle
Spracherkennungs-Engine.



alphaspeech **μ**

Der kleinste
Spracherkenner der Welt.



alphaspeech **NLU**

Extrahiere Informationen
aus Spracheingaben.



alphaspeech **HMI**

Gerätesteuerung,
Dialogsysteme und
Assistenten.



alphaspeech **consult**

Wir finden Lösungen.
Versprochen.

alphaspeech **pro**

The professional speech recognition engine

- Grammar Speech Recognition OR Large Vocabulary Continuous Speech Recognition (LVCSR)
- **Grammar development** and optimization for **professional vocabulary** and **operating conditions**
- Speaker-independent
- Real-time (**On-the-Fly**)
- Provided as **SaaS, On-Premises** or as **Software Library/SDK**

Text-to-Speech

Transcription

Voice User Interfaces

Dialog Systems

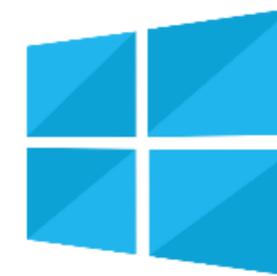


Custom API and interface development on demand

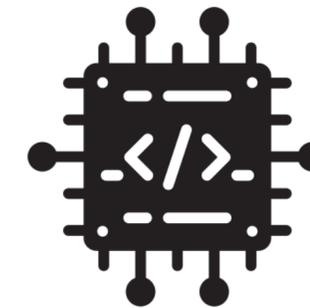
alphaspeech **SDK**

Our ASR Library in your application.

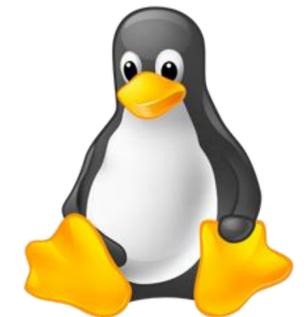
- Easy development and integration with **Software Development Kit (SDK)**
- Comprehensible **documentation** and **demo code** for a quick start
- For **Microsoft Windows, Linux** and **embedded systems**
- APIs for **C++** and **Python**
- Application-specific **ASR models** (acoustic, language)
- Signal preprocessing / denoising



*Microsoft
Windows*



*Embedded
(z. B. ARM)*



Linux



APIs for C++ and Python

alphaspeech μ

The world's smallest speech recognition system

- Continuous ASR of **words, phrases or sentences**
- Speaker- and platform-independent
- Optimized in terms of **RAM memory** and **power consumption**
- Available languages: English, German, Japanese (others on demand)



<https://linguino.de>

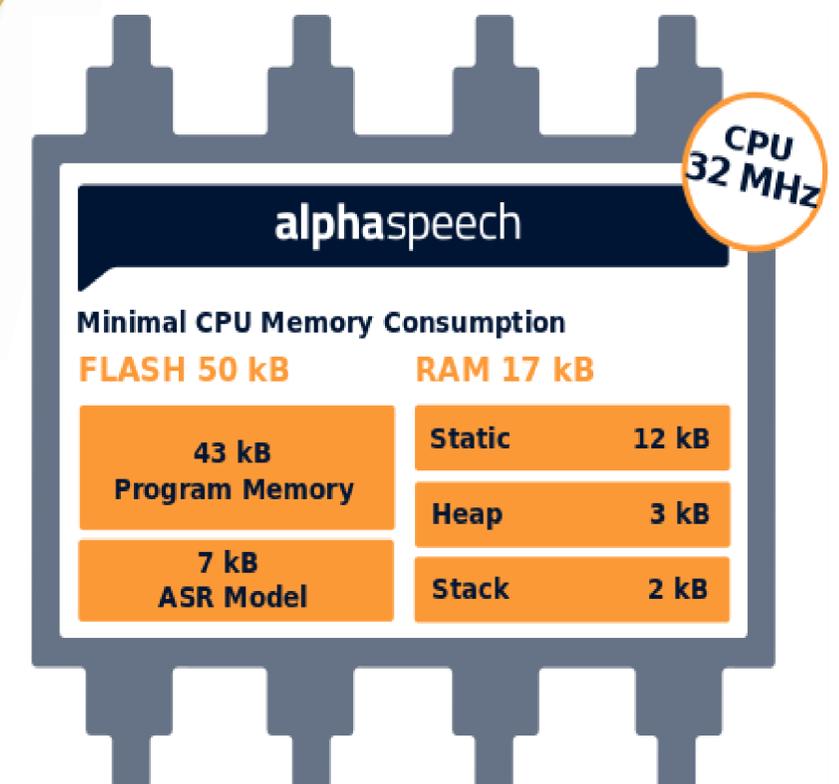
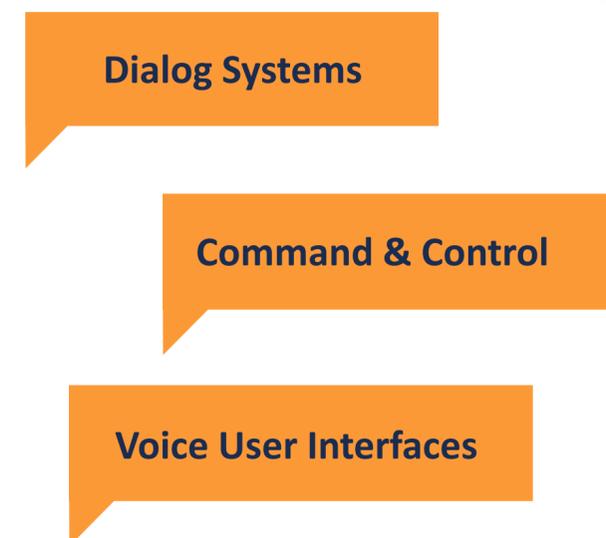


Figure 1: minimal CPU memory consumption of **alphaspeech μ** .

alphaspeech **Speech-to-Form**

Extract information from free speech input

- Natural Language Understanding (**NLU**)
 - **intent recognition**
 - **Information extraction (slot filling)**
- Quick development of application-specific **NLU** and **ASR** models
- Efficient and accurate due to **optimal interaction** of ASR and NLU
- **Easy integration** with standard APIs

Voice User Interfaces

Data Entry and Form Filling

The screenshot shows a medical form titled 'Wundprotokoll Nr. 2 für Patient Meier, Michael, 12.01.1988'. The form is displayed on a laptop screen. At the top, there are fields for 'Patient' (Meier, Michael [12.01.1988]), 'Erfasst durch' (Müller, Fritz), 'Datum' (01.03.2022), and 'Wunde Nr.' (2). A 'Protokoll exportieren' button is visible. The main form is divided into several sections:

- Wundgröße:** Fields for 'Länge (cm)' (2,5), 'Breite (cm)' (1,0), and 'Tiefe (cm)' (25). Below these are radio buttons for 'Methodik': 'Perpendicular' (selected) and 'Längsachse zum Körper'.
- Wundgrund:** Radio buttons for 'Nekrose' (keine, ja), 'Epithelinseln' (ja, nein), 'trocken', 'feucht', and a dropdown for 'Granulation (Gewebe Neubildung)'.
- Anmerkungen:** A text area for notes.
- Wundumgebung:** Checkboxes for 'unauffällig/Haut intakt', 'gerötet', 'geschwollen', 'verhärtet', 'mazeriert (aufgeweicht/zersetzt)', and 'schmerzhaft'.

At the bottom of the form, a summary line reads: 'Länge der Wunde 2,5 cm, Breite 1 cm, Tiefe 25 mm. Perpendicularmethode.' A microphone icon is located at the bottom center of the screen.

alphaspeech **Voicebot**

Customized free speech dialogs

- Natural Language Understanding (**NLU**)
 - **intent recognition**
 - **Information extraction (slot filling)**
- Quick development of application-specific **Dialogs, NLU** and **ASR models**
- Efficient and accurate due to **optimal interaction** of ASR and NLU
- Modular composition of **ASR, TTS** and **DMS**
- **Easy integration** with standard APIs

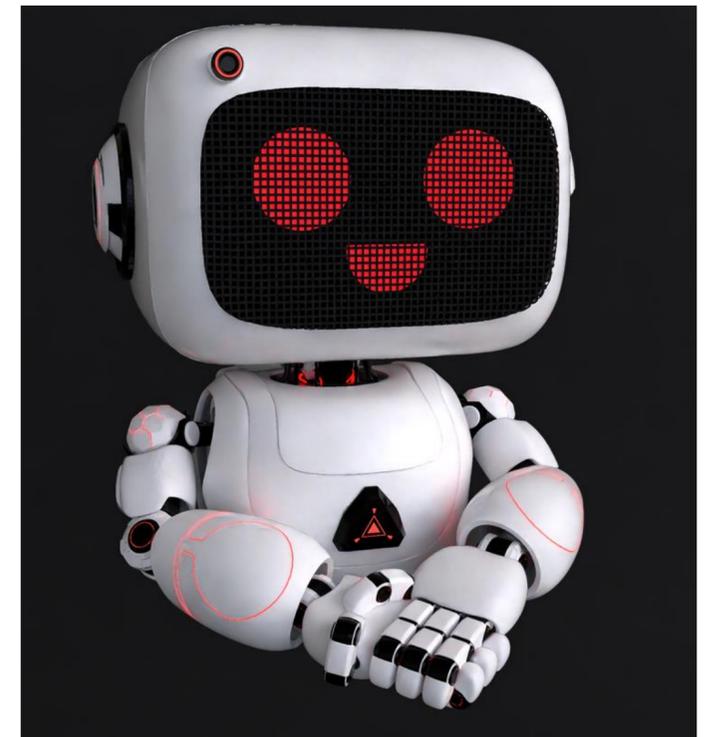
Dialog Systems

Voice User Interfaces

Voicebots

Chatbots

Phonebots

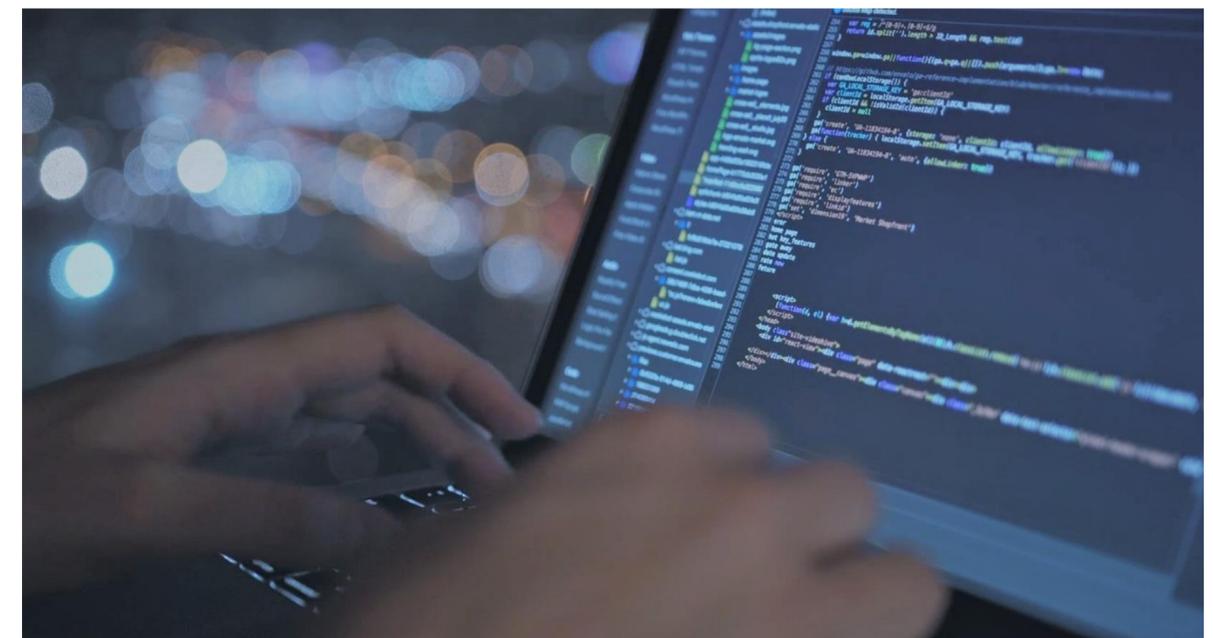


alphaspeech **consult**

We find solutions. We promise.

20 years experience in speech recognition and processing

- **Consulting and development services**
- **Optimization and adaptation of ASR systems**
- **Custom APIs, wrapper or bindings**
- **Performance analyses**
- **Software development, e.g. ASR integration**
- **Data preparation, e.g. cleaning, normalization, annotation**

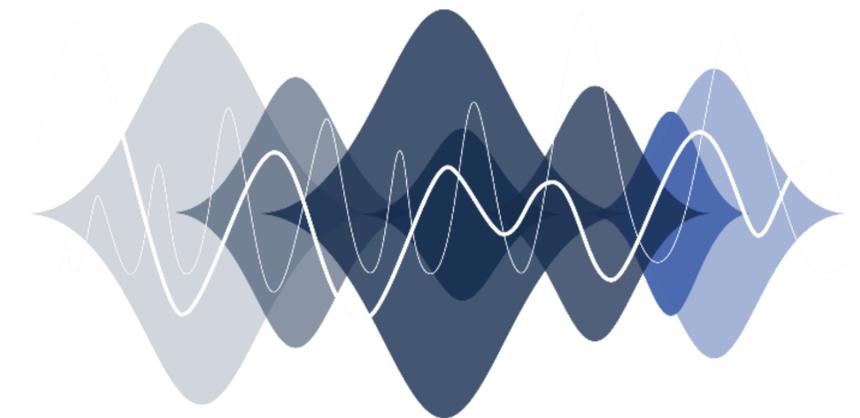


Closing Statement

Vision and Mission

Speech is the most natural form of human communication.

By recognizing and interpreting spoken language, we want to contribute to a revolution of machine and computer interfaces.



alphaspeech

Spracherkennung. Made in Germany.

Noch Fragen zu alphaspeech?

Sprich mit uns!



+49 351 6533 3808

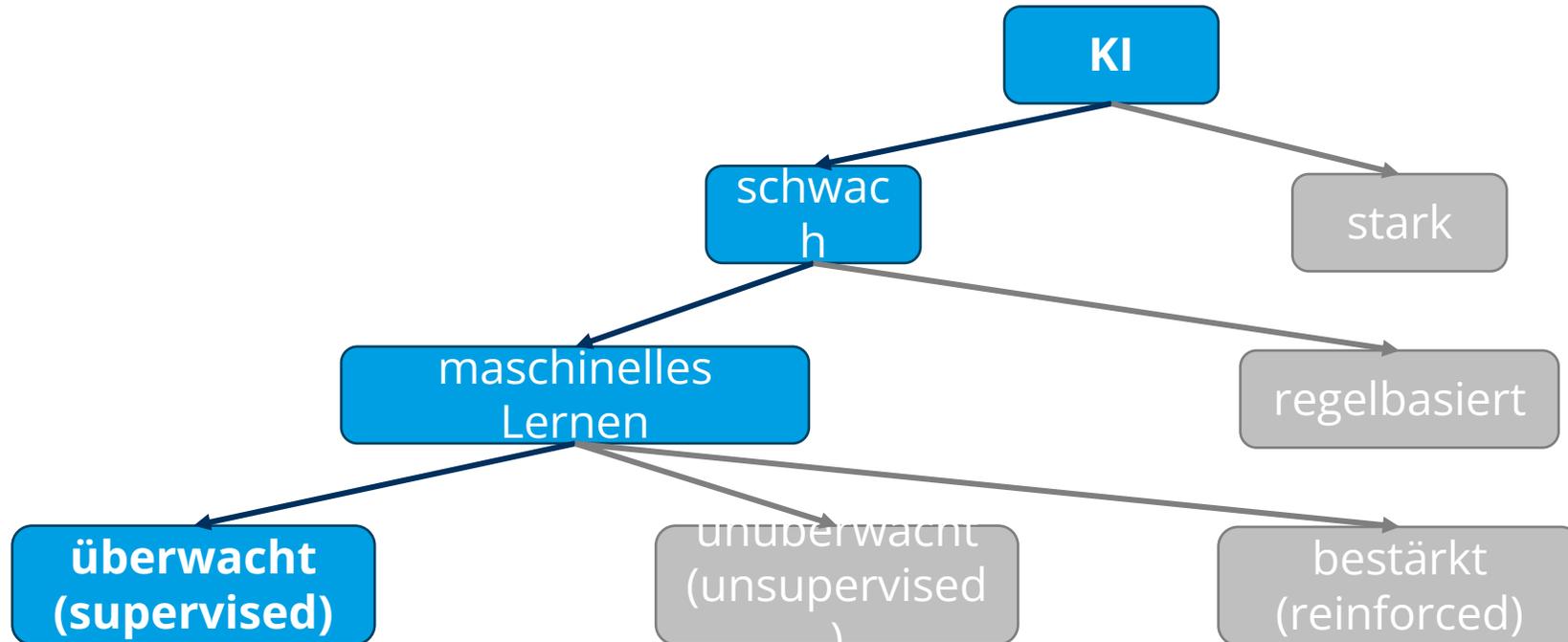


sprich-mit-uns@alphaspeech.de

Edgar Scherstjanoi
Labelling - Das Menschliche in der künstlichen Intelligenz

Überwachtes maschinelles Lernen

Einordnung in „künstlicher Intelligenz“



Überwachtes maschinelles Lernen

Supervised Learning

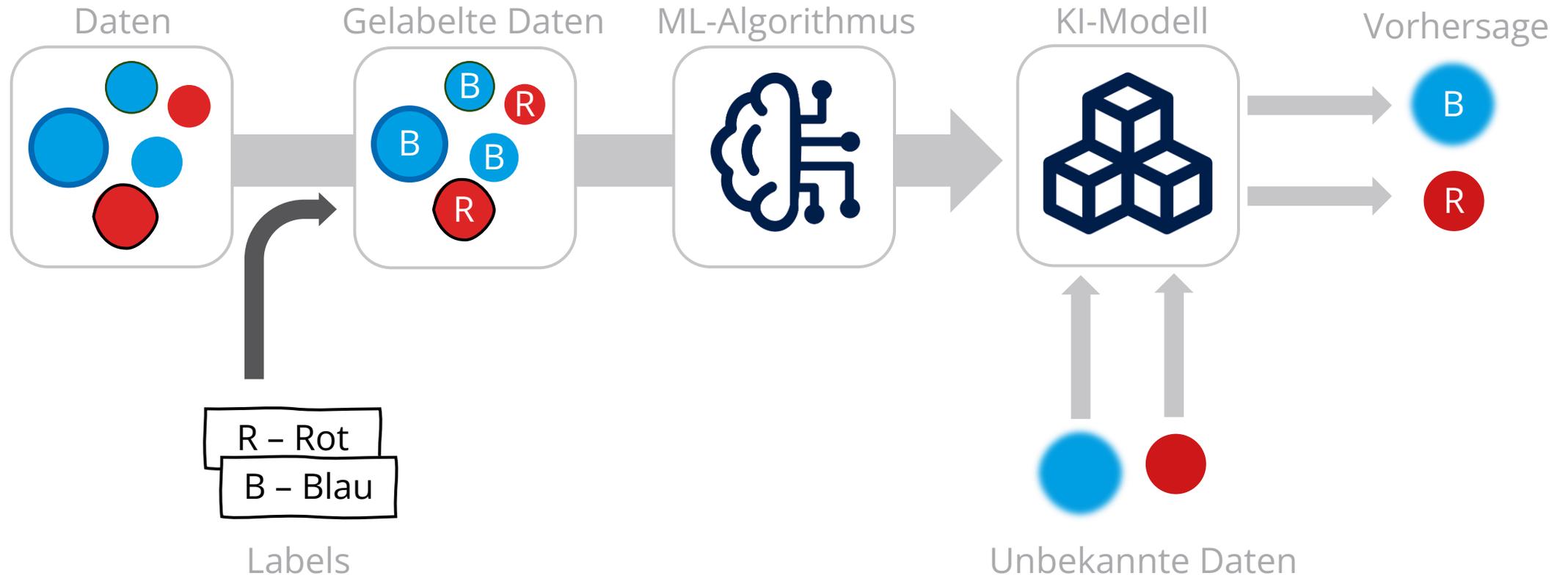


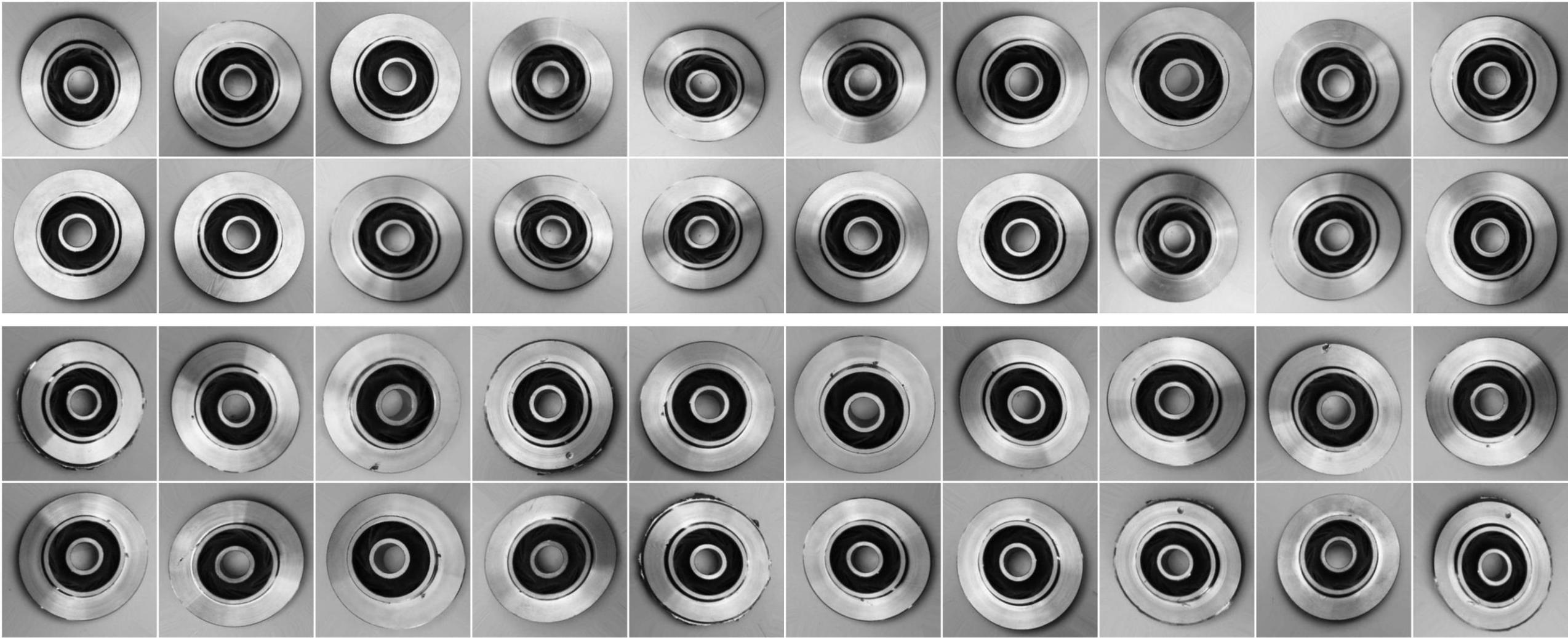
Image Classification

MNIST Erkennung handschriftlicher Ziffern



Image Classification

Erkennung von Gussfehlern



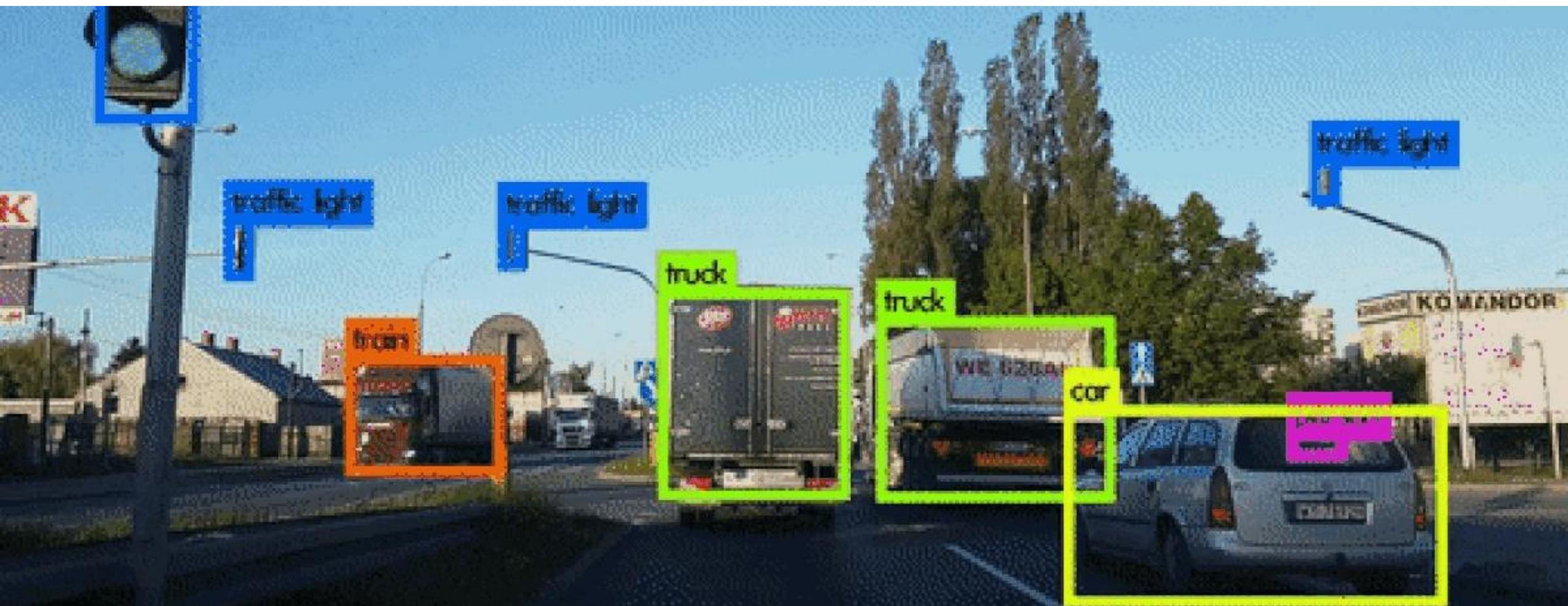
Object Detection

Datensatz: COCO – Common Objects in Context



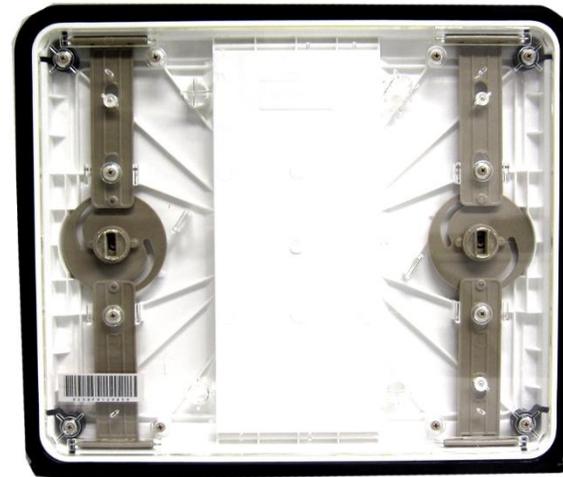
Object Detection

Algorithmus: You Only Look Once (YOLO)



ResponsiveFab YOLO

Detektion von Transportbehältern



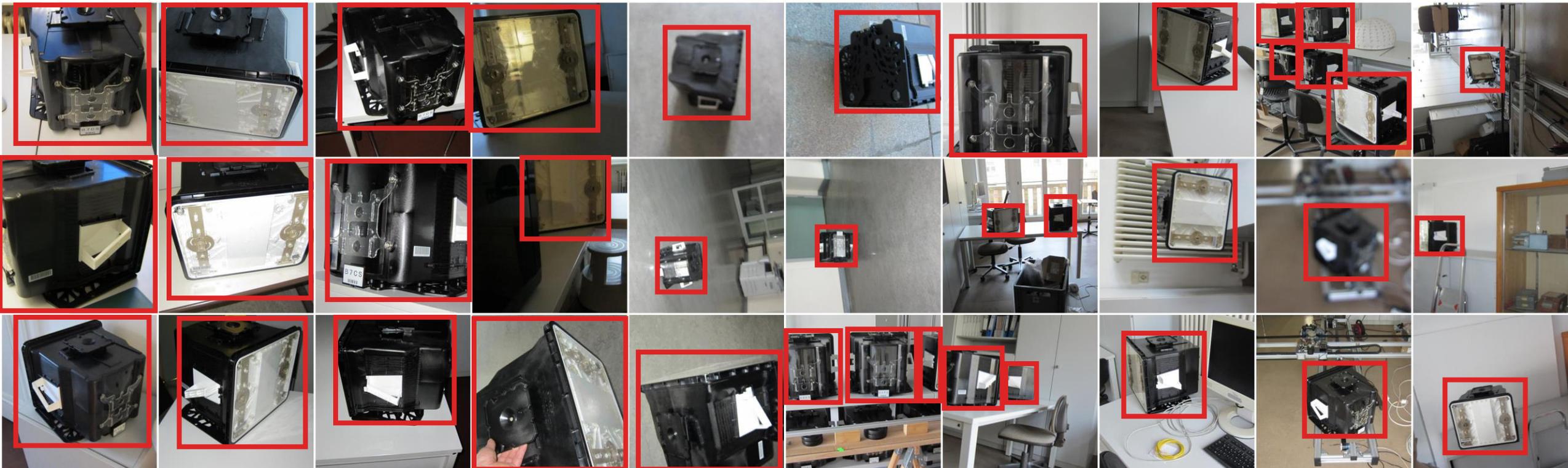
ResponsiveFab YOLO

Detektion von Transportbehältern



ResponsiveFab YOLO

Detektion von Transportbehältern



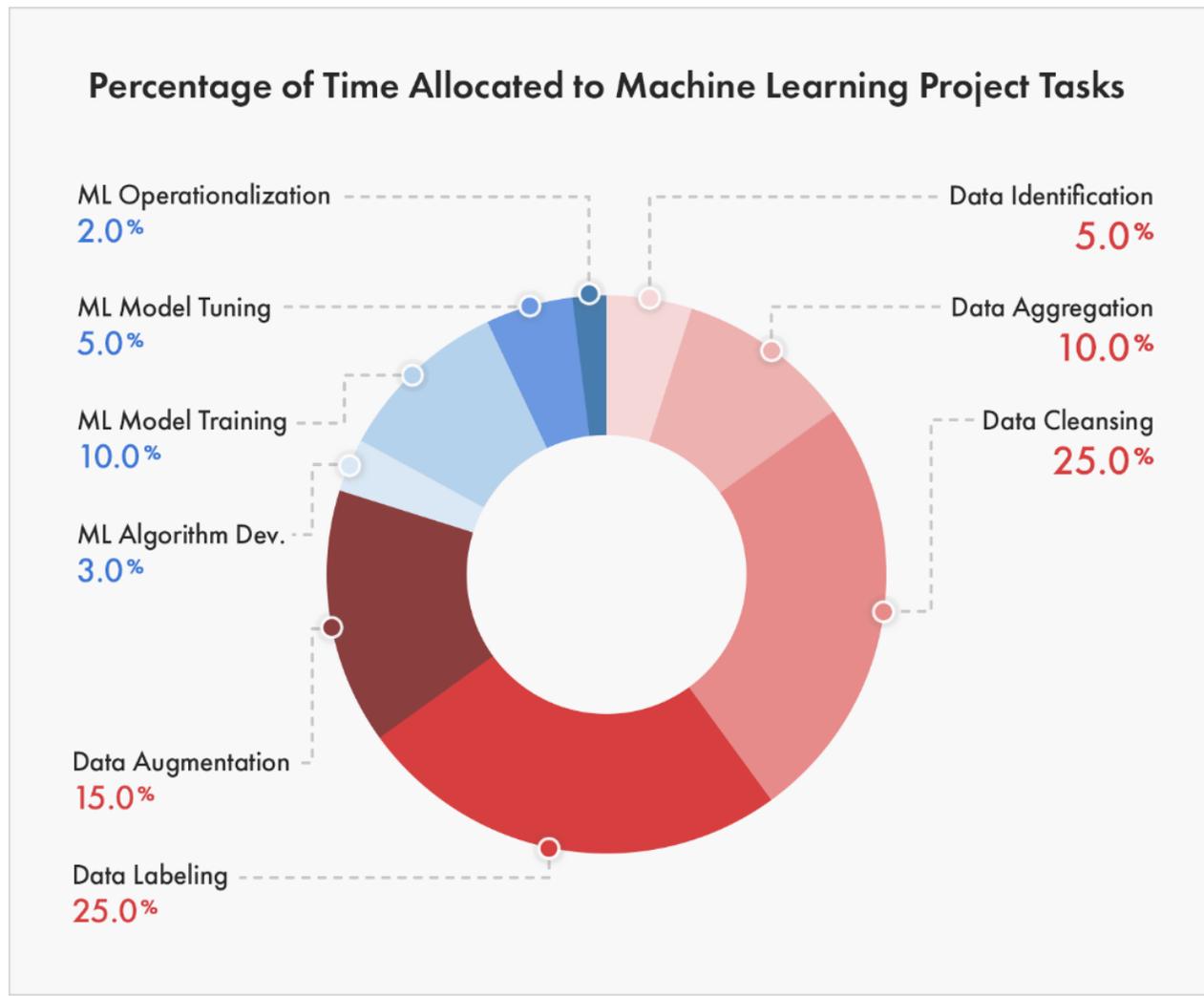
ResponsiveFab YOLO

Detektion von Transportbehältern



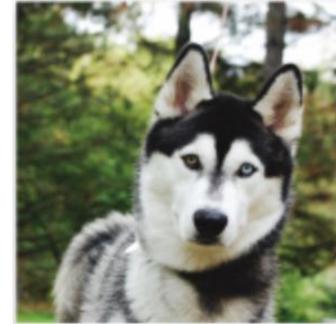
Was ist das Problem?

Zeitaufwand



Was ist das Problem?

Fehleranfälligkeit



Predicted: Husky
True: Husky



Predicted: Husky
True: Husky



Predicted: Wolf
True: Wolf



Predicted: Husky
True: Wolf



Predicted: Wolf
True: Wolf



Predicted: Wolf
True: Husky

Was ist das Problem?

Tätigkeit

- Monotone Wiederholungen
- Aktualität/ Instandhaltungen

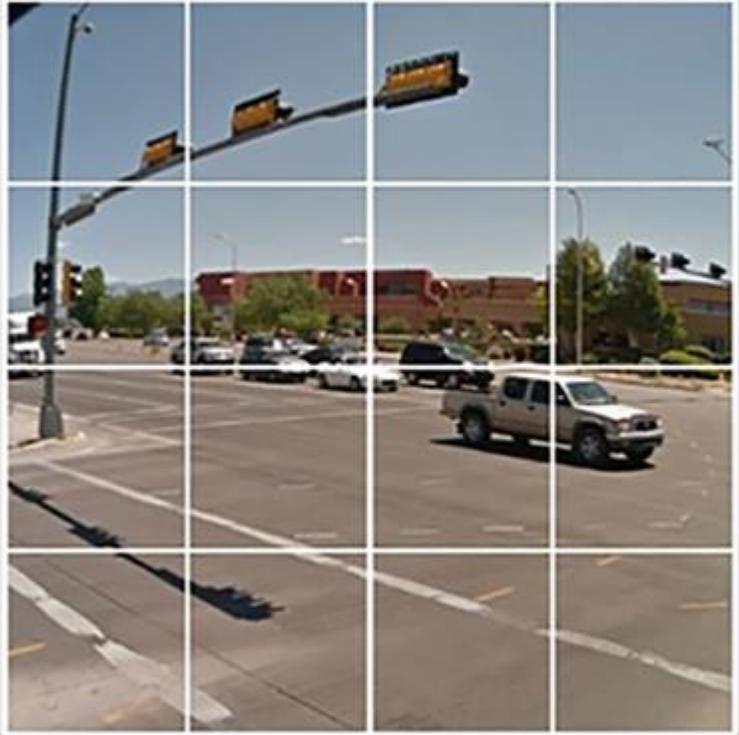
Vorraussetzung: Domänenspezifisches Wissen

- Grundlagen maschinelles Lernen
 - Verständnis zur produktionstechnischen Problematik
- Abschätzung Aufwand/ Nutzen

Wer?

- Angestellte
- Spezielle Abteilung
- Extern (Outsourcing)
- Crowdsourcing

Select all the sections with traffic lights

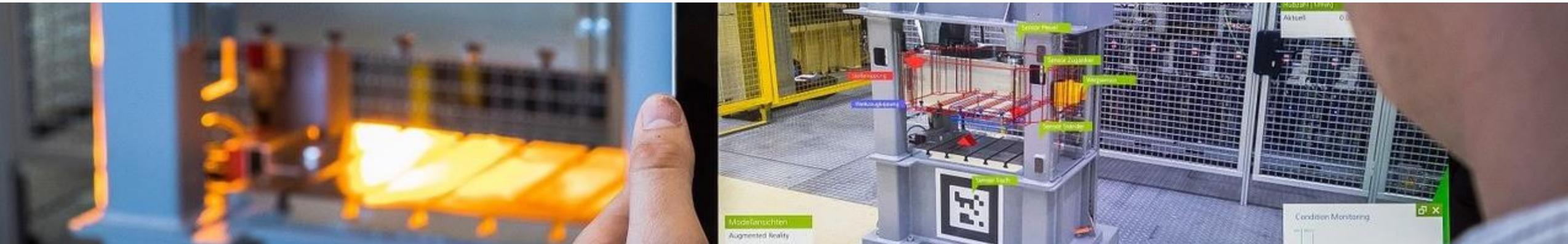


Submit Your Response

Submit

Zusammenfassung

- Datenqualität → Modellqualität
- Labeller/ Annotator → neue Tätigkeitsform
- Kein Masterplan → Individuelle Lösungen
- Angebot ProKI-Workshops
→ <https://tu-dresden.de/ing/proki/unsere-angebote/workshop>

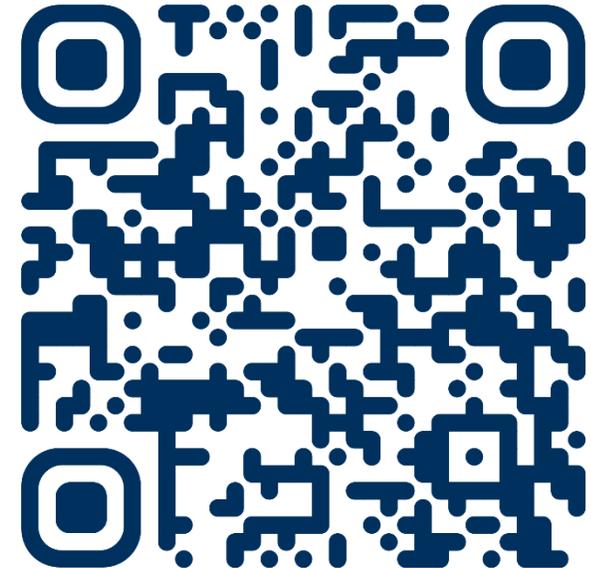


Quellen

- [1] EMNIST Dataset. <https://www.nist.gov/itl/products-and-services/emnist-dataset>
- [2] Casting Product. Dataset <https://www.kaggle.com/datasets/ravirajsinh45/real-life-industrial-dataset-of-casting-product/data>
- [3] COCO Dataset. <https://cocodataset.org/>
- [4] You Only Look Once. <https://pjreddie.com/darknet/yolo/>
- [5] ResponsiveFab YOLO. Scherstjanoi et al. (2019). Automatic FOUP Recognition for Autonomous Transport Robots. 19th European Advanced Process Control and Manufacturing Conference (apc | m) at Villach, Austria.
- [6] Zeitwaufwand. <https://dzone.com/articles/an-introduction-to-data-labeling-in-artificial-int>
- [7] Husky vs. Wolf. <https://carpentries-incubator.github.io/data-science-ai-senior-researchers/05-Problems-with-AI/index.html>
- [8] ReCaptcha. Moradi, M. and Keyvanpour, M.R. (2022), "CAPTCHA for crowdsourced image annotation: directions and efficiency analysis", Aslib Journal of Information Management, Vol. 74 No. 3, pp. 522-548. <https://doi.org/10.1108/AJIM-08-2021-0215>
- [9] Active Learning. Monarch, R. M. (2021). Human-in-the-Loop Machine Learning: Active learning and annotation for human-centered AI. Simon and Schuster.
- [10] Semi-Supervised Learning. <https://www.ibm.com/topics/semi-supervised-learning>
- [11] InstructGPT. <https://arxiv.org/abs/2203.02155>

Feedback

Wie hat Ihnen der 16. ProKI-InfoPoint gefallen?



17. ProKI-InfoPoint

11. April 2024 16.00Uhr

Welche Rolle spielt Künstliche Intelligenz bei der Gewährleistung und Verbesserung der Produktqualität?



Vielen Dank für die Aufmerksamkeit!

<https://tu-dresden.de/ing/proki>

<https://proki-darmstadt.de/>

<https://proki-netz.de/>



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

ANSPRECHPARTNER

Technische Universität Dresden

Dr. Hajo Wiemer

Telefon: +49 (0) 351 463 32004

E-Mail: hajo.wiemer@tu-dresden.de



ANSPRECHPARTNER

Technische Universität Darmstadt

Sven Varchmin

Telefon: +49 (0) 6151 16 23148

E-Mail: sven.varchmin@ptu.tu-darmstadt.de

