



ZAKKI

Zentrale Anlaufstelle für innovatives Lehren und Lernen interdisziplinärer Kompetenzen der KI

Anwendungsorientierte Forschung zu Künstlicher Intelligenz

Prof. Dr.-Ing. Sebastian von Enzberg
sebastian.von.enzberg@h2.de

h2-Science-Day
Magdeburg, 20.11.2024

KI-Sprachmodelle in der Forschung?

EFFICIENT TRANSFER LEARNING IN LOW-DIMENSIONAL DIFFUSION MODELS: A FINE-TUNING APPROACH

Anonymous authors
Paper under double-blind review

ABSTRACT

This paper addresses the challenge of efficiently adapting low-dimensional diffusion models to new datasets with limited data through transfer learning. Diffusion models, while powerful for generative tasks, typically require extensive training on large datasets, making them impractical for scenarios with data or computational constraints. We propose a novel transfer learning framework that fine-tunes a pre-trained diffusion model on small amounts of target data, significantly reducing training time while maintaining competitive performance. Our approach involves training a base MLPDenoiser model on multiple source datasets and systematically fine-tuning it on varying amounts of a held-out target dataset. Experiments demonstrate substantial improvements in both computational efficiency and model performance. Fine-tuning on just 10% of the target dataset reduced the KL divergence by 45% and training time by 57% compared to the base model. Our results, supported by KL divergence comparisons and generated sample visualizations, show that transfer learning can greatly enhance the efficiency and adaptability of diffusion models in low-dimensional spaces. This work opens up new possibilities for applying diffusion models in resource-constrained environments and rapid prototyping scenarios, potentially accelerating research in fields such as drug discovery and financial modeling where low-dimensional data is prevalent.

1 INTRODUCTION

Diffusion models have emerged as a powerful class of generative models, demonstrating remarkable success in various domains such as image synthesis, audio generation, and molecular design Yang et al. (2023). However, their application to low-dimensional data and their adaptability to new datasets with limited samples remain underexplored. This paper investigates the potential of transfer learning techniques to enhance the efficiency and adaptability of diffusion models in low-dimensional spaces, addressing a critical gap in the current research landscape.

The application of diffusion models to low-dimensional data presents unique challenges. Unlike high-dimensional data such as images, low-dimensional datasets often lack the complex hierarchical

3.1 PROBLEM SETTING

Let $\mathcal{D}_s = \{(x_i, y_i)\}_{i=1}^{N_s}$ denote a source dataset consisting of N_s samples, where $x_i \in \mathbb{R}^d$ represents a d -dimensional data point and y_i its corresponding label. We consider a pre-trained diffusion model M_s that has been trained on \mathcal{D}_s . Our goal is to adapt this model to a target dataset $\mathcal{D}_t = \{(x_j, y_j)\}_{j=1}^{N_t}$, where $N_t \ll N_s$, i.e., the target dataset is significantly smaller than the source dataset.

The transfer learning process involves fine-tuning the pre-trained model M_s on the target dataset \mathcal{D}_t to obtain an adapted model M_t . We denote the fine-tuning procedure as $\mathcal{F}: (M_s, \mathcal{D}_t) \rightarrow M_t$. Our objective is to optimize \mathcal{F} such that M_t performs well on the target dataset while requiring minimal fine-tuning data and computational resources.

Key assumptions and challenges in our problem setting include:

- The source and target datasets are low-dimensional (typically $d \leq 3$), which presents unique challenges compared to high-dimensional data like images.
- The target dataset \mathcal{D}_t may have different statistical properties or underlying distributions compared to the source dataset \mathcal{D}_s .
- We aim to achieve good performance in scenarios where only a small fraction of the target dataset is available for fine-tuning.

By addressing these challenges, our work opens up new possibilities for applying diffusion models to new datasets while maintaining limited data and computational resources.

Our experiments demonstrate the effectiveness of our approach. Fine-tuning on just 10% of the target dataset reduced the KL divergence by 45% and training time by 57% compared to the base model. Our results, supported by KL divergence comparisons and generated sample visualizations, highlight the potential of transfer learning in diffusion models to new datasets while maintaining limited data and computational resources.

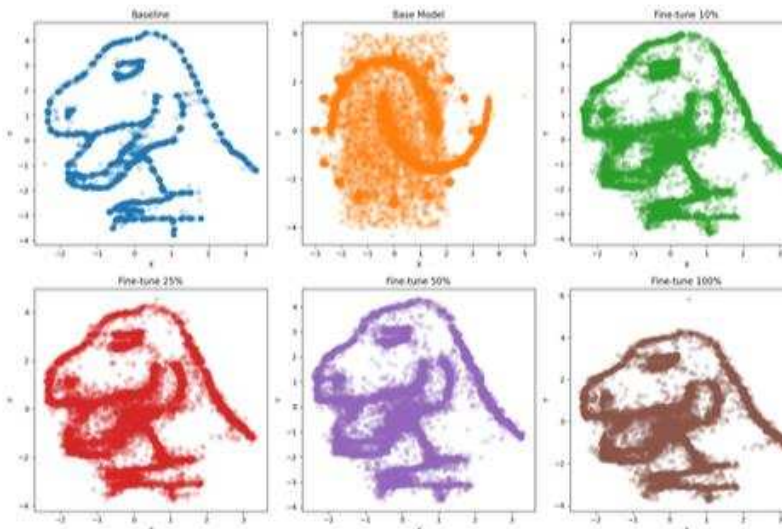


Figure 4: Generated samples for the Dino dataset from different experimental runs, showing the improvement in sample quality as we increase the amount of fine-tuning data.

REFERENCES

Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger (eds.), *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL <https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf>.

Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT Press, 2016.

Ho, Jonathan, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851. Curran Associates, Inc., 2020. URL <https://proceedings.neurips.cc/paper/2020/file/ca4b-Paper.pdf>.

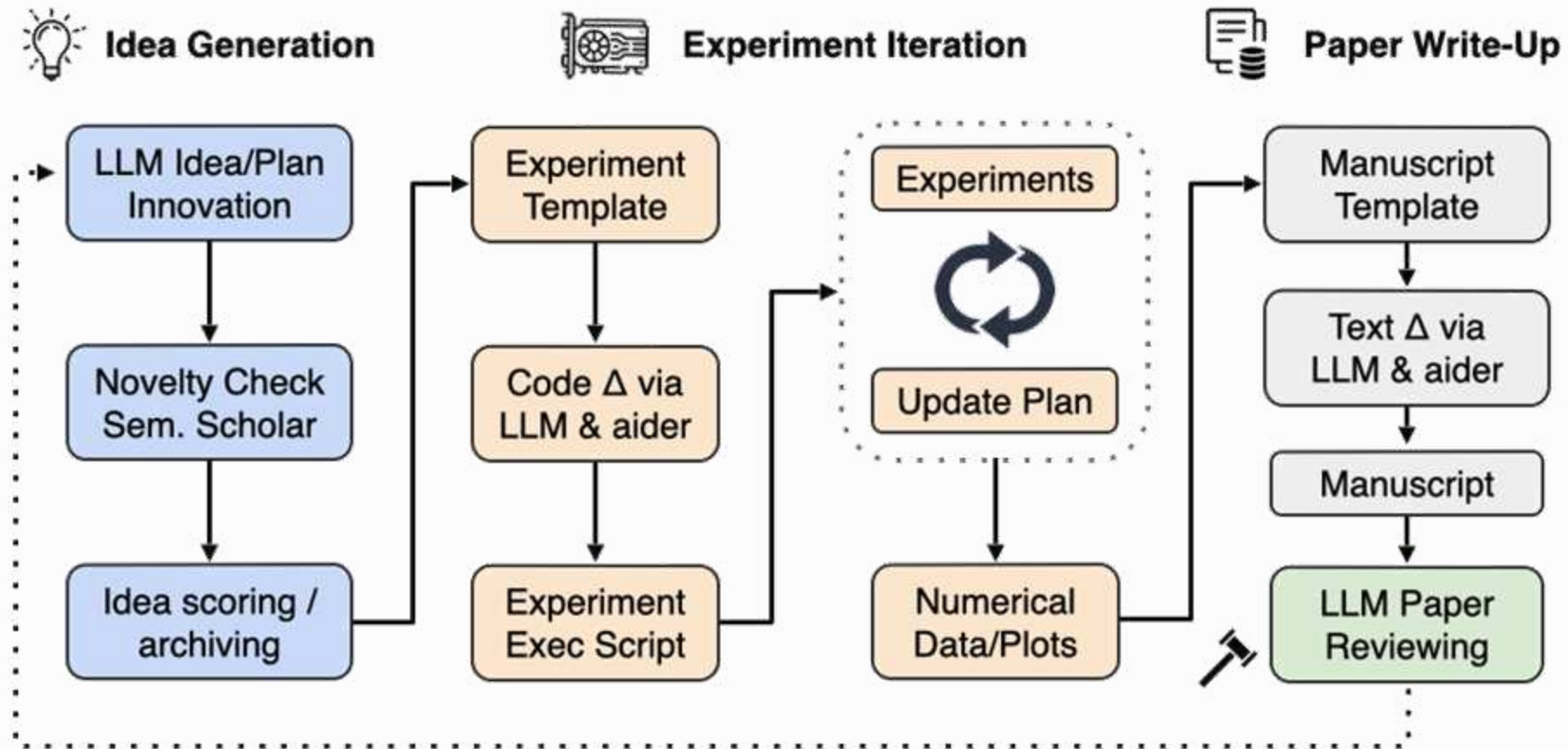
Oh, Alice H., Alekh Agarwal, Danielle Belgrave, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Neural Information Processing Systems*, 2022. URL <https://openreview.net/forum?id=k7FuTOWMoc7>.

Kingma, Diederik P. and Max Welling. Auto-Encoding Variational Bayes. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014*.

Rubachev, and Artem Babenko. Tabddpm: Modelling

Diffusion Models from Limited Labeled Data (11d) workshop:

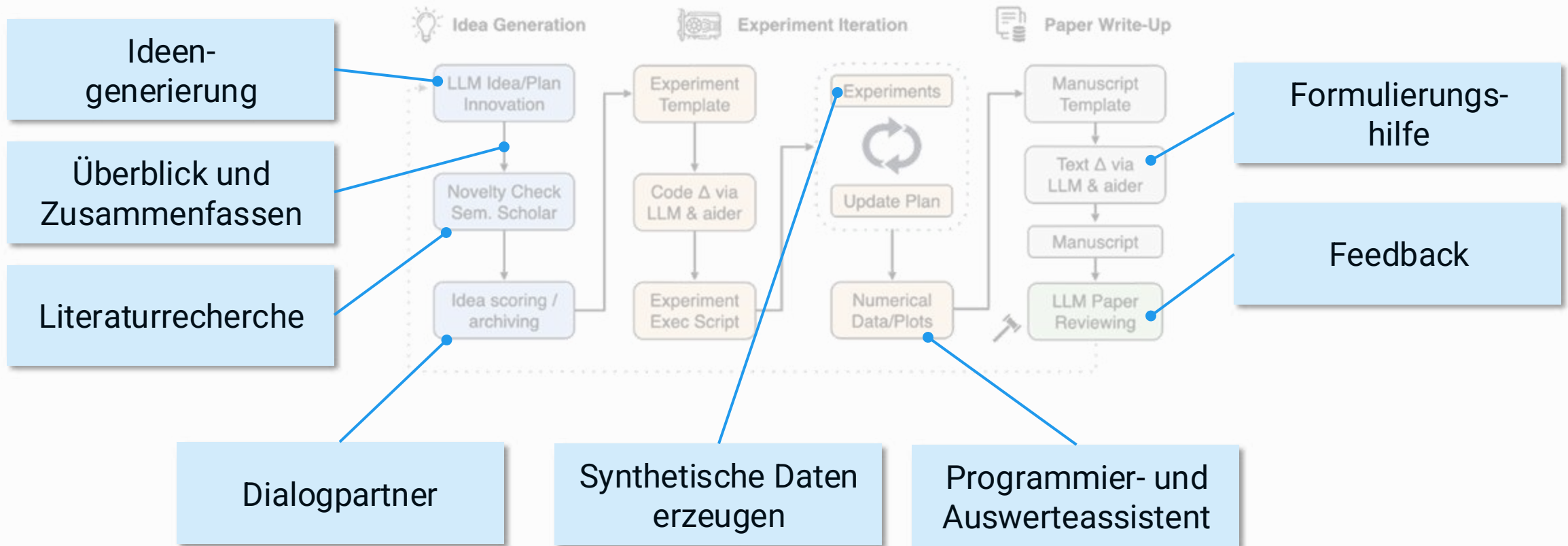
KI-Sprachmodelle in der Forschung?



Lu, Chris, et al. "The AI scientist: Towards fully automated open-ended scientific discovery." *arXiv:2408.06292*, 01.07.2024

Bildquelle: <https://github.com/SakanaAI/AI-Scientist>

KI-Sprachmodelle in der Forschung?



Lu, Chris, et al. "The AI scientist: Towards fully automated open-ended scientific discovery." *arXiv:2408.06292*, 01.07.2024

Bildquelle: <https://github.com/SakanaAI/AI-Scientist>

KI-Sprachmodelle in der Forschung?

Oliver picks 44 kiwis on Friday.
Then he picks 58 kiwis on Saturday.
On Sunday, he picks double the number of kiwis he did on Friday, but *five of them were a bit smaller* than average.

How many kiwis does Oliver have?

$$44 + 58 + 2 \cdot 44 = 190$$

GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models

Iman Mirzadeh[†] Keivan Alizadeh Hooman Shahrokhi*
Oncel Tuzel Samy Bengio Mehrdad Farajtabar[†]

Apple

Abstract

Recent advancements in Large Language Models (LLMs) have sparked interest in their formal reasoning capabilities, particularly in mathematics. The GSM8K benchmark is widely used to assess the mathematical reasoning of models on grade-school-level questions. While the performance of LLMs on GSM8K has significantly improved in recent years, it remains unclear whether their mathematical reasoning capabilities have genuinely advanced, raising questions about the reliability of the reported metrics. To address these concerns, we conduct a large-scale study on several state-of-the-art open and closed models. To overcome the limitations of existing evaluations, we introduce GSM-Symbolic, an improved benchmark created from symbolic templates that allow for the generation of a diverse set of questions. GSM-Symbolic enables more controllable evaluations, providing key insights and more reliable metrics for measuring the reasoning capabilities of models. Our findings reveal that LLMs exhibit noticeable variance when responding to different instantiations of the same question. Specifically, the performance of all models declines when only the numerical values in the question are altered in the GSM-Symbolic benchmark. Furthermore, we investigate the fragility of mathematical reasoning in these models and demonstrate that their performance significantly deteriorates as the number of clauses in a question increases. We hypothesize that this decline is due to the fact that current LLMs are not capable of genuine logical reasoning; instead, they attempt to replicate the reasoning steps observed in their training data. When we add a single clause that appears relevant to the question, we observe significant performance drops (up to 65%) across all state-of-the-art models, even though the added clause does not contribute to the reasoning chain needed to reach the final answer. Overall, our work provides a more nuanced understanding of LLMs' capabilities and limitations in mathematical reasoning.

Mirzadeh, Iman, et al. "GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models." *arXiv:2410.05229*. 07.10.2024

KI-Sprachmodelle in der Forschung?

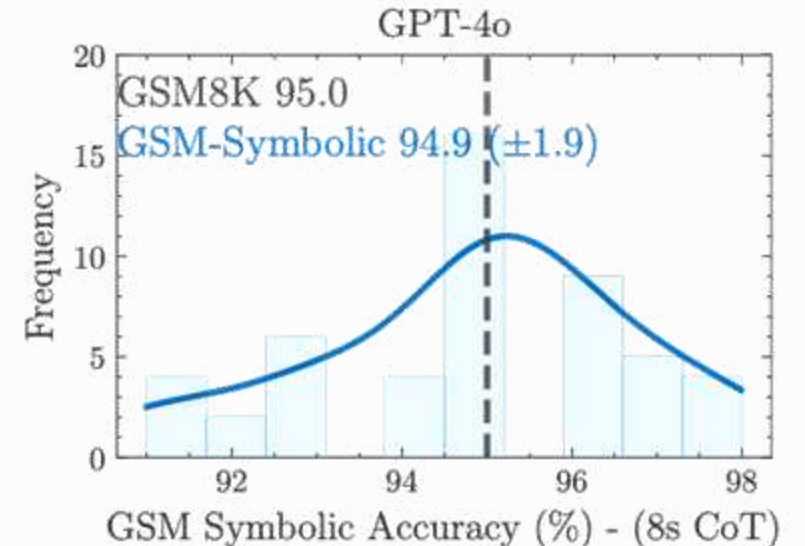
Oliver picks 44 kiwis on Friday.
Then he picks 58 kiwis on Saturday.
On Sunday, he picks double the number of kiwis he did on Friday, but *five of them were a bit smaller* than average.

How many kiwis does Oliver have?

o1-mini: Oliver picks kiwis over three days as follows:

- Friday: 44 kiwis
- Saturday: 58 kiwis
- Sunday: Double the number he picked on Friday, which is $2 \times 44 = 88$ kiwis

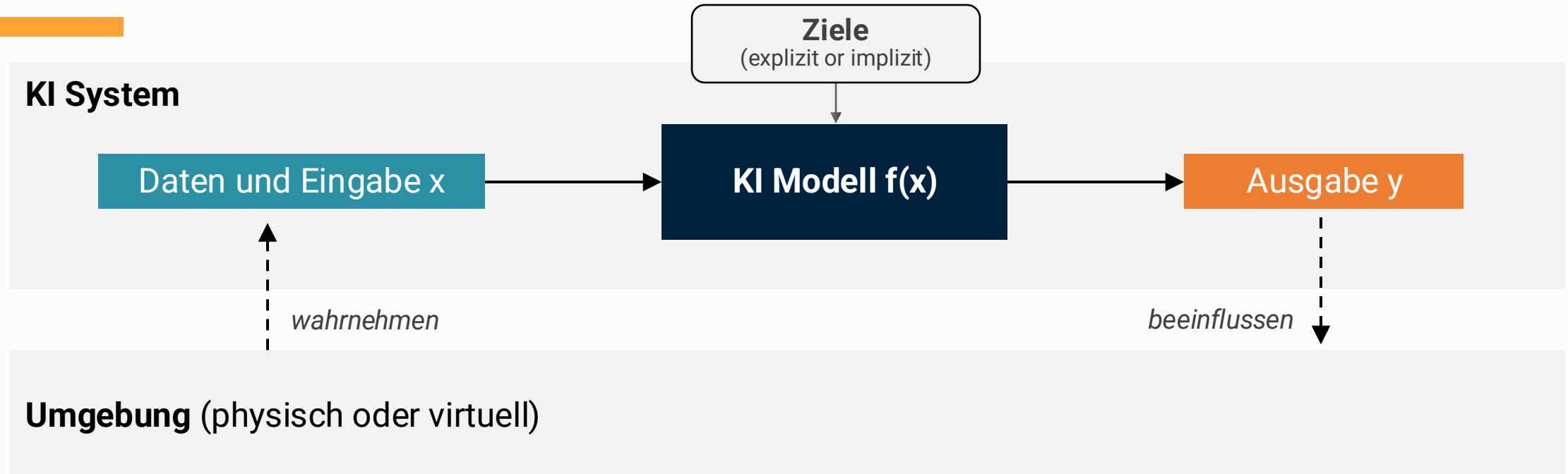
However, on Sunday, 5 of these kiwis were smaller than average. We need to subtract them from the Sunday total: 88 (Sunday's kiwis) - 5 (smaller kiwis) = 83 kiwis. Now, summing up the kiwis from all three days: 44 (Friday) + 58 (Saturday) + 83 (Sunday) = 185 kiwis. Oliver has a total of 185 kiwis.





Was ist eigentlich künstliche Intelligenz?

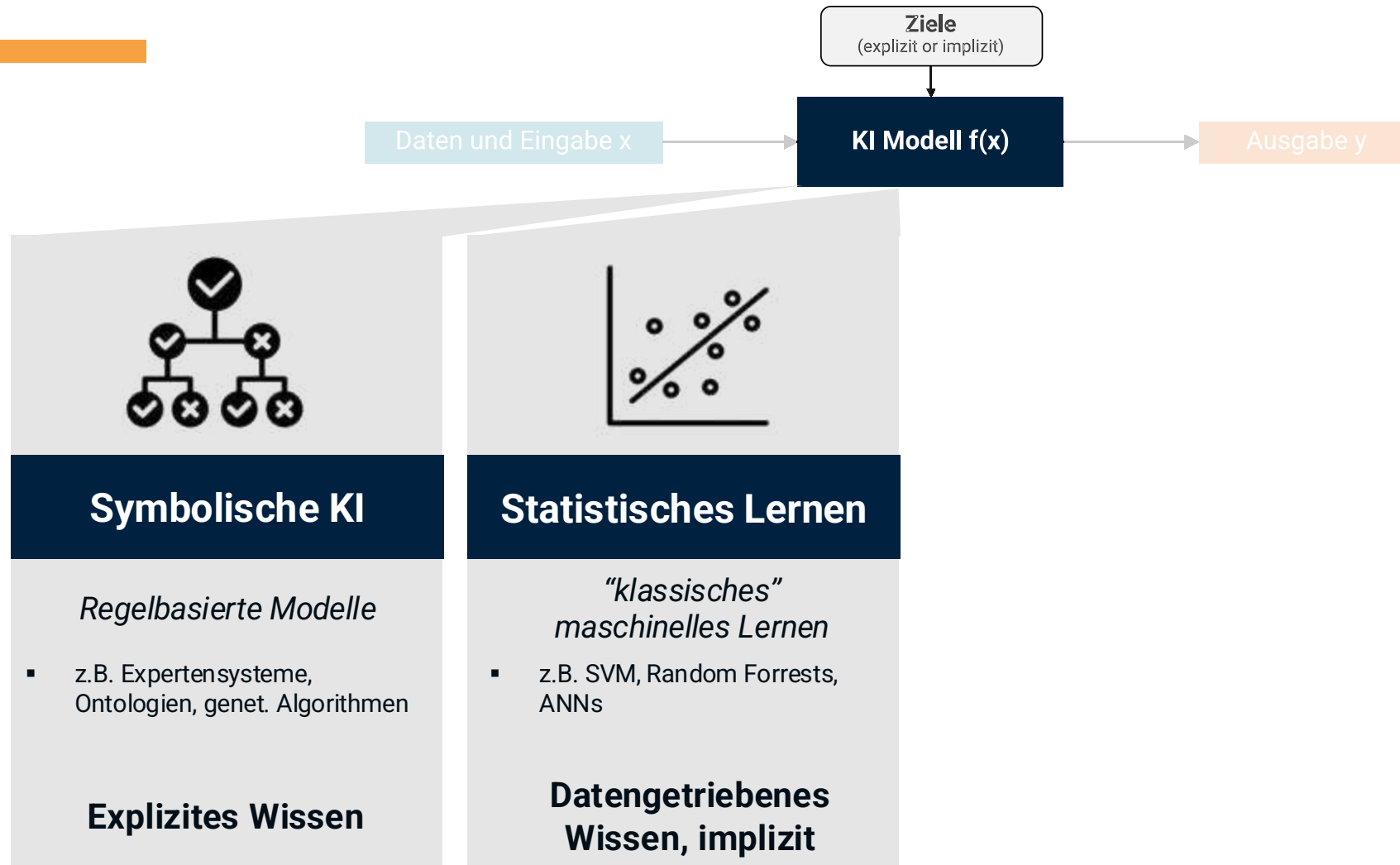
Was ist Künstliche Intelligenz?



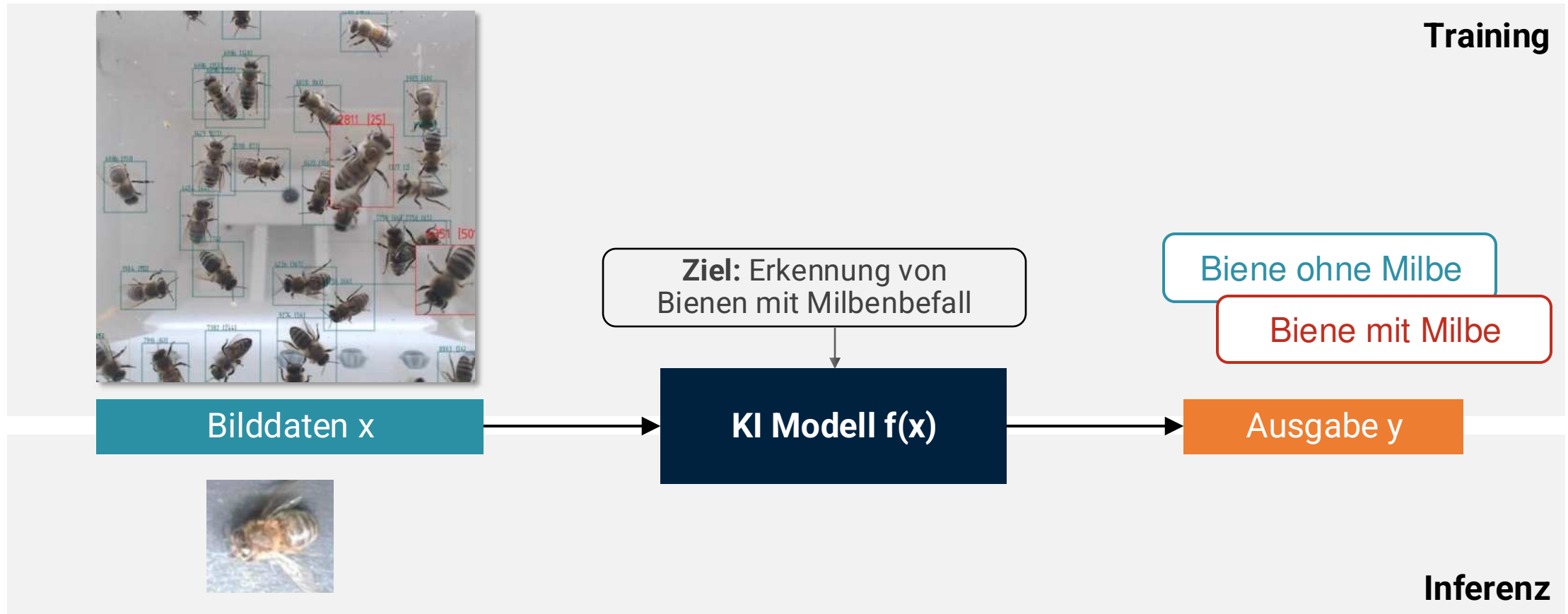
„Ein KI-System ist ein **maschinenbasiertes System**, das **expliziten oder impliziten Zielsetzungen** dient und **aus erhaltenen Inputs darauf schließt**, wie Vorhersagen, Inhalte, Empfehlungen, Entscheidungen oder andere **Outputs zu erzeugen sind**, die die **physische oder virtuelle Umgebung beeinflussen** können.“

Abbildung nach: OECD AI definition, updated 2023, <https://oecd.ai/en/wonk/ai-system-definition-update>

Teilgebiete der KI: Modellierungsparadigmen

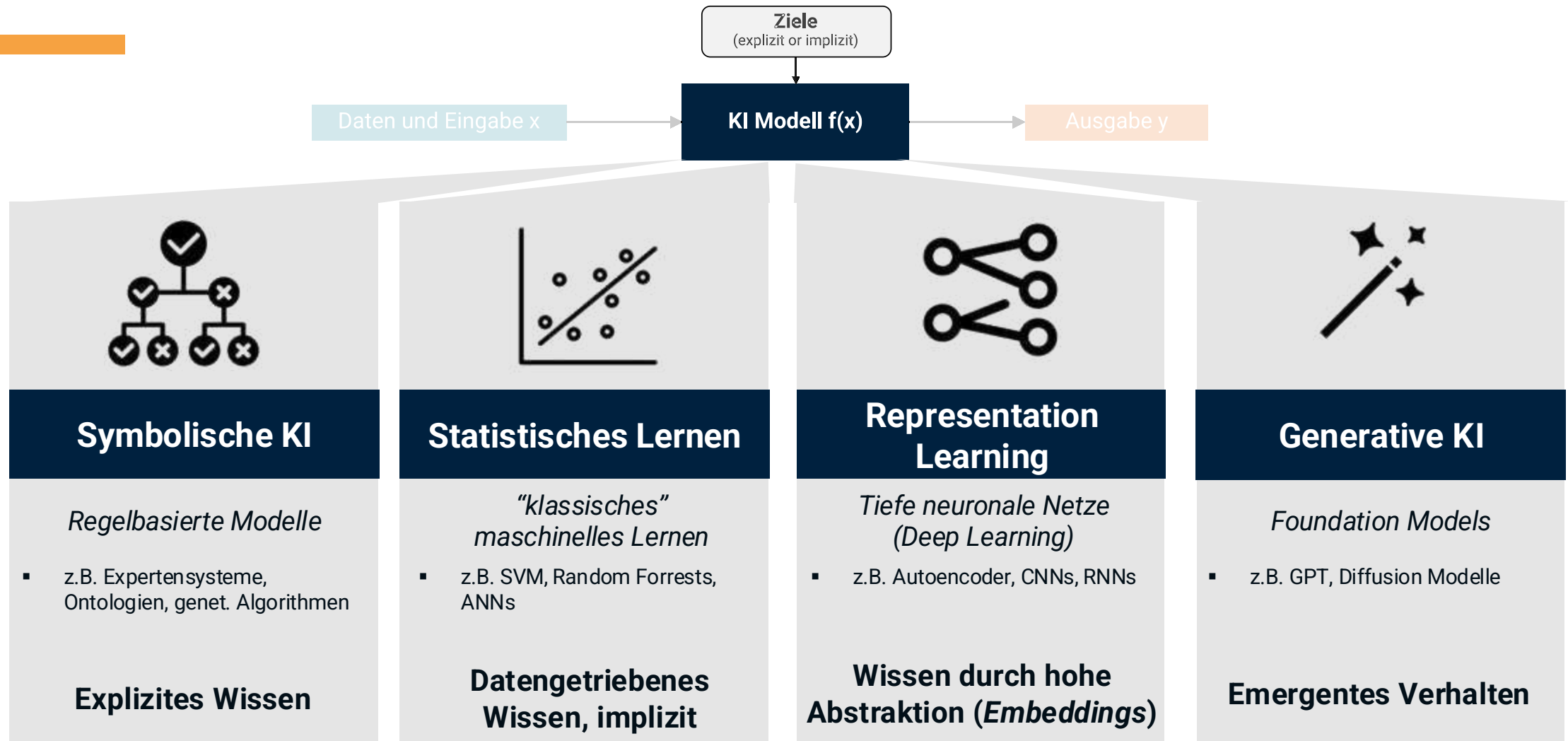


Beispiel: Erkennung der Varroa Milbe

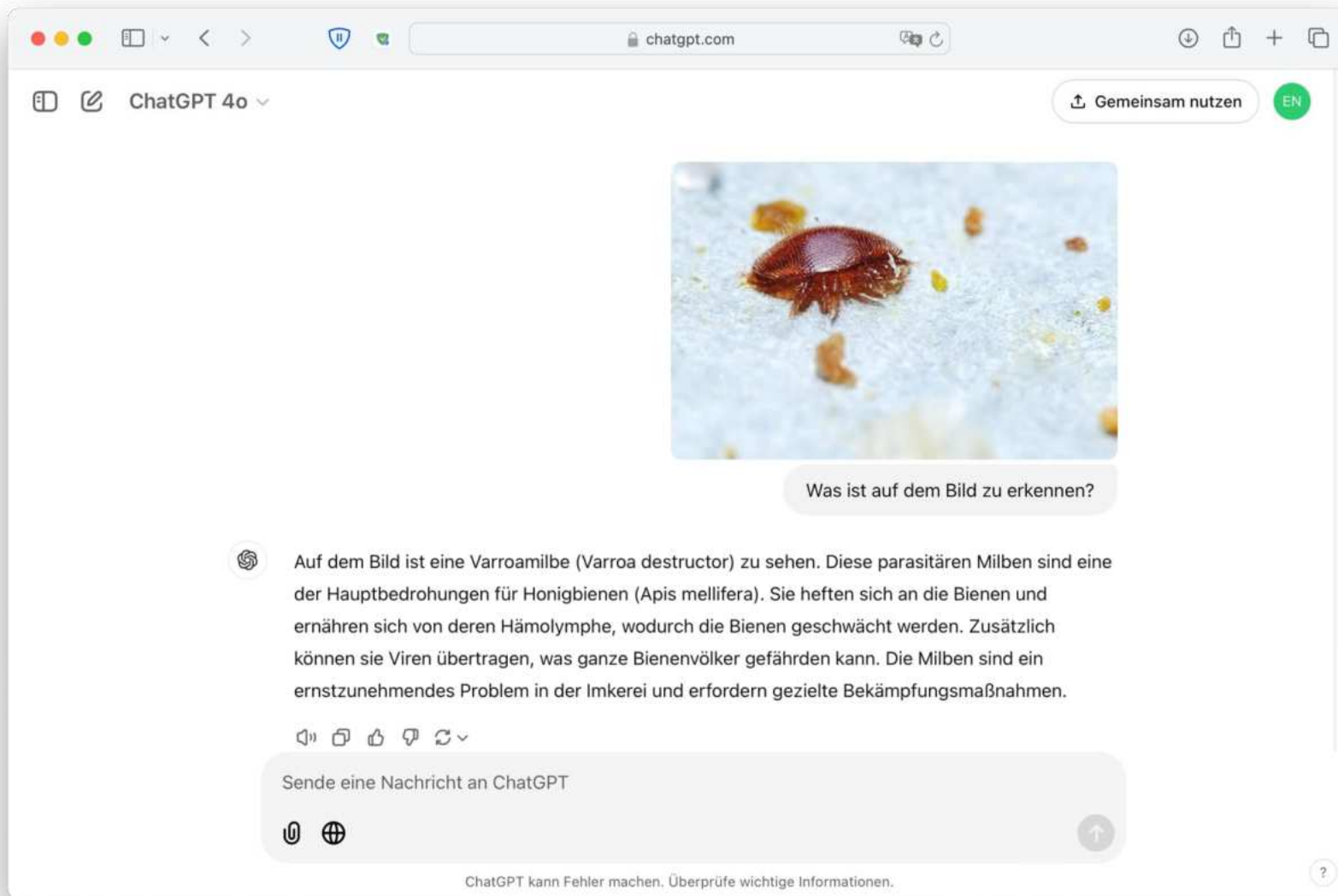


Bildquelle: h2, Projekt API Scan // apisys.de

Teilgebiete der KI: Modellierungsparadigmen



Beispiel: Erkennung der Varroa Milbe



The screenshot shows a web browser window with the URL 'chatgpt.com'. The interface includes a 'ChatGPT 4o' header, a 'Gemeinsam nutzen' button, and a language selector set to 'EN'. A user has uploaded a photograph of a Varroa mite and asked, 'Was ist auf dem Bild zu erkennen?'. The AI response identifies the mite as *Varroa destructor*, a parasitic mite that attacks honeybees (*Apis mellifera*). It explains that the mites feed on the bees' hemolymph, weakening them, and can also transmit viruses, posing a significant threat to beekeeping. Below the response are icons for voice, copy, like, dislike, and refresh. A text input field at the bottom contains the placeholder 'Sende eine Nachricht an ChatGPT' and a send button. A disclaimer at the bottom states 'ChatGPT kann Fehler machen. Überprüfe wichtige Informationen.'

Was ist auf dem Bild zu erkennen?

Auf dem Bild ist eine Varroamilbe (*Varroa destructor*) zu sehen. Diese parasitären Milben sind eine der Hauptbedrohungen für Honigbienen (*Apis mellifera*). Sie heften sich an die Bienen und ernähren sich von deren Hämolymphe, wodurch die Bienen geschwächt werden. Zusätzlich können sie Viren übertragen, was ganze Bienenvölker gefährden kann. Die Milben sind ein ernstzunehmendes Problem in der Imkerei und erfordern gezielte Bekämpfungsmaßnahmen.

Sende eine Nachricht an ChatGPT

ChatGPT kann Fehler machen. Überprüfe wichtige Informationen.



Wie kann anwendungsorientierte Forschung zu KI aussehen?

Anwendungsorientierte Forschung zu KI

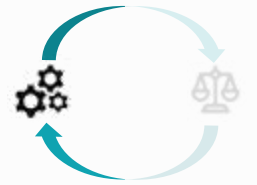
- **Forschungsgegenstand:**
KI-System
- **Fokussierte Dimensionen:**
Daten, Algorithmen, Infrastruktur



Daneben: **Forschung mit KI**
Einsatz von KI-Verfahren für die
Forschung

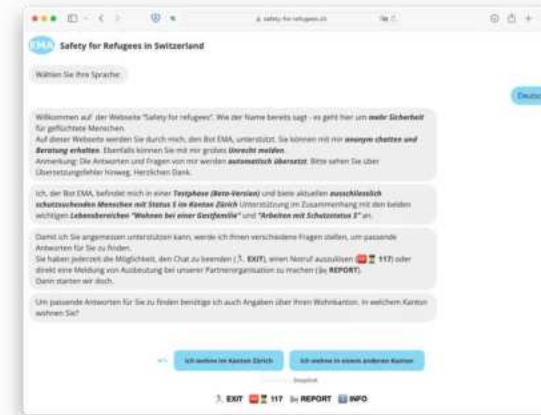
- **Forschungsgegenstand:**
KI-System in seiner Umgebung
- **Fokussierte Dimensionen:**
Anwendungsintegration, Mensch,
Organisation

Interdisziplinäre Anwendungsbeispiele



KITT: Demonstrationsprojekt
Automatische Erstellung von
Kartendaten aus Luftbildern
(h2 Magdeburg)

<https://h2.de/kitt>



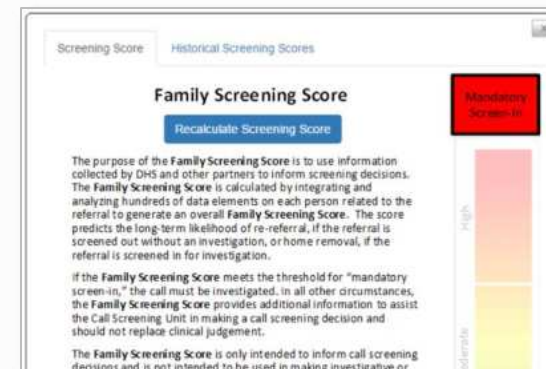
EMA Chatbot des Projektes
safety-for-refugees
(HAW Zürich)

<https://safety-for-refugees.ch>



KIADEKU: Bilderkennung zur
Dekubitusversorgung
(Unikliniken Essen, München)

<https://www.interaktive-technologien.de/projekte/kiadeku>



AFST: Allegheny Family
Screening Tool für das
Department of Human
Services, Pennsylvania
(Auckland Univ. of Tech.)

<https://www.allegheycounty.us/Services/Human-Services-DHS/DHS-News-and-Events/Accomplishments-and-Innovations/Allegheny-Family-Screening-Tool>

Beforschung von KI-Anwendungen



KI Auswirkungen auf die industrielle Arbeitswelt
(it's OWL / Fraunhofer IEM)
<https://arbeitswelt.plus>



KI für NGO: Bedarf, Akzeptanz und Umsetzungsmöglichkeiten
(OTH Regensburg)

<https://sozial-gesundheitswissenschaften.oth-regensburg.de/forschung/ist-institut-fuer-technikfolgenabschaetzung/projekte/projekte-im-bereich-mensch-technik-interaktion/kiniro>



Biotechnologie und künstliche Intelligenz: Risiken der Forschung für Sicherheit und Proliferation von Biowaffen

Projektteam: Harald König, Arnold Sauter

Themenfeld: Biotechnologie und Gesundheit, Infrastrukturen und Sicherheit

Themeninitiative: Verteidigungsausschuss

Analyseansatz: TA-Projekt

Starttermin: September 2024

Endtermin: 2026



<https://www.tab-beim-bundestag.de>

Büro für Technikfolgen-Abschätzung
beim Deutschen Bundestag



Technikfolgen: Systemische und existenzielle Risiken von KI

Forschungsprojekt am KIT entwickelt Handlungsempfehlungen, um Risiken bei der Entwicklung von KI-basierten Technologien zu reduzieren

<https://www.kit.edu/kit/202404-systemische-und-existenzielle-risiken-von-ki.php>

Auswahl: Forschungsprojekte an der h2

- **SynerKI:** Synergetische KI-Lösungen für ganzheitliche Prozesse in Operations, Finance, HR, Management und Marketing
(Prof. Behrendt, FB Wirtschaft)
- **KID:** Autonom agierende detektierende KI-Drohntechnologie für den Wasser- und Umweltsektor
(Prof. Bernd Ettmer, FB WUBS)
- **ChatGPT / Co. im medialen Diskurs** - soziale Gerechtigkeit ausgeblendet?
(Prof.in Elke Grittmann, FB SGM)

...und viele weitere!

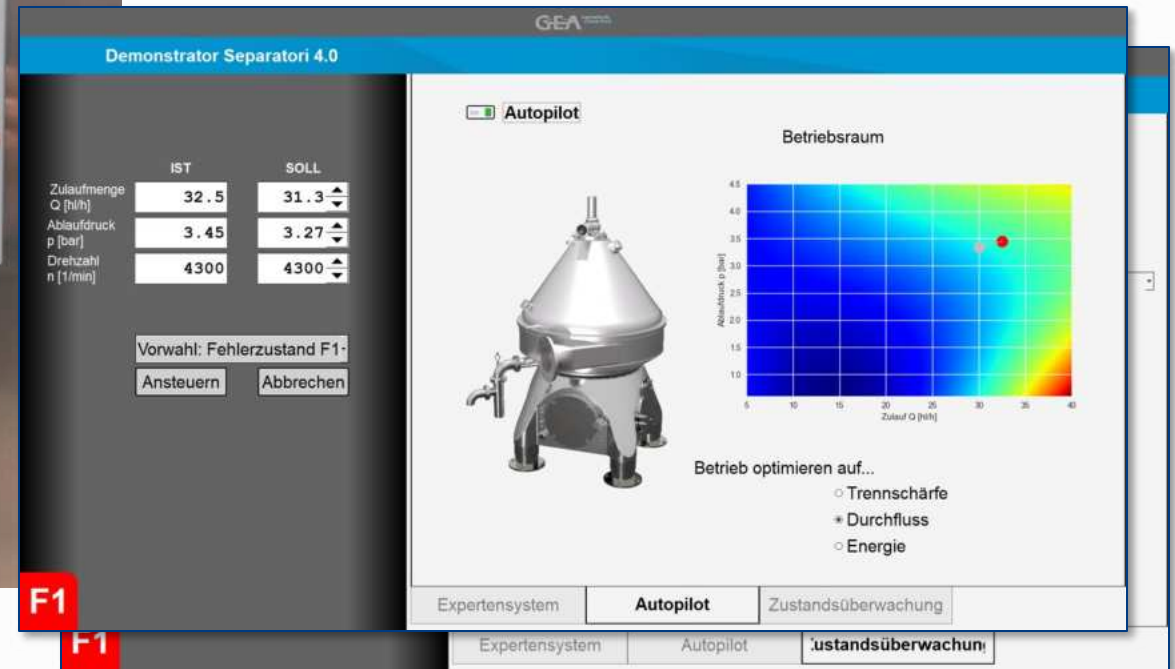


Prozesssteuerung an der industriellen Zentrifuge

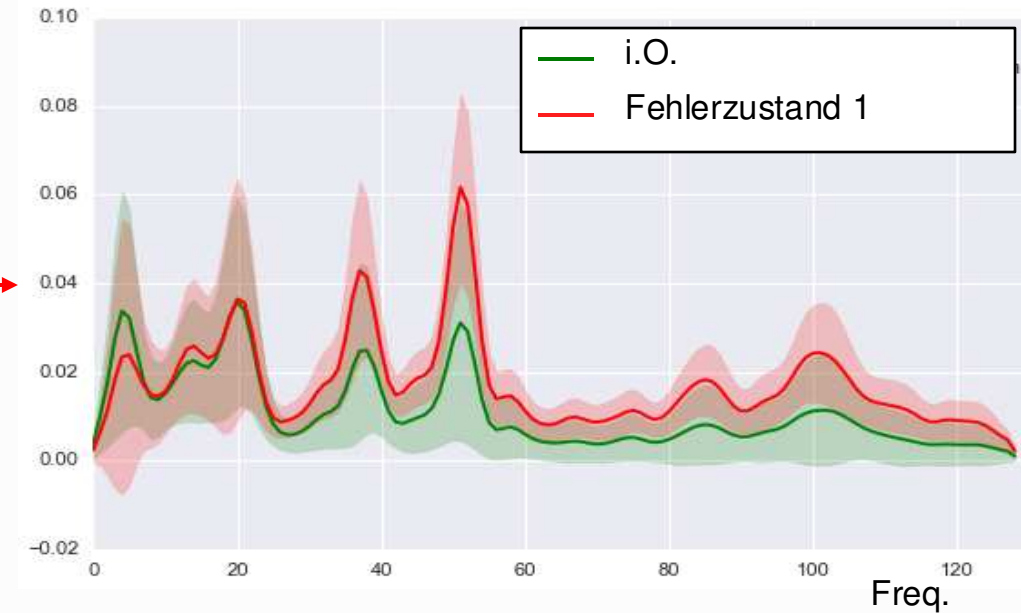
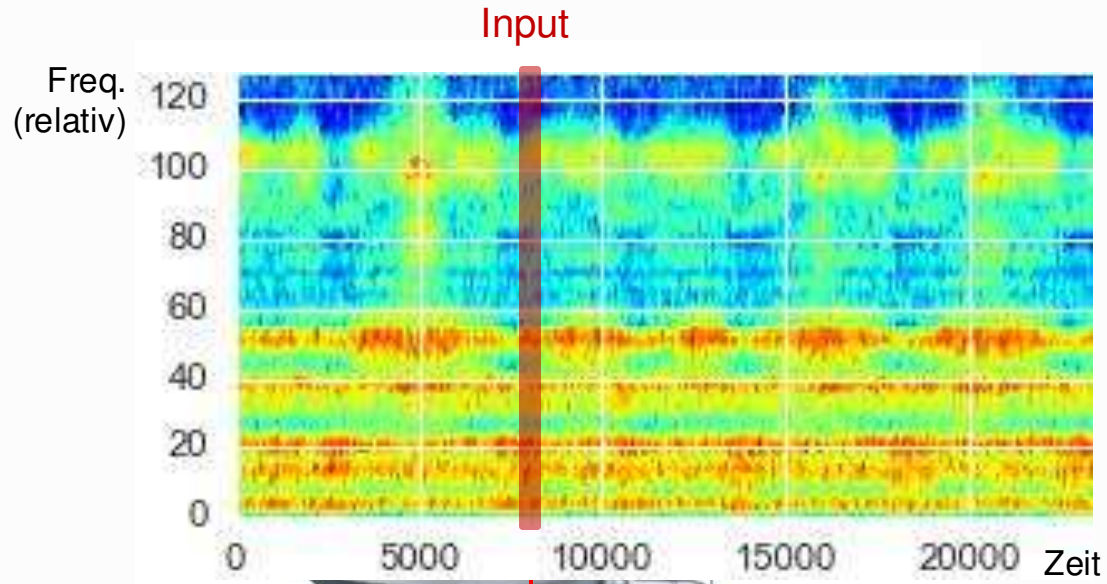


Nutzerakzeptanz erreichen durch hohe Usability und Erklärbarkeit!

- Erkennung von Maschinenfehlern mittels Maschinellen Lernverfahren
- Assistenzsystem für den Bediener
- Automatische Betriebspunktoptimierung



Prozesssteuerung an der industriellen Zentrifuge



Domänenwissen ist essentiell für die Anwendungsentwicklung!

- > 250 Messreihen @ 48kHz
30 x 1 Sekunde Zeitfenster
- Manuelle Annotation (Testlauf Wasser)
- Siemens S7 + IPC für Datenaufnahme

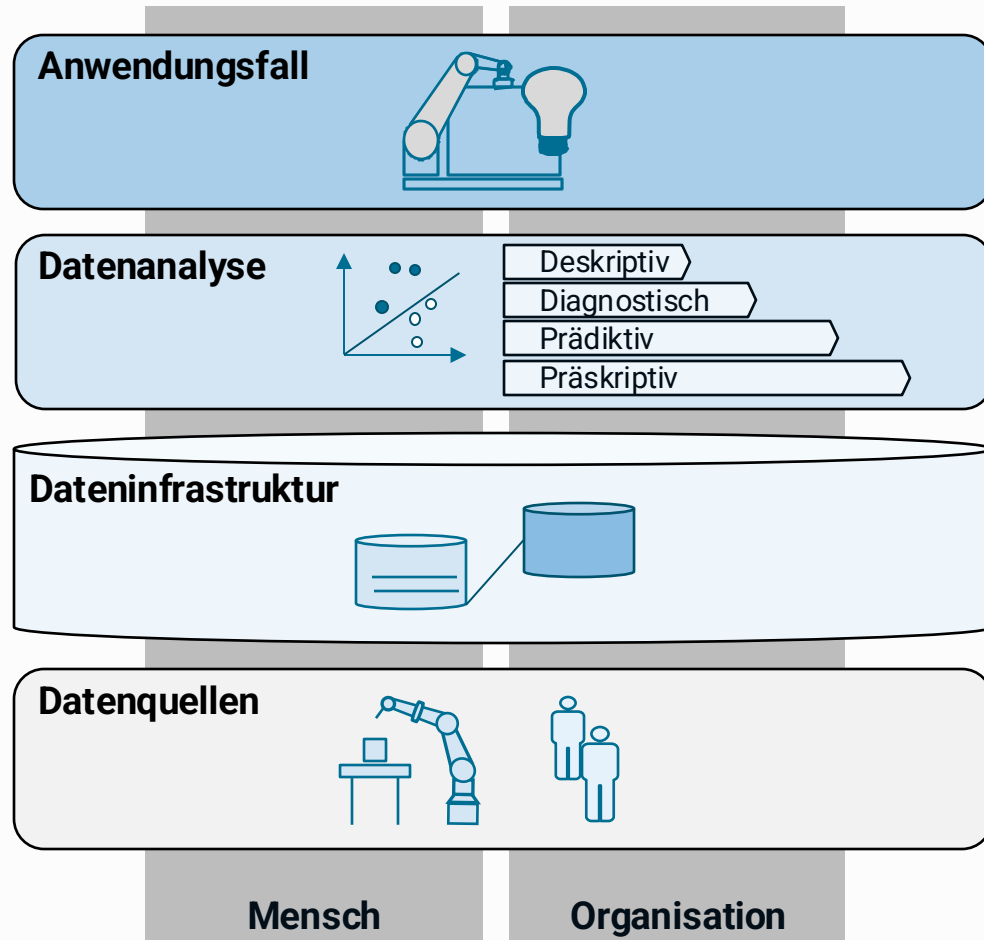
Typische Herausforderungen



- ▶ Wie gestaltet man sinnvolle Anwendungsfälle für den Nutzer?
- ▶ Welche Algorithmen setzt man ein? Wie trainiert man erfolgreich?
- ▶ Wie und wo speichere und (vor-)verarbeite ich meine Daten?
- ▶ Welche Eigenschaften haben meine Daten? Was ist der genaue Kontext?

Typische Herausforderungen

Angewandte KI ist
interdisziplinär!



Wie gestaltet man sinnvolle Anwendungsfälle für den Nutzer?



Welche Algorithmen setzt man ein?
Wie trainiert man erfolgreich?



Wie und wo speichere und (vor-)verarbeite ich meine Daten?



Welche Eigenschaften haben meine Daten?
Was ist der genaue Kontext?



Welche Auswirkungen ergeben sich?
Wie geht man mit diesen um?

Thesen zu anwendungsorientierter KI-Forschung

Nutzerakzeptanz erreichen durch **hohe Usability** und **Erklärbarkeit!**



Für eine erfolgreiche KI-Anwendung sind hohe **Usability**, **Erklärbarkeit** und **verantwortungsvolle** Entwicklung entscheidend.

Domänenwissen ist essentiell für die Anwendungsentwicklung!



KI ist ein Werkzeug, dessen Nutzen sich nur in geeigneten Anwendungen mit **ausreichendem Domänenwissen** entfaltet.

Angewandte KI ist **interdisziplinär!**



Angewandte KI ist durch und durch **interdisziplinär** und vereint **technische**, **domänenspezifische**, **ethische** und **soziale** Aspekte.



ZAKKI

Zentrale Anlaufstelle für innovatives Lehren und Lernen interdisziplinärer Kompetenzen der KI

Vielen Dank für Ihre Aufmerksamkeit!

Prof. Dr.-Ing. Sebastian von Enzberg
sebastian.von.enzberg@h2.de

h2-Science-Day
Magdeburg, 20.11.2024