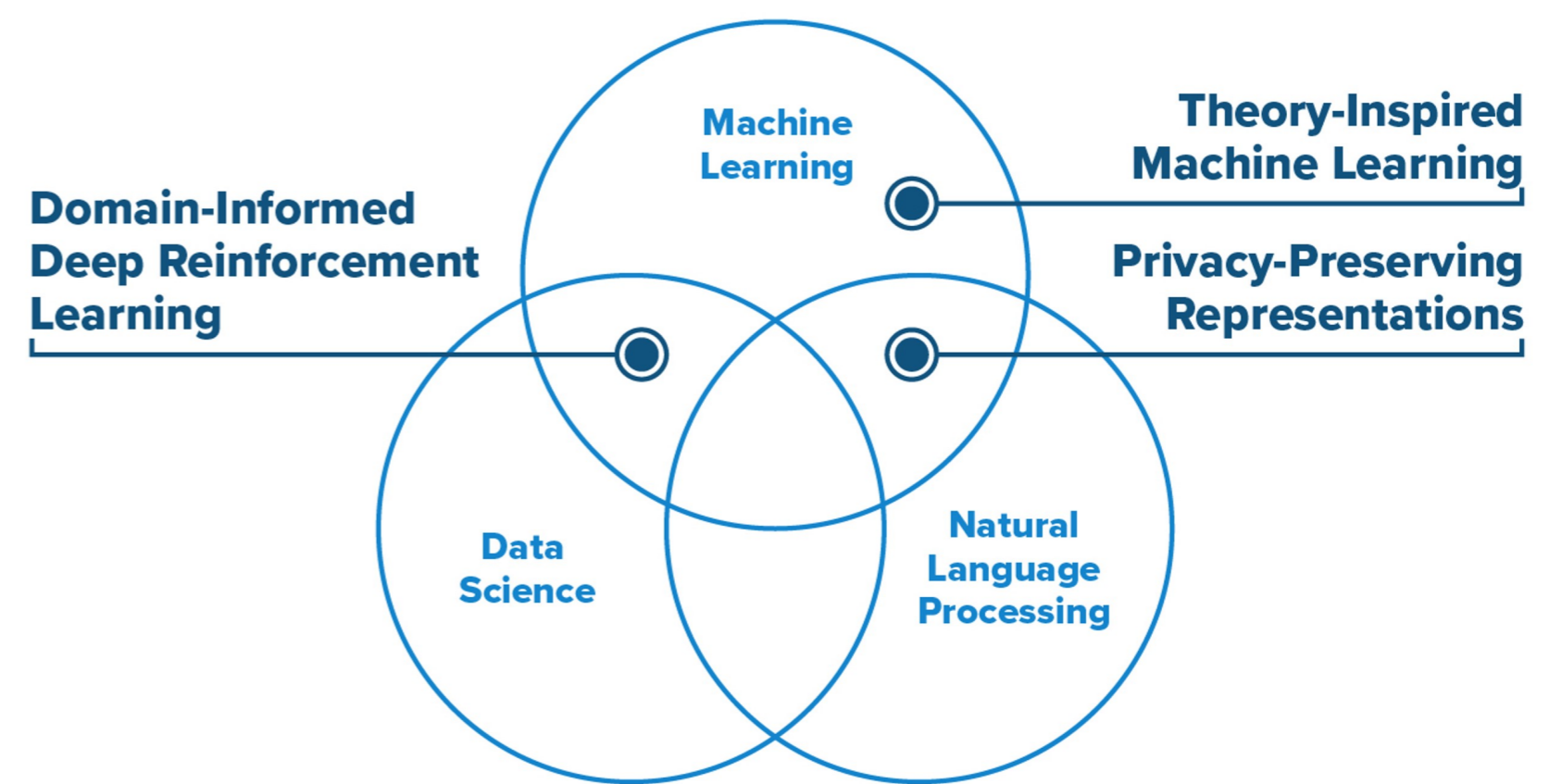# KNOW Center

AREA METHODS & ALGORITHMS FOR AI

The research area Methods & Algorithms for AI (AAI) is part of the **Know-Center GmbH, a research center for trustworthy AI and data**. Our area AAI

- ► covers a wide range of topics in data science, machine learning, and natural language processing
- ► conducts **research on the theoretical underpinnings of artificial intelligence and provides algorithmic solutions** for industry and the scientific community
- ► incorporates domain knowledge from respective disciplines during algorithmic model development

Domain-Informed Deep Reinforcement Learning · Machine Learning · Theory-Inspired Machine Learning · Privacy-Preserving Representations · Data Science · Natural Language Processing

MACHINE LEARNING & DEEP LEARNING

Objectives

- ► include domain knowledge in ML models (theory-inspired ML)
- ► develop in-depth understanding of ML systems

Keywords: Physics-informed ML, probabilistic ML (representation learning, variational approaches, Gaussian processes), time series analysis (forecasting, anomaly detection), transfer learning, reinforcement learning, image segmentation & classification, information theory & causality, Bayesian optimization

Spotlight: Physics-Informed Neural Networks
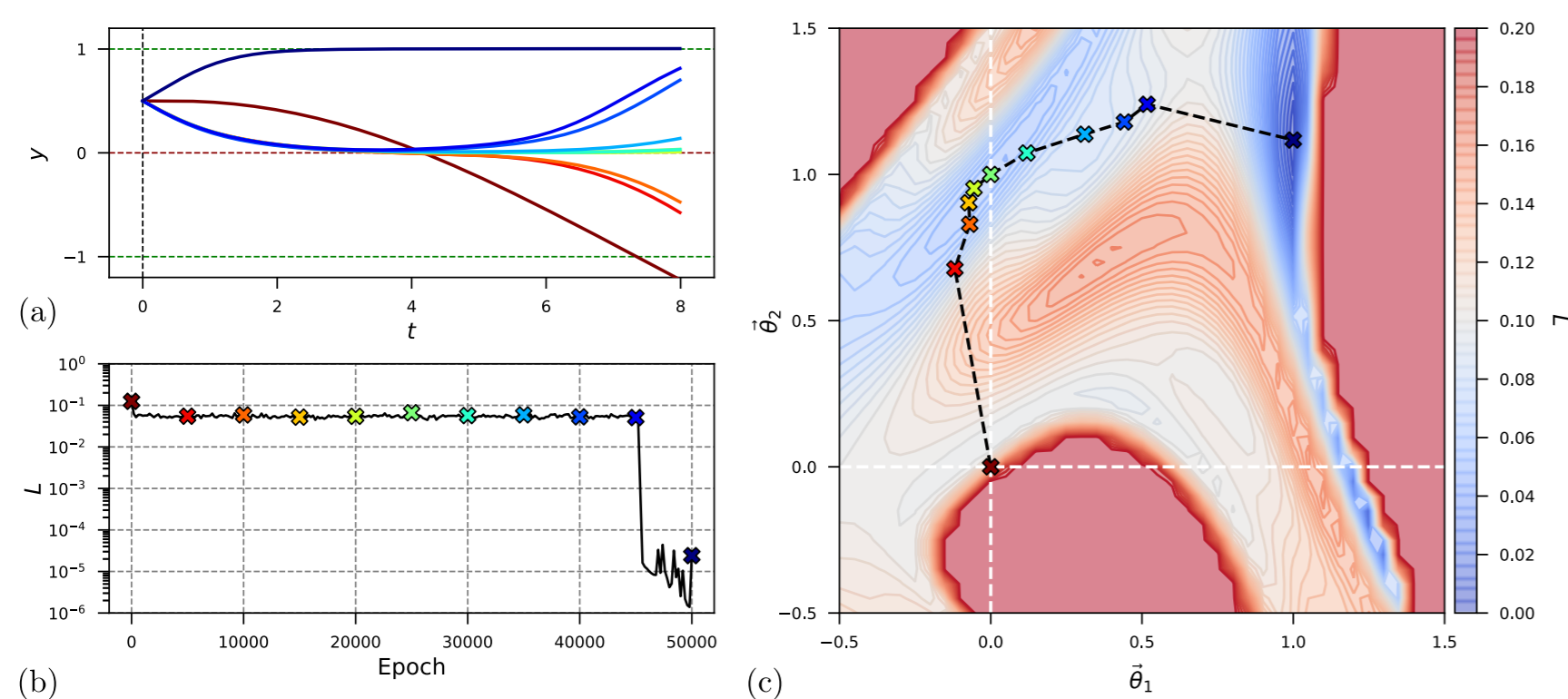
Training PINNs sometimes stalls (b) at incorrect solutions (a), because unstable fixed points of the differential equation create local minima in the loss landscape (c) [1]. The loss landscape is smoother for smaller computational domains, which explains the success of domain decomposition approaches.

[1] F. M. Rohrhofer et al., On the role of fixed points of dynamical systems in training physics-informed neural networks; Trans. Machine Learning Research, 2023
[2] F. M. Rohrhofer et al., On the Pareto-Front of Physics-Informed Neural Networks (arXiv:2105.00862), submitted

Spotlight: Image Segmentation for Wear Detection
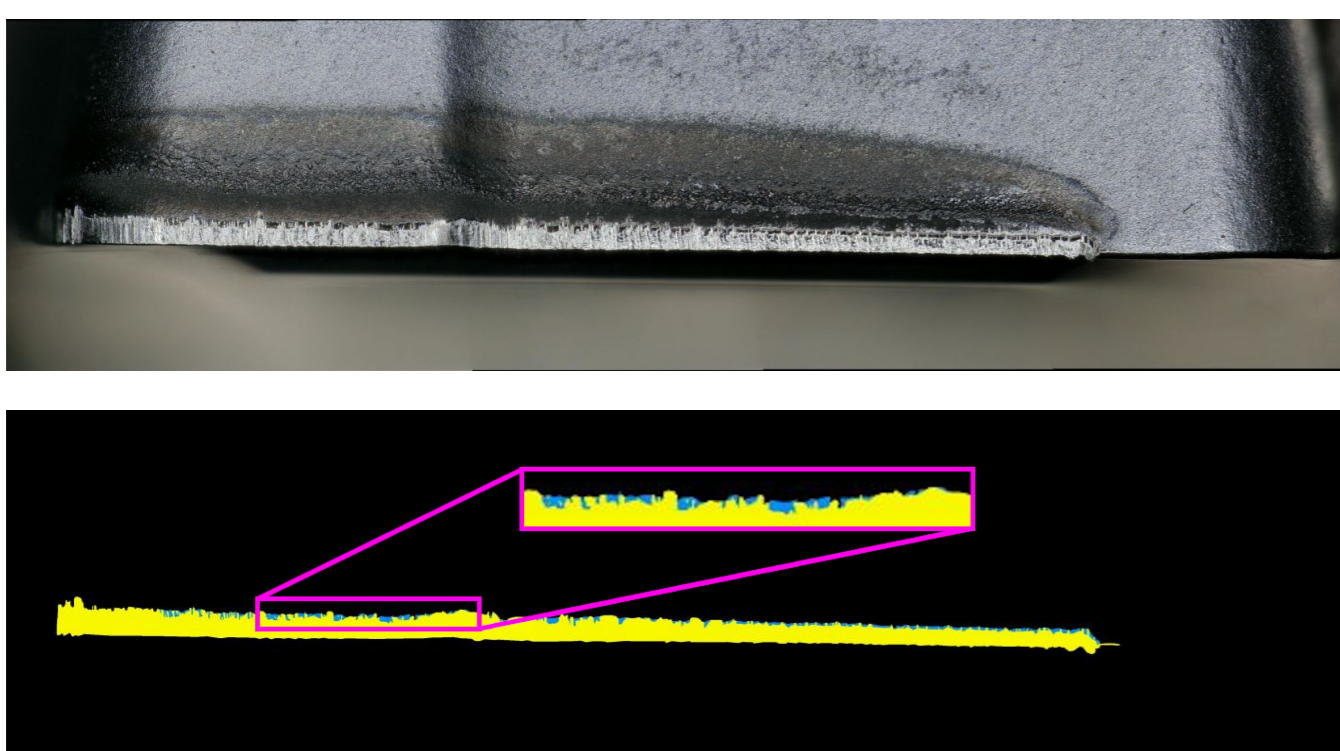
Light optical microscopy image of cutting insert (top) and associated mask (bottom) with abrasive wear area in blue and area of transferred work piece material in yellow. Image segmentation was performed with a U-Net.

[3] E. Schlager et al., Evaluation of Data Augmentation and Loss Functions in Semantic Image Segmentation for Drilling Tool Wear Detection (arXiv:2302.05262), submitted

NATURAL LANGUAGE PROCESSING

Objectives

- ► understand what constitutes privacy and preserve it
- ► explore balancing effects between preserving privacy vs. feasibility/requirements of NLP tasks
- ► reduce language complexity according to problem setting

Keywords: Natural language understanding, text simplification, dialogue systems, large language models, information & relation extraction, privacy preservation and knowledge graphs

Spotlight: Privacy Preservation in Deep Neural Networks
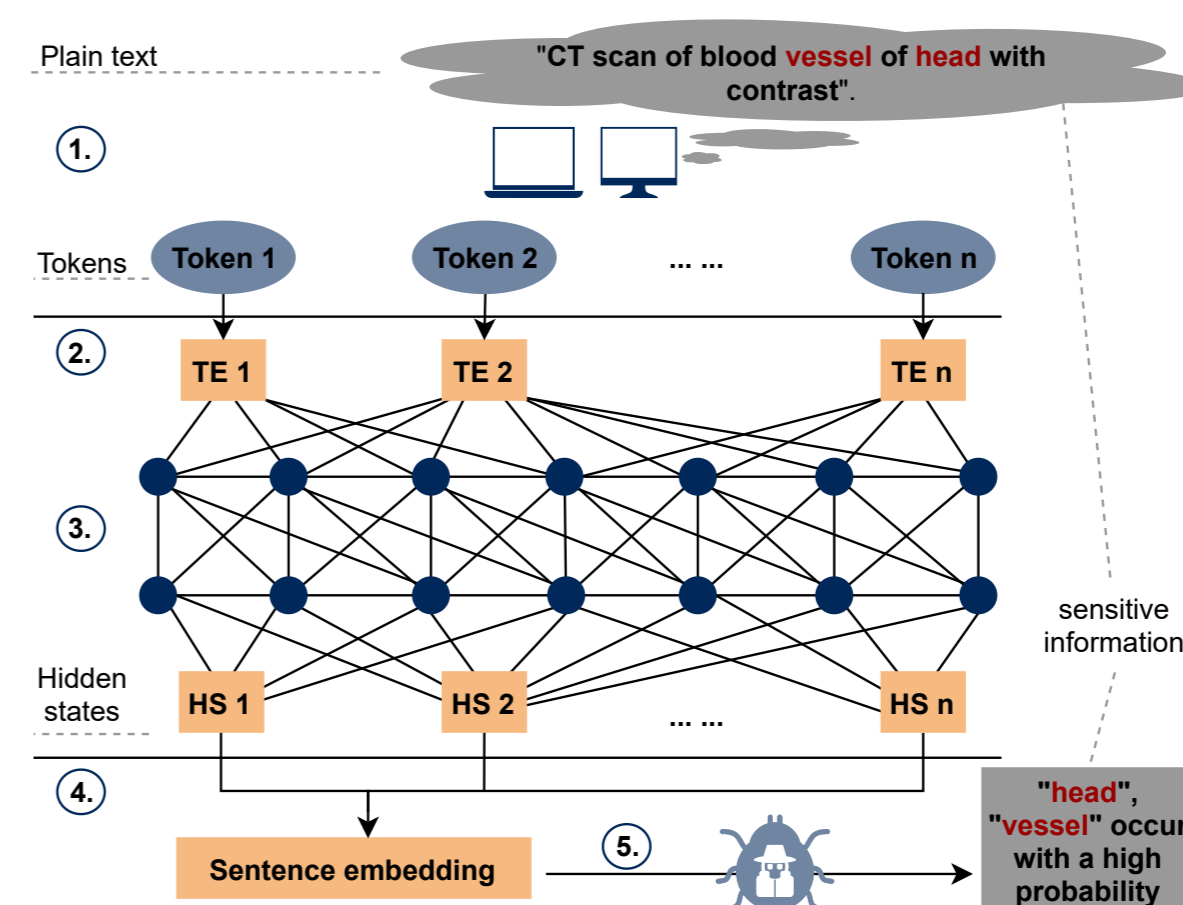
Privacy risks [4]: the unintended memorization and successive reconstruction (step 5) of sensitive information (in red) from sentence embeddings (steps 1–4).

[4] S. Sousa et al., How to keep text private? A systematic review of deep learning methods for privacy-preserving natural language processing; Artificial Intelligence Review, 2022.
[5] S. Sousa et al., Assessing the impact of differential privacy in NLP tasks based on transfer learning with deep neural networks, in preparation.

DATA SCIENCE

Objectives

- ► Support ML model development and operation
- ► Provide data engineering services for many different types of data

Keywords: Combinatorial and non-convex optimization, data engineering, ML engineering & MLOps, signal processing

Spotlight: Optimization applications in route planning

[6] C. Schweimer et al., A route pruning algorithm for an automated geographic location graph construction; Scientific Reports 11:11547, 2021.