Constantin Kühne

Email | GitHub | LinkedIn | Google Scholar

Education

MSc Data Engineering (2023 - present)
Hasso Plattner Institute/ University of Potsdam

- Exchange semester at Eindhoven University of Technology (2025)
- Relevant Coursework: Generative AI Models, Deep Learning, Machine Learning for Image Analysis, Reinforcement Learning & Algorithm Discovery, Knowledge Engineering, Advanced Machine Learning Seminars

BSc Wirtschaftsinformatik/ Business Intelligence (2019 - 2023)
Hochschule der Bayerischen Wirtschaft

Research Experiences

Semantically Complex Group Object Detection with Vision Language Models (06/2025 - present)
Hasso Plattner Institute

- Ongoing master thesis under Prof. Dr. Gerard de Melo at the chair of Artificial Intelligence and Intelligent Systems
- Research on adapting vision language models for semantic group object detection tasks (e.g., detecting a group of people and their actions in street view photos)
- Development of novel neural network architectures and training strategies for this task as well as evaluation on relevant datasets

Icewafl: A Configurable Data Stream Polluter (11/2024 - 04/2025)
Hasso Plattner Institute

- Research assistant role under Dr. Lisa Ehrlinger at the Information Systems chair of Prof. Dr. Felix Naumann
- Research and software development on a framework for data pollution via Apache Flink [Link]
- Evaluation of the effects of data errors on machine learning models

Beyond Balanced: Learning Static Search Trees with Tree-MDPs (10/2024 - 04/2025)
Hasso Plattner Institute

- Research seminar under Alexander Kastius; chair for Artificial Intelligence and Sustainability of Prof. Dr. Ralf Herbrich
- Designed a reinforcement learning framework (Tree-MDP + A2C) to construct static binary search trees for database indexing that adapt to skewed query distributions [Link]
- Achieved up to 40% faster average lookup times on skewed query distributions vs. traditional binary search
- Demonstrated scalability and adaptability to varying data distributions

TrueSkill/MLPs Beyond Gaussians (10/2024 - 03/2025)
Hasso Plattner Institute

- Master's project under Prof. Dr. Ralf Herbrich at the chair for Artificial Intelligence and Sustainability
- Research on more accurate factor graph inference via mixture models (uniform mixtures and Gaussian mixtures) [Link]
- Evaluation of these methods on TrueSkill and Multilayer Perceptrons

Learning Disentangled Representations with Identifiable Diffusion Models (10/2023 - 09/2024)
Hasso Plattner Institute

- Research seminar under Alexander Rakowski and Eshant English at the Digital Health & Machine Learning chair of Prof. Dr. Christoph Lippert
- Research on combining identifiable Variational Autoencoders with Denoising Diffusion Probabilistic Mode to gain disentangled latent representations [Part1] [Part2]

Creation of knowledge graphs based on natural language texts in a domain-specific context (10/2022 - 02/2023)
Hochschule der Bayerischen Wirtschaft & Infineon Technologies AG

- Bachelor's thesis on automatically creating knowledge graphs of domain-specific file content
- Fine-tuning of language models (RoBERTa) for triple extraction
- Design and implementation of a semi-automatic triple extraction pipeline via seeds and semi-supervised machine learning

Industry Experiences

System Engineer (03/2023 - 08/2023)
Infineon Technologies AG

- Implementation and establishment of a Customer Data Platform
- Analysis and value creation out of customer data
- Consolidation and quality assurance of different data sources

Dual Studies at the Analytics Department (09/2019 - 03/2023)
Infineon Technologies AG

- Development, implementation, and operationalization of software solutions in the field of Data Science with a focus on natural language processing
- Creation of semantic search for domain-specific technical documents via Elasticsearch and language models
- Design and implementation of an automatic matching engine for finding and matching funding programmes to fund managers, reducing their workload by up to 50%
- Creation and operationalization of a framework to train machine learning algorithms for the optimization of stock management, resulting in 500K€ (with prospects up to 6.5M€) scrap avoidance and 1600 working hours saved per year

Technical Skills

- Python
- Java
- Typescript
- SQL
- PyTorch
- Lightning
- Weights & Biases
- Stable Baselines3
- Docker
- Databricks
- Elasticsearch
- Tableau

Awards

- Deutschlandstipendium, 2023-2025

Publications

1. Christoph Schinninger and Fabian Panse and Constantin Kühne and Lisa Ehrlinger. Icewafl: A Configurable Data Stream Polluter. In Proceedings 28th International Conference on Extending Database Technology (EDBT 2025). https://doi.org/10.48786/edbt.2025.64

Languages

- German - (C2)
- English - (C1)
- French - (A2)