

Jan-Christoph Klie

✉ jck@klie-ai.de  [jcklie](https://github.com/jcklie)  klie-ai.de  [Google Scholar](https://scholar.google.com/citations?user=jcklie)  [in jcklie](https://in.jcklie)

Soon-to-graduate doctoral student and freelance AI/IT consultant. My research focuses on improving annotation efficiency and quality, enabling quicker and more cost-effective data collection for new and exciting applications. I am also interested in applying large language models to novel use cases, continual and interactive machine learning, and bringing ML research into production.

EDUCATION

TU DARMSTADT

PH.D. IN COMPUTER SCIENCE

Jan 2018 - Apr 2024 | Darmstadt, DE

Thesis Topic: *Improving NLP Annotation*

Quality and Efficiency, supervised by

[Prof. Dr. Iryna Gurevych](#)

TU DARMSTADT

M.Sc. IN COMPUTER SCIENCE

Oct 2014 - Oct 2017 | Darmstadt, DE

Robotics, computer vision, natural language processing and software engineering. Minor in control theory.

DHBW MANNHEIM

B.Sc. IN COMPUTER SCIENCE

Oct 2011 - Sep 2014 | Mannheim, DE

Integrated degree program with IBM.

SKILLS

PROGRAMMING

Experienced:

Java • \LaTeX • Python • TypeScript

Advanced:

C/C++ • HTML/CSS • Rust

Elementary:

Go • Kotlin • MatLab • R

TOOLS

Experienced:

Sklearn • PyTorch • Transformers

Advanced:

Android • React • Spring • Docker

Elementary:

AWS • Tensorflow • Vue

LANGUAGES

German ● ● ● ● ● ●
English ● ● ● ● ● ●
Chinese ● ● ● ○ ○ ○
Vietnamese ● ○ ○ ○ ○ ○

EXPERIENCE

Klie AI | FREELANCE MACHINE LEARNING/SOFTWARE ENGINEER

Oct 2022 - Present | Darmstadt, Germany

I am available for hire as a freelance AI/IT consultant. My expertise is consulting customers on data annotation, especially efficiency and quality. I also enjoy working on applied machine learning, software engineering or DevOps/MLOps projects.

Annotation Quality Annotation Efficiency Machine Learning Natural Language Processing
LLM Python Go React Java Typescript Rust Slurm DevOps MLOps

Apple | RESEARCH INTERN

May 2022 - Sep 2022 | Heidelberg, Germany

- Evaluated annotation error detection for OCR data annotation tasks.
- Investigated sample size computation and reduction for quality estimation through manual inspection for desired statistical guarantees.

Python Pytorch Statistics Data Annotation Data Quality Management

UKP Lab, TU Darmstadt | RESEARCH PROGRAMMER&SYSADMIN LEAD

Jan 2018 - May 2023 | Darmstadt, Germany

- This was the position funding my Ph.D studies in natural language processing.
- Worked on the [INCEpTION](#) project, a Java-based, extensible text annotation platform offering annotation assistance and knowledge management.
- Lead the system administration team that provides the IT for around 60 people in our lab (mail, chat, personal hardware, compute, support) so that they have the infrastructure and compute available for their research.
- Managed and participated in the IT modernization of our lab, including building a GPU cluster, rolling out modern monitoring, introducing a new chat platform, migrating and deprecating old services, digitizing workflows.
- Supervised sysadmin apprentices, student helper and thesis students.

Annotation Interactive NLP PyTorch LLM Java Spring Slurm Prometheus

Senacor | SOFTWARE DEVELOPER ONLINE PAYMENT SYSTEM

Aug 2016 - Sep 2017 | Frankfurt, Germany

Worked on a new online payment system backed by the largest banks in Germany.

Java Spring Android Scrum Event Sourcing CQRS

RMR, TU Darmstadt | SOFTWARE DEVELOPMENT ROBOTICS

Jul 2015 - Mar 2016 | Darmstadt, Germany

Developed a ROS package for a stereo camera mounted on Robotinos as well as a marker-based localization system leveraging four ceiling cameras.

Python C++ OpenCV ROS Matlab Aruco

IBM NGNCC via KONTIP GmbH | SOFTWARE DEVELOPMENT IoT

Oct 2014 - Jul 2016 | Frankfurt, Germany

Built a scalable IoT-Middleware to connect sensors and actuators of any manufacturer. Starting point was the platform developed during my B.Sc. thesis.

C++ Java Arduino Python SIP Linux PostgreSQL LoRa

PUBLICATIONS

1. Richard Eckart de Castilho, **Jan-Christoph Klie**, Iryna Gurevych. 2023. Integrating annotation tools into larger processes. *Under submission*.
2. **Jan-Christoph Klie**, Richard Eckart de Castilho, Iryna Gurevych. 2023. [Analyzing Dataset Annotation Quality Management in the Wild](#). In: *arXiv*. *Under submission*.
3. **Jan-Christoph Klie**, Ji-Ung Lee, Kevin Stowe, Gözde Gül Sahin, Nafise Sadat Moosavi, Luke Bates, Dominic Petrak, Richard Eckart de Castilho, Iryna Gurevych. 2023. [Lessons Learned from a Citizen Science Project for Natural Language Processing](#). In: *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 6982–6993.
4. **Jan-Christoph Klie**, Bonnie Webber, Iryna Gurevych. 2023. [Annotation Error Detection: Analyzing the Past and Present for a More Coherent Future](#). In: *Computational Linguistics*, 49 (1): 157–198.
5. Ji-Ung Lee*, **Jan-Christoph Klie***, Iryna Gurevych. 2022. [Annotation Curricula to Implicitly Train Non-Expert Annotators](#). In: *Computational Linguistics*, 48 (2): 343–373.
(*: equal contribution)
6. **Jan-Christoph Klie**, Richard Eckart de Castilho, Iryna Gurevych. 2020. [From Zero to Hero: Human-In-The-Loop Entity Linking in Low Resource Domains](#). In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6982–6993.
7. Richard Eckart de Castilho, Nancy Ide, Jin-Dong Kim, **Jan-Christoph Klie**, Keith Suderman. 2019. [A Multi-Platform Annotation Ecosystem for Domain Adaptation](#). In: *Proceedings of the 13th Linguistic Annotation Workshop*, pages 189–194.
8. Richard Eckart de Castilho, Nancy Ide, Jin-Dong Kim, **Jan-Christoph Klie**, Keith Suderman. 2019. [Towards cross-platform interoperability for machine-assisted text annotation](#). In: *Genomics & Informatics*, 17(2):e19.
9. Beto Bouldosa, Richard Eckart de Castilho, Naveen Kumar, **Jan-Christoph Klie**, Iryna Gurevych. 2018. [Integrating Knowledge-Supported Search into the INCEpTION Annotation Platform](#). In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 127–132.
10. **Jan-Christoph Klie**, Michael Bugert, Beto Bouldosa, Richard Eckart de Castilho, Iryna Gurevych. 2018. [The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation](#). In: *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*, pages 5–9.
11. Teresa Botschen, Iryna Gurevych, **Jan-Christoph Klie**, Hatem Mousselly-Sergieh, Stefan Roth. 2018. [Multimodal Frame Identification with Multilingual Evaluation](#). In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1481–1491.