

Education

- 2021–Present **PhD, Computer Science**, *University of Copenhagen*
Research Focus: Natural Language Processing, Multimodal Machine Learning
Advisor: Prof. Anders Søgaard
- 2018–2020 **M.Sc., Computer Science**, *TU Darmstadt*, *Cumulative GPA: 1.34**
Focus: Natural Language Processing, Machine Learning, Artificial Intelligence
Thesis Supervisors: Dr. Jonas Pfeiffer, Prof. Iryna Gurevych
- 2015–2018 **B.Sc., Computer Science**, *DHBW Stuttgart*, *Cumulative GPA: 1.4* (ECTS Grade A)*

*On a scale of 1-5, 1.0 being the best possible grade

Professional Experience

- 03/2025–
04/2025 **University of Cambridge**, *Visiting PhD Student*, Cambridge, UK
Research visit at Language Technology Lab (LTL), hosted by Prof. Anna Korhonen.
- 08/2024–
02/2025 **Amazon AGI**, *Applied Scientist Intern*, Cambridge, UK
Post-training (RLHF/RLAIF) for Amazon Nova Sonic. Research on modular and multi-objective preference optimization for conversational speech in multimodal LLMs.
- 06/2023–
12/2023 **Meta, Fundamental AI Research (FAIR)**, *Research Scientist Intern*, Menlo Park, USA
Research on multimodal representation learning (video/text). Developed state-of-the-art technique for sign language translation through self-supervised video pretraining at scale.
- 10/2018–
03/2021 **DXC Technology**, *Technical Consultant – DualStudy Master*, Frankfurt a.M., Germany
Technical consulting (data engineering and DevOps) for client in the area of autonomous driving.
- 10/2015–
09/2018 **Hewlett Packard Enterprise**, *Intern – DualStudy Bachelor*, Bad Homburg, Germany
Integrated program that combined studies in computer science at DHBW Stuttgart with six internships in various departments and locations at Hewlett Packard Enterprise. Included an internship with Dr. Harumi Kuno at Hewlett Packard Labs in Palo Alto, USA.

Selected Publications

- ACL 2025 Antonia Karamolegkou, Malvina Nikandrou, Georgios Pantazopoulos, Danae Sanchez Villegas, **Phillip Rust**, Ruchira Dhar, Daniel Hershcovich, and Anders Søgaard. Evaluating Multimodal Language Models as Visual Assistants for Visually Impaired Users.
- HuCLLM @
ACL 2024 Antonia Karamolegkou, **Phillip Rust**, Yong Cao, Ruixiang Cui, Anders Søgaard, and Daniel Hershcovich. Vision-Language Models under Cultural and Inclusive Considerations.
- ACL 2024 **Phillip Rust**, Bowen Shi, Skyler Wang, Necati Cihan Camgöz, and Jean Maillard. Towards Privacy-Aware Sign Language Translation at Scale.
- EMNLP 2023 Jonas Lotz, Elizabeth Salesky, **Phillip Rust**, and Desmond Elliott. Text Rendering Strategies for Pixel Language Models.
- EMNLP 2023 Nadav Borenstein, **Phillip Rust**, Desmond Elliott, and Isabelle Augenstein. PHD: Pixel-Based Language Modeling of Historical Documents.

- ICLR 2023 **Phillip Rust**, Jonas F. Lotz, Emanuele Bugliarello, Elizabeth Salesky, Miryam de Lhoneux, and Desmond Elliott. Language Modelling with Pixels. (notable top 5%).
- ICML 2023 **Phillip Rust** and Anders Søgaard. Differential Privacy, Linguistic Fairness, and Training Data Influence: Impossibility and Possibility Theorems for Multilingual Language Models.
- ACL 2022 Daniel Hershcovich, Stella Frank, Heather Lent, Miryam de Lhoneux, Mostafa Abdou, Stephanie Brandl, Emanuele Bugliarello, Laura Cabello Piqueras, Ilias Chalkidis, Ruixiang Cui, Constanza Fierro, Katerina Margatina, **Phillip Rust**, and Anders Søgaard. Challenges and Strategies in Cross-Cultural NLP.
- ACL 2021 **Phillip Rust**, Jonas Pfeiffer, Ivan Vulić, Sebastian Ruder, and Iryna Gurevych. How Good is Your Tokenizer? On the Monolingual Performance of Multilingual Language Models.
- ACL 2020 Gözde Gül Şahin, Yova Kementchedjhieva, **Phillip Rust**, and Iryna Gurevych. PuzzLing Machines: A Challenge on Learning From Small Data.

Summer & Winter Schools

- 2021 **Lisbon Machine Learning Summer School (LxMLS)**, Virtual
- 2021 **Advanced Language Processing Winter School (ALPS)**, Virtual
- 2023 **HPLT & NLPL Winter School**, Skeikampen, Norway

Invited Talks

- 08/2022 **Amazon AWS AI (Tel Aviv)**, *Language Modelling with Pixels*, Virtual
- 09/2022 **Pioneer Centre for AI**, *Language Modelling with Pixels*, Copenhagen, Denmark
- 01/2023 **Amazon AWS AI (New York)**, *Language Modelling with Pixels*, Virtual
- 01/2023 **COLT Group, UPF Barcelona**, *Language Modelling with Pixels*, Virtual
- 03/2023 **ML Collective Reading Group (NLP & CV)**, *Language Modelling with Pixels*, Virtual

Teaching

- 21/22; 22/23 **Advanced Topics in Natural Language Processing (NDAK19001U)**,
Teaching Assistant, University of Copenhagen
- 21/22 **Introduction to Data Science (NDAK16003U)**,
Teaching Assistant, University of Copenhagen

Academic Service

Reviewing: ACL Rolling Review, ACL, EMNLP, EACL, MRL @ EMNLP, TiE @ ECCV, TL4NLP @ NeurIPS
Workshop Program Committee: MRL @ EMNLP, TL4NLP @ NeurIPS

Technical Skills

Programming & ML: Python, PyTorch, HuggingFace Ecosystem, DeepSpeed, vLLM
Miscellaneous: Unix/Linux, Docker, AWS (S3, SageMaker, EC2), Slurm, Kubernetes, Airflow

Languages

German (native), English (fluent, TOEFL iBT 116/120), Japanese (elementary), Danish (elementary)