Leon Weber-Genzel

leonweber@posteo.de
https://www.leonweber.me

Education

2018–current **PhD studies**, *Humboldt-Universität in Berlin*, Berlin

2016–2018 **MSc Computer Science**, *Humboldt-Universität in Berlin*, Berlin
Grade 1.0

2013–2016 **BSc Computer Science**, *Freie Universität*, Berlin, *Grade 1.4*

2010–2013 **BA Philosophy/English studies**, *University of Bamberg*, Bamberg, *Grade 1.3*

2000–2009 **Abitur (equivalent to A level)**, *Ehrenbürg-Gymnasium*, Forchheim, *Grade 1.8*

Professional Experience

Research

11/2022–current **Research Assistant**, *Human-in-the-Loop Natural Language Processing*, Ludwig-Maximilians-Universität
Researcher in the lab of Prof. Barbara Plank. Developing human-in-the-loop methods for Natural Language Processing.

11/2021–current **Open Source Research Contributor**, *Biomedical Language Modelling*, Huggingface BigScience
Core contributor in the Huggingface BigScience biomedical working group. Designed, organized and carried out hackathon with the goal of standardizing all biomedical NLP datasets with the help of roughly 50 participants.

11/2018–11/2022 **Research Assistant**, *Text Mining for Pathway Curation*, Humboldt-Universität zu Berlin
PhD studies on developing information extraction methods to assist biochemical pathway curation. Jointly supervised by Ulf Leser and Jana Wolf. Developed deep learning methods for extracting biochemical knowledge from scientific literature and to extend pathway models based on the extracted knowledge.

07/2018–10/2018 **Research Assistant**, *Trend Detection in Biological and Biomedical Patent Documents*, Humboldt-Universität zu Berlin
Joint project with Bayer R&D Information. Developed methodology and evaluation framework for trend detection in biomedical patent documents based on large-scale representation learning of patent texts and outlier detection. Unpublished

10/2017–07/2018 **MSc Thesis**, *NLProlog – Reasoning with Weak Unification for NLP*
Supervised by Pasquale Minverini (University College London) and Ulf Leser. Developed neuro-symbolic method for question answering that is capable of learning interpretable deduction rules from natural language data.

11/2017–02/2018 **Student Research Assistant**, *Friedrich-Alexander-Universität*, Erlangen-Nürnberg
Deep Learning for biomedical time-series segmentation. Supervised by Kilin Shi. Found architecture that performed on par with the state-of-the-art method without the need for extensive preprocessing and feature engineering.

04/2016–10/2018 **Student Research Assistant**, *Humboldt Universität zu Berlin, Knowledge Management in Bioinformatics Group*, Berlin
Machine Learning in Natural Language Processing and Genetics. Supervised by Maryam Habibi and Ulf Leser. Developed deep learning methods for Biomedical Named Entity Recoginition that surpassed state of the art performance. Built machine learning models for variant effect prediction in cooperation with Dominik Seelow from Charitè

02/2017 – 05/2017 **Research Intern**, *Max Dehlbrück Center, Computational Regulatory Genomics Group*, Berlin
Deep Learning for Regulatory Genomics. Supervised by Philipp Drewe-Boß. Evaluated various models integrating heterogeneous sources of information for polyA-site prediction which surpassed state-of-the-art models. Unpublished

02/2017 – 05/2017 **Contribution to Project**, *Free University*, Berlin
Automated Higher Order Logic Theorem Proving in Philosophy. Supervised by Christoph Benzmüller. Modelled and verified different (semi-)formal arguments in the debate about Gödels proof of god. Showed applicability of automated theorem proving to Philosophy

Teaching

10/2022–02/2023 **Exercise**, *Vertiefung in die Computerlinguistik*
Developed and led the exercise sessions for the lecture *Vertiefung in die Computerlinguistik* that focusses on Deep Learning for Natural Language Processing.

10/2020–03/2021 **Seminar**, *New Developments in Deep Learning*
Seminar on recent developments in deep learning research with a focus on graph neural networks and probabilistic methods.

02/2018 **Tutorial**, *Social Bioinformatics Deep Learning Workshop*, Berlin
Hands-on session on Deep Learning for Bioinformaticians

02/2014–03/2014 **Student Teaching Assistant**, *Freie Universität*, Berlin
Object-oriented programming and Java (block course)

Other

03/2015–12/2015 **Student Employee**, *Carmeq GmbH*, Berlin
Software Engineering in Java and Ruby

04/2014–03/2015 **Student Employee**, *Freie Universität, Physics Department*, Berlin
Software Engineering and System Administration

Awards & Competitions

2021 **First Place at Shared Task**, *DrugProt - BioCreative VII*, Shared task on extracting chemical-protein relations from the biomedical literature, First ranking team out of 30

2021 **Second Place at Shared Task**, *BioNLP - MEDIQA 2021*, Shared task on summarizing consumer health questions, Second ranking team out of 23

2019 **First Place at Shared Task**, *CLEF eHealth Multilingual Information Extraction*, Shared task on code assignment for German animal experiment summaries, First ranking team out of six

2018 **Award for Best Master's Thesis**, *Department of Computer Science, Humboldt-Universität zu Berlin*, NLProlog – Reasoning with Weak Unification for NLP, One out of two awarded theses

Languages

German native

English fluent

Computer Skills

Languages Python, Java

Machine Learning Frameworks PyTorch, huggingface transformers, scikit-learn, flair, Keras, Tensorflow

Other GNU/Linux, Docker, Git, LaTeX, Intellij IDEs, VIM, Microsoft Office

Community Service

Reviewer ACL, EMNLP, OUP Bioinformatics, Big Data, BioNLP, BMC Bioinformatics, BMC Supplements, CIKM, COINS, NSNJI, EDBT

Open source contributions flair, huggingface BigScience