Dr.
Steffen Remus

📍 Hamburg, Germany
📞 + 49 179 4730952
✉ sremus@pm.me
ⓘ Personal Profile
in LinkedIn
◯ GitHub
G Google Scholar
ⓘD ORC ID

Core Skills & Expertise

Machine Learning & AI:
NLP, Deep Learning, GenAI, Information Extraction, LLMs, RAG

Software Engineering:
IaS, SaaS, PaaS, CI/CD, DevOps, Containerization, Microservices, Webservices

Data Engineering:
SQL, NoSQL, Vector Databases, MapReduce, Cluster Management, DFS, ETL

Management & Leadership:
Research Collaborations, Teaching, Supervision, Project Management, Workshop Organization

Problem-Solving & Analysis:
Critical Thinking, Algorithmic Design

Technical Skills

Programming Languages:
Python, Java, JavaScript, Scala, C#, R, LaTeX

Data & ML Frameworks:
PyTorch, Spark, Hadoop, Scikit-Learn

Databases:
PostgreSQL, PgVector, Milvus, ChromaDB, Elasticsearch, Redis, ...

Cloud & DevOps:
AWS, GCP, Azure, Docker, Kubernetes, Linux, SSH

Languages & Interests

Languages:
German (native), English (fluent), Spanish & French (basic)

Interests:
🏃 Running
🌊 Kitesurfing
🚴 Road Biking
🖥 Technology
🏋 Problem-Solving

§ Professional Summary

Machine Learning and NLP expert with 10+ years of experience in ML algorithms, large-scale data engineering, and software development. Passionate about research, teaching, and consulting. Strong leadership skills with a proven track record in mentoring, project management, and fostering collaborations.

§ Professional Experience

Language Technology Group & HITeC e.V., University of Hamburg, Germany
since 07/2023 . . . . **Postdoctoral Research Scientist**
10/2016 to 06/2023 . . **Pre-doctoral Research Scientist**
▸ Conducted original research in NLP and AI with Prof. Chris Biemann
▸ Led and developed ML-driven projects, collaborating with industry partners
▸ Supervised 30+ theses, interns, and student researchers
▸ Maintained key infrastructure, including a Hadoop cluster and GPU servers
▸ Organized research workshops and served as a reviewer for 50+ issues
▸ *Key Achievements:* Successfully defended Ph.D. (2023)

Freelance Consulting Services, Germany
since 2016, occasional & part-time
▸ Provided expert guidance on NLP and ML projects for clients such as Glanos GmbH, T-Systems, and Merck KGaA

International Business Machines (IBM)
05/2015 to 11/2015 . . **Research Intern at the Zurich Research Lab, Rüschlikon, Switzerland**
▸ Developed systems for Named Entity Recognition and Relation Extraction
07/2014 to 10/2014 . . **Research Intern at the Thomas J. Watson Research Center, Yorktown Heights, NY, USA**
▸ Enhanced knowledge graphs using distributional semantics

Technical University of Darmstadt & DIPF (Germany)
04/2013 to 06/2016 . . **Researcher**
▸ Worked on Knowledge Discovery in Scientific Literature (§KDSL) projects

ConWeaver GmbH, Darmstadt, Germany
10/2012 to 05/2013 . . **Software Engineer (full time)**
2009 to 09/2012 . . . . . **Software Engineer (part-time)**
🔗 **full list**

§ Education

Dr.rer.nat. (Ph.D.) in Computer Science, University of Hamburg, Germany (2023)
Thesis: "Domain Defining Context: On Domain-Dependent Corpus Expansion and Contextualized Semantic Structuring"

M.Sc. in Computer Science, Technical University of Darmstadt, Germany (2012)
Thesis nominated for best thesis award at GSCL
🔗 **full list**

§ Research & Scientific Contributions

▸ **Research Projects:** JOIN-T2, WebAnno meets EXMARaLDA, BIMDanube, KDSL
▸ **Visiting Researcher:** Bar Ilan University (Israel), IIT Kharagpur (India)
▸ **Conference & Journal PC:** EMNLP, ACL, AAAI, TPAMI, TACL, ...
▸ **Workshops & Shared Tasks Organized:** GermEval 2020, GermEval 2019, Biomedical Information Management Workshops, IBM Watson workshops

§ Awards & Grants

▸ **2022:** Supervised a DAAD-funded student research project (Eur 10,000 grant)
▸ **2019:** DAAD-funded international research collaboration in India (Eur 3,000 grant)
▸ **2015:** Best mentoring award, "Algorithms of Language Technology" course
▸ **2013 – 2016:** KDSL Graduate School Fellowship

§ Public Work

▸ (Co-)authored over 30+ peer-reviewed papers. (🔗 full list)
▸ Contributed to 50+ open source project repositories on 3+ version control platforms (github, bitbucket, sourceforge, ...), in 7+ organizations of which I am maintainer/creator of 20+ projects.