Dr.
Steffen Remus

- Hamburg, Germany
- + 49 179 4730952
- sremus@pm.me
- Personal Profile
- LinkedIn
- GitHub
- Google Scholar
- ORC ID

Skillset

Core Skills

Communications,
Analysis & Algorithmic Thinking,
Team Work,
Creativity,
Problem Solving,
Research,
Critical Thinking,
Organizations,
Leadership,
Project management,
Self-management,
Presentations / Aesthetics

Knowledgable Concepts

Machine Learning, Deep Learning, Containerization, Generative AI, Information Extraction, Information Retrieval, Data Modeling, Data Structures, Large Language Models (LLMs), SQL, MapReduce, Artificial Intelligence (AI), (Computational) Linguistics, Data Science, Data Management, Software Architecture, Data Visualization, Software Development, Big Data, NLP (Natural Language Processing), NLU (Natural Language Understanding), Computational Linguistics, Knowledge Discovery, Knowledge Representation, Information Discovery, Text Mining, Data Mining,

Professional Profile / Summary

A highly dedicated research engineer with over 10 years of professional hands-on experience in:

- Developing machine learning algorithms and managing large-scale data engineering projects, with a primary focus on natural language processing.
- Software engineering, utilizing industry-standard tools and processes.
- Leading and managing individuals and teams, fostering individual growth and collaboration.
- Administering and maintaining server infrastructure.

I actively share my expertise through my current role as a (co-) lecturer and advisor, and as a freelance consultant for industry partners. As a researcher, I thrive in fast-paced environments and I am committed to staying at the forefront of advancements in machine learning and AI. I approach challenges with patience and a strong attention to detail, and I excel in bridging technical expertise with effective communication, ensuring strong teamwork and maintaining high levels of motivation. My leadership style is collaborative and adaptive, focusing on fostering a positive team environment, encouraging open communication, and empowering individuals to reach their full potential.

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**

Outline: Conducted original research under the guidance of Prof. Dr. Chris Biemann, participated in project-related events, offered consultancy services for industry partners, developed and delivered courses in machine learning, NLP, and software engineering, supervised theses, studies, student helpers and interns, maintained technical infrastructure and ensured its uptime.

Key Responsibilities:

- §Teaching & Supervision
- §Research Projects
- Server infrastructure maintenance: Responsible for maintaining key components of the group's server infrastructure, including the management of a Hadoop cluster that supports a distributed Elasticsearch cluster, two GPU servers, and several web servers. One of these web servers hosts demo applications as Docker containers.

Key Achievements:

- Successfully defended Ph.D. in June, 2023
- Organized four research workshops and served 50+ times as a reviewer
- Gained leadership qualifications during 19 teaching activities, supervision of over 30 theses, interns, and student helpers, providing academic guidance and mentorship at various levels in well organized regular SCRUM-style meetings.
- Contributed to over 30 small to medium sized software projects.
- (Co-)authored over 30 peer-reviewed §publications.

Freelance Consulting Services (Occasional & Part-Time)

occasional since 2016 . **Research Consultant (Ad hoc engagements)**

Outline: Provided project-based freelance consultancy services, specializing in NLP research and offering expert guidance on various related topics.

Key Responsibilities: Worked with various clients, including:

- Glanos GmbH (Munich, Germany)
- T-Systems (Darmstadt, Germany)
- Merck KGaA (Darmstadt, Germany)

International Business Machines (IBM)

05/2015 to 11/2015 . . **Research Intern at the Zurich Research Lab, Rüschlikon, Switzerland**

Supervisor: Dr. Abdel Labbi & Dr. Alfio Gliozzo

Outline: Implemented a system for named entity recognition, cross-sentence relation extraction, and event detection using distributional features, and managed an internal crowdsourcing task for relation annotation.

07/2014 to 10/2014 . . **Research Intern at the Thomas J. Watson Research Center, Yorktown Heights, NY, USA**

Supervisor: Dr. Alfio Gliozzo

Outline: Implemented, tested, and evaluated distributional semantic methods to enhance knowledge graphs for both open and closed domains.

Tech Stack

PyTorch, Docker, Elasticsearch, Lucene, Solr, Hadoop (Administration & Applications), MySQL, PostgreSQL, PgVector, ParadeDB, MongoDB, Apache Spark, REST, Websockets, Webservices, Microservices, Swagger, Redis, Large Language Models (LLMs), AWS (Amazon Web Services), LaTeX, . . .

Programming Languages

Python

Java

JavaScript

R

Scala

C#

Other

Research and Development (R&D), Qualitative & Quantitative Research Methodologies

Languages

German (native)

Englisch (fluent)

Spanish (basic)

French (basic)

Hobbies & Interests

🏃 Running

🌊 Kitesurfing

🚴 Biking (road)

🖥 Tech Stuff

🍃 Any kind of sports activity

👾 Solving Riddles

💡 Keenly interested to try new things

Dr.
Steffen
Remus

Technical University of Darmstadt, Germany & Deutsches Institut für Pädagogische Forschung (DIPF), Frankfurt (Main), Germany

04/2013 to 06/2016 . . **Researcher in the §KDSL Graduate School Program**

Outline: Conducted original research under the supervision of Prof. Dr. Chris Biemann in a team of 5 core research members and 5 research principles in the framework of the Knowledge Discovery in Scientific Literature (§KDSL) graduate school.

Key Responsibilities:

▸ Project Crawling & Semantic Structuring.

ConWeaver GmbH, Darmstadt, Germany

10/2012 to 05/2013 . . **Software Engineer (full time)**

2009 to 09/2012 . . . . . **Software Engineer (part-time student assistant position)**

Other

2009 to 2012...Tutor (part-time) at Technical University of Darmstadt, Darmstadt, Germany
2007 to 2008...Systems Administrator (part-time) at DeltaSquare GmbH, Frankfurt (Main), Germany
2002 to 2006...Support staff (part-time) at Greater Union Filmpalast, Frankfurt (Main), Germany
2000 to 2003...Systems Specialist (full-time) at Siemens AG, Frankfurt (Main), Germany
1997 to 2000...Apprenticeship (full-time) at Siemens AG, Frankfurt (Main), Germany

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", Advisor: Prof. Dr. Chris Biemann

M.Sc. in Computer Science – Technical University of Darmstadt, Germany – 2012
Thesis: "Automatically Identifying Lexical Chains by Means of Statistical Methods – A Knowledge-Free Approach" – Nominated for best thesis award at the German conference for computational linguists (GSCL), Advisor: Prof. Dr. Chris Biemann

B.Sc. in Computer Science – Technical University of Darmstadt, Germany – 2009

State Certified Engineer (Techniker) CS&N – Frankfurt (Main), Germany – 2005

Certified Apprenticeship (Ausbildung) ITSE – Siemens AG, Frankfurt (Main), Germany – 2000

High School – Brüder Grimm Realschule, Frankfurt (Main), Germany – 1997

Scientific Experience

Research Projects

03/2019 to 02/2022 . . **JOIN-T2: Joining Ontologies with Text**

06/2017 to 01/2019 . . **WebAnno meets EXMARaLDA**

01/2017 to 03/2019 . . **BIMDanube: Network of Excellence on Biomedical Information Management in the Danube Region**

04/2013 to 06/2016 . . **Knowledge Discovery in Scientific Literature (§KDSL)**

Visiting Researcher

04/2014 to 06/2014 . . **Bar Ilan University, Ramat Gan, Tel Aviv, Israel**

02/2019 to 04/2019 . . **Indian Institute of Technology (IIT) Kharagpur, Kharagpur, West Bengal, India**

Organizing Comittee

Organizer Shared Task on Classification and Regression of Cognitive and Motivational Style from Text (GermEval Task 1) held in conjunction with the SwissText & Konvens Joint Conference, 2020 (link)

Main Organizer Shared Task on Hierarchical Classification of Blurbs (GermEval Task 1) held in conjunction with the GSCL Conference, 2019 (link)

Main Organizer 2nd Workshop on Biomedical Information Management: Data-Driven Innovations held in Hamburg, Aug. 2018 (link)

Main Organizer 1st Workshop on Biomedical Information Management: Challenges and Open Problems held in Hamburg, Feb. 2018 (link)

Co-Organizer One Day Workshop on IBM Watson held in Darmstadt, Apr. 2014

Programm Comittee

I regularly serve as a voluntary reviewer for the venues listed below and I contributed to numerous other venues (not listed), with a total of over 50 reviewing activities to date.
Conferences: EMNLP, ACL, AAAI, CONLL, *SEM, NAACL, EACL, AACL-IJCNLP, Coling, LREC, GSCL/KONVENS, COLM; *Journals*: TPAMI, TACL, NLE; *Workshops*: TextGraphs, SemEval, WAC.

Teaching Experience

Curricula

Conducted 19 teaching activities including 3 Seminars, 7 practice labs and 2 lectures with lab exercises at the University of Hamburg, led 10 teaching activities at the Technical University of Darmstadt, including 2 seminars and 3 lecture exercise labs and participated 5× as a tutor.

Supervision

Provided dedicated supervision for 2 student helpers, 3 inters and 24 academic theses, including 10 Bachelor's theses, 11 Master's theses and 3 independent studies. Led regular one-on-one progress meetings on a weekly or bi-weekly basis, ensuring consistent academic development, timely completion, and overall project success.

Awards & Grants

2022 Mentor for a student research project (Förderung studentischer Forschungsgruppen), supervising a team of 4 students awarded with €10,000 to support a novel language learner application with eye-tracking on consumer hardware (iPads).

2019 Awarded over €3,000 by the German DAAD office to work in a top international research program in India.

2015 Best mentoring award for "Algorithms of Language Technology" exercise lab.

2013, 2014, 2016 Association of Computational Linguists (ACL) Travel Grants.

2013 – 2016 Knowledge Discovery in Scientific Literature (KDSL) Graduate School Fellowship from the Technical University of Darmstadt & German Institute for International Educational Research (DIPF) (link)

2013 Awarded a travel grant of €1,000 from Freunde der TU-Darmstadt e.V.

Selected Publications

1 **Remus** S, Wiedemann G, Anwar S, Petersen-Frey F, Muhie Yimam S, & Biemann C. (2022). More like this: Semantic retrieval with linguistic information. *Proceedings of the 18th Conference on Natural Language Processing (KONVENS 2022)*, 156–166. pdf.

2 Wiedemann G, **Remus** S, Chawla A, & Biemann C. (2019). Does bert make any sense? interpretable word sense disambiguation with contextualized embeddings. *Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)*, 161–170. pdf, bib.

3 **Remus** S, Kaufmann M, Ballweg K, von Landesberger T, & Biemann C. (2017). Storyfinder: Personalized knowledge base construction and management by browsing the web. *CIKM '17: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2519–2522. pdf, poster, web/git.

4 **Remus** S, & Biemann C. (2016). Domain-specific corpus expansion with focused webcrawling. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, 23–28. pdf, bib, web/git.

Find the full list of publications here.

Dr.
Steffen
Remus