

Anja Reusch

Curriculum Vitæ

Research Interests

- Natural Language Processing and Generation
- Neural Information Retrieval
- Interpretability, Explainability, Multilinguality

Education

- 05/2024 **Doktoringenieurin (Dr.-Ing.)**, *Technische Universität Dresden, Germany*
Doctor of Engineering in Computer Science, thesis topic 'Transformer-Encoders for Mathematical Answer Retrieval', Supervisor: Prof. Dr.-Ing. Wolfgang Lehner, *passed with award nomination*
- 10/2013 - **Diplom Computer Science**, *Technische Universität Dresden, Germany*
- 09/2019 Integrated Bachelor and Master course, thesis topic 'A Semi- and Unsupervised Approach for Grammatical Error Correction', final grade: *1.0 (summa cum laude)*
- 08/2018 - **Research Intern**, *National University of Singapore, NLP Group, Singapore*
- 12/2018 Research project regarding semi- and unsupervised methods for Grammatical Error Correction with a convolutional Encoder-Decoder Network and Variational Auto-Encoders.
- 10/2015 - **Junior Year Program in English**, *Tohoku University, Sendai, Japan*
- 03/2016 Exchange semester, interdisciplinary courses in science and Japanese language, which were held in English, Individual Research Training

Experience

Vocational

- Since **Doctoral Researcher**, *Technion - IIT, NLP Group, Haifa, Israel*
- 08/2024 ◦ Azrieli International Postdoctoral Fellow
◦ Research in Interpretability and Information Retrieval
- 11/2019 - **Doctoral Researcher**, *Technische Universität Dresden, Database Research Group*
- 03/2024 ◦ Research concerning neural information retrieval for scientific documents including \LaTeX notation and programming code, pre-training methods for transformer encoder models, interpretability of transformer encoders.
◦ Leading a team of student assistants as part of the Software Campus Leadership Program
◦ Visits at Technion IIT, Israel (five months), and NM-AIST, Tanzania (four weeks).
- 12/2017 - **R&D Student Assistant**, *interface projects GmbH, Dresden, Germany*
- 08/2018 and ◦ Incorporation of Chinese and Japanese language in the company's enterprise search system
- 01/2019 - ◦ Evaluation of several strategies to improve the enterprise search engine using PyTorch,
- 10/2019 gensim and fastText in Python

10/2016 - **Student Assistant – Research**, *Technische Universität Dresden, Chair for Foundations of Programming*, Dresden

- Language Processing using Moses Decoder to support the chair's research

Miscellaneous

03/2017 - **Teaching Assistant**, *Richard English*, Shaoxing, China

- Supporting children of grade 2 to 9 in public and private schools to learn and apply English, preparing and evaluating materials as well as various class room activities

Languages

Native German

Fluent English

Intermediate Mandarin Chinese

HSK 4 - 5

Basic Japanese, Modern Hebrew

Scholarships and Awards

- **Azrieli International Postdoctoral Fellowship** (2024 - 2026)
- **Ali Kaufmann Postdoctoral Fellowship** (since 2024)
- **SIGIR Student Travel Grant** (2022)
- **ARQMath-2 Best Paper Award** (2021)
- **Enno Heidebroek Award** (2020) – Awarded for outstanding performance in the course of study as one of the best graduates of the computer science program
- **Saxonia Women Award** (2019) – Awarded for outstanding achievements in studies, as well as social commitment
- **Business Analyst Nanodegree Program** – Scholarship by Bertelsmann SE & Co. KGaA (2018 - 2019), Application of advanced analytics and statistics in several practical projects
- **Deutschlandstipendium** – Public-Private scholarship, funded by the Federal Ministry of Education and Research, and IBM, Carl Zeiss AG, T-Systems Multimedia Solutions GmbH and SAP SE (2014 - 2019)
- **PROMOS** – Scholarship to support the mobility of students by Technische Universität Dresden (2015)
- **JASSO** – Japan Student Services Organization by the Ministry of Education, Culture, Sports, Science and Technology Japan (2015 - 2016)

Memberships and Funding

- **Software Campus Project Funding** (2021 - 2023), Project Funding (100k€) by the German Ministry of Education and Research, PI for SELMA: Semantic Code Search using Deep Learning
- Associate Member of the DFG-funded Research Training Group "Role-based Software Infrastructures for continuous-context-sensitive Systems" (2020 - 2022)
- Member of the DFG-funded Cluster of Excellence "Physics of Life" (2019 - 2024)

Mentoring and Teaching Experience

- **Mentoring** of 7 Bachelor Theses, 11 Master/Diploma Theses, 15 student theses
- **Information Retrieval** (2020/21, 2021/22, 2022/23, 2023/24) lecturer (Preprocessing of text, practical implementation of search systems, neural network-based search, word embeddings), and (2020/21, 2021/22, 2022/23) teaching assistant for exercises, implementing Jupyter-notebook based practical exercises, online quizzes (German)
- **Scalable Data Engineering** (2021/22, 2022/23, 2023/24) lecture about text processing, preprocessing, Neural Network based approaches for text processing, word embeddings, and (2023/24) teaching assistant for 130 registered students (English)
- **Guest Teaching** at NM-AIST, Arusha, Tanzania (four weeks, 2023) lecturer on Information Retrieval, including exercises and practical project (English)
- **Information Retrieval Lab** (2022/23, 2023, 2024), mentoring of various group projects including "Clickbait Spoiling" as part of the SemEval Shared Task Clickbait, and Open Web Search components for WOWS'24 (German)
- **Scalable Data Management** (2020, 2021, 2022) teaching assistant (German)
- **Database Foundations** (2021) transition of teaching materials to online-course, creation of practical group exercises (German)
- **Programming** (2015, 2016) teaching assistant for first year students in functional programming, introduction to lambda calculus, program verification, and unification algorithm (German)

Invited Talks

- **MPI Saarbrücken - Group for Databases and Information Systems** (02/2024) Transformer-Encoders for Math Answer Retrieval
- **Webis Group** (10/2023) - Transformer-Encoders for Math Answer Retrieval
- **SAP SE** (10/2022) - Neural Information Retrieval
- **Technion NLP Group** (06/2022) - Pre-Training for Math Answer Retrieval

Academic Service

- **CIKM** - PC member, reviewing short papers, 2024
- **SIGIR** - PC member, reviewing full and short papers, 2023, 2024
- **ARQMath@CLEF** - PC member, reviewer 2021, 2022
- **MathNLP 2022** - Panel Discussion Participant

Publications

- [1] Linda Erben, Maria Hampel, Malte-Christian Kuns, Vincent Melisch, Per Natzschka, Wilhelm Pertsch, Lina Razouk, Reiner Stolle, Robert Thomas Thoss, Tuan Giang Trinh, Julius Gonsior, and Anja Reusch. Assembling four open web search components. In *1st Workshop for Open Web Search*, 2024.
- [2] Anja Reusch, Guilherme C. Lopes, Wilhelm Pertsch, Hannes Ueck, Julius Gonsior, and Wolfgang Lehner. Selma: A semantic local code search platform. In *46th European Conference on Information Retrieval*, 2024.
- [3] Anja Reusch, Julius Gonsior, Claudio Hartmann, and Wolfgang Lehner. Investigating the usage of formulae in mathematical answer retrieval. In *46th European Conference on Information Retrieval*, 2024.
- [4] Simon Birkenheuer, Jonathan Drechsel, Paul Justen, Jimmy Phlmann, Julius Gonsior, and Anja Reusch. Sabrina spellman at semeval-2023 task 5: Discover the shocking truth behind this composite approach to clickbait spoiling! In *Proceedings of the The 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 969–977, 2023.
- [5] Christian Falkenberg, Erik Schönwälder, Tom Rietzke, Chris-Andris Görner, Robert Walther, Julius Gonsior, and Anja Reusch. Mr-fosdick at semeval-2023 task 5: Comparing dataset expansion techniques for non-transformer and transformer models: Improving model performance through data augmentation. In *Proceedings of the The 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 88–93, 2023.
- [6] Julius Gonsior, Christian Falkenberg, Silvio Magino, Anja Reusch, Claudio Hartmann, Maik Thiele, and Wolfgang Lehner. Comparing and improving active learning uncertainty measures for transformer models. In *European Conference on Advances in Databases and Information Systems*, pages 119–132. Springer, 2023.
- [7] Anja Reusch and Wolfgang Lehner. Extracting operator trees from model embeddings. In *Proceedings of the 1st Workshop on Mathematical Natural Language Processing (MathNLP)*, pages 40–50, Abu Dhabi, United Arab Emirates (Hybrid), December 2022. Association for Computational Linguistics.
- [8] Anja Reusch. Pre-Training for Mathematics-Aware Retrieval. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 3496–3496, 2022.
- [9] Anja Reusch, Maik Thiele, and Wolfgang Lehner. Transformer-Encoder and Decoder Models for Questions on Math. *Proceedings of the Working Notes of CLEF 2022*, pages 5–8, 2022.
- [10] Anja Reusch, Maik Thiele, and Wolfgang Lehner. Transformer-Encoder-Based Mathematical Information Retrieval. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 175–189. Springer, 2022.

- [11] Anja Reusch, Maik Thiele, and Wolfgang Lehner. An ALBERT-based Similarity Measure for Mathematical Answer Retrieval. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1593–1597, 2021.
- [12] Anja Reusch, Maik Thiele, and Wolfgang Lehner. Teaching BERT Mathematics. In *2nd International Workshop on Natural Language Processing for Scientific Texts SciNLP 2021*, 2021.
- [13] Anja Reusch, Maik Thiele, and Wolfgang Lehner. TU_DBS in the ARQMath Lab 2021, CLEF. In *CLEF (Working Notes)*, pages 107–124, 2021.
- [14] Anja Reusch, Alexander Weber, Maik Thiele, and Wolfgang Lehner. RecipeGM: A Hierarchical Recipe Generation Model. In *2021 IEEE 37th International Conference on Data Engineering Workshops (ICDEW)*, pages 24–29. IEEE, 2021.