Leonard Traeger

leotraeg.github.io/me/
leonard.traeger@th-koeln.de
Cologne, Germany

Education

University of Maryland, Baltimore County (UMBC) — Baltimore, USA
Ph.D. in Information Systems: Collaborative Synthesis of Data Systems — *Aug. 2022 – May 2026*

Cologne University of Applied Sciences (Cologne UAS) — Cologne, Germany
M.Sc. in Computer Science / Information Systems; grade 1.2/5 equiv. GPA: 3.8/4 — *Mar. 2019 – Apr. 2022*

Macquarie University (MU) — Sydney, Australia
Exchange semester in Information Systems; grades HD, HD, C equiv. GPA: 3.2/4 — *June. 2017 – Dec. 2017*

Cologne University of Applied Sciences (Cologne UAS) — Gummersbach, Germany
B.Sc. in Business Information Systems; grade 1.7/5 equiv. GPA: 3.3/4 — *Aug. 2015 – Feb. 2019*

Professional Experience

Big Data Analytics Department at Cologne UAS — Cologne, Germany
Research Associate (On-site) — *Oct. 2023 - Present*

- Research on Data Management using Machine Learning and Language Models
- Summer 2025, Winter 2025/2026, Summer 2026: Teaching Assistant for Databases I and II
- Supervision of undergraduate theses and projects

Dortmund University of Applied Sciences and Arts — Dortmund, Germany
Lecturer (Hybrid) — *Oct. 2023 - Apr. 2025*

- Fall 2023 and Fall 2024: Programming in Data Science with Python and R (Programming II Syllabus)

Information Systems Department at UMBC — Baltimore, USA
Graduate Teaching and Research Assistant (Hybrid) — *Aug. 2022 - Jan. 2026*

- Spring 2024: Semantic Mapper for Application Integration (Technology Catalyst Fund)
- Fall 2022, Spring 2023, Fall 2024, Spring 2025, Fall 2025: Laboratories for Extended Databases (IS420 and IS620)
- Fall 2022: Guest lecture about Inteplato in Systems and Information Integration (IS722)

IT Consultancy (BWI GmbH) for Bundeswehr (German Armed Forces) — Bonn, Germany
Data Warehouse Application Consultant (On-site) — *Mar. 2019 - Jul. 2022*

- Developed data pipelines for materials, missions, and unit-critical dashboards on BW/4HANA for >300K users
- Designed role authorization, testing routines, training materials, and workshop leader of >20 participants

Awards and Services

VLDB Summer School (2023) awarded with **Best Poster** in Cluj, Romania

UMBC Graduate Student Association Conference Travel Grant for CSCI (2022) in Las Vegas, USA

UMBC Graduate Student Representative for Tenure and Promotion Committee (2022) in Baltimore, USA

Oracle Industry Speaker at DOAG (2022) "Machine Learning for Data Integration" in Duesseldorf, Germany

Fulbright Germany Graduate Scholarship at UMBC (2021) in Baltimore, USA

Reaccreditation Auditor for "M.Sc. Digital Sciences" at Cologne UAS (2019 - 2022) in Cologne, Germany

Foerderverein Gummersbach Study Abroad Funds at MU (2017) in Sydney, Australia

Programming Skills

SQL, Python (numpy, pandas, matplotlib, sklearn, keras), R, Javascript (npm, vite, react), MATLAB, ABAP

Publications

Khwaileh, Enas; **Traeger, Leonard**; Albert, Gatt; Yannis, Velegrakis. TABERTA: A Language Model for Dataset Discovery. 30th International Conference on Extending Database Technology (EDBT 2027). DOI:10.48786/edbt.2027.02 HuggingFace

Traeger, Leonard; Khwaileh, Enas; Behrend, Andreas; Karabatis, George. RACT: Retrieval Augmented Column-Table Learning for Multi-Table Schema Matching. Technical Report (ArXiv 2026). https://arxiv.org/abs/2606.07843 GitHub

Traeger, Leonard. Collaborative Synthesis of Data Systems: Automated, Contextual, and Streamlined Approaches Using Language Models. Dissertation (University of Maryland, Baltimore County 2026). ISBN 9798244895346

Traeger, Leonard; Behrend, Andreas; Karabatis, George. Scoper: Streamline Linkable Schemas for Matching. IEEE 42st International Conference on Data Engineering (ICDE 2026). DOI:10.1109/ICDE65706.2026.00319 GitHub YouTube Demo Application

Traeger, Leonard; Behrend, Andreas; Karabatis, George. Collaborative Scoping: Self-Supervised Linkability Assessment for Schema Matching. 29th International Conference on Extending Database Technology (EDBT 2026). DOI:10.48786/edbt.2026.03 GitHub

Traeger, Leonard; Behrend, Andreas; Karabatis, George. Collective Scoping: Streamlining Entity Sets Towards Efficient and Effective Entity Linkages. Springer Nature Computer Science Volume 6 (SNCS 2025). DOI:10.1007/s42979-025-03734-7 GitHub

Traeger, Leonard; Behrend, Andreas; Karabatis, George. SEALM: Semantically Enriched Attributes with Language Models for Linkage Recommendation. 27th International Conference on Enterprise Information Systems (ICEIS 2025). DOI:10.5220/0013217700003929 GitHub

Traeger, Leonard; Behrend, Andreas; Karabatis, George. Scoping: Towards Streamlined Entity Collections for Multi-Sourced Entity Resolution with Self-Supervised Agents". 26th International Conference on Enterprise Information Systems (ICEIS 2024). **Best Paper Award** DOI:10.5220/0012607500003690 GitHub

Traeger, Leonard; Omar, Faraque. Autonomous Object Detection and Landing with Parrot Rolling Spider in Matlab Simulink". IEEE International Conference on Embedded Systems, Cyber-physical Systems & Applications (ESCS 2023). DOI 10.1109/CSCE60160.2023.00160 GitHub YouTube

Traeger, Leonard; Behrend, Andreas; Karabatis, George. Inteplato: Generating Mappings of Heterogeneous Relational Schemas using Unsupervised Learning". IEEE International Conference on Computational Science and Computational Intelligence (CSCI 2022). DOI:10.1109/CSCI58124.2022.00083 GitHub

Patents

Karabatis, George; Behrend, Andreas; **Traeger, Leonard**. Systems and methods for organizing data of different local data schemas based on similarity rankings, scoring, and signatures. US Patent App. 18/600,828 (2026).