

# FABIAN WENZ

## Master of Mathematics in Data Science

✉ [fabian@wenz.org](mailto:fabian@wenz.org)  
📍 Munich, Germany  
🌐 [github.com/fabian-wenz](https://github.com/fabian-wenz)

☎ 0049-176-49387111  
🌐 [fabian-wenz.github.io](https://fabian-wenz.github.io)

✉ Bunzlauer Straße 17, 80992 Munich  
🌐 [fabian-wenz](https://fabian-wenz.github.io)



## EXPERIENCE

### PhD in database systems

#### Technical University of Munich (TUM)

📅 Sept 2025 – Mar 2026

- Research on AI-enhanced data placement
- Continued development and presentation of BENCHPRESS for text-to-SQL benchmark curation.
- Contributed to Athena-tools, a plug-and-play advisor for Retrieval-Augmented Generation.
- Teaching Assistant for Advanced Algorithms for Competitive Programming.
- Designed and delivered lectures on RAG and HNSW indexing.

### Research Assistant (under scholarship)

#### Technical University of Munich (TUM)

📅 April 2024 – Sept 2025

- Contributed to the development of *Athena-tools*, a plug-and-play advisor for Retrieval-Augmented Generation using vector databases.
- Co-authored a publication accepted at IISWC 2025, focusing on practical RAG pipeline optimization and evaluation.
- Supported benchmarking and tool development for integration with vector database backends.

### Research Assistant (under scholarship)

#### Massachusetts Institute of Technology (MIT)

📅 Oct 2024 – Sept 2025

- Contributed to LLM research on enterprise text-to-SQL tasks under the guidance of Prof. Michael Stonebraker.
- Co-authored a paper on data integration and benchmark curation for CIDR 2025.
- Developed human-in-the-loop annotation systems for Text-to-SQL benchmark generation.

### C++ Database Engineer (part-time)

#### Celonis

📅 Sep 2022 – Mar 2024

- Develop and implement new, innovative features and algorithms for celonis' query engine in C++, Java and Python
- Increase the reliability and performance of existing components of celonis' query engine

### Data Scientist – Fixed Income (part-time)

#### Amundi S.A.

📅 Jul 2021 – Mar 2022

- visualization of fond-mandates

## STRENGTHS

Persistent

Logical-thinking

Detail-driven

Algorithms

Optimization

Efficiency

Data Structures

AI

## LANGUAGES

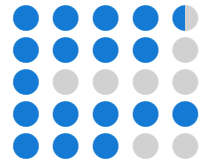
English (C1)

Italian (B2)

Chinese (A1.1)

German (C2)

Latin (Big Latinum)



## EDUCATION

### Master Thesis: Leveraging LLMs for Querying Large Datawarehouses: A New Paradigm

#### Massachusetts Institute of Technology

📅 April 2024 – Oct 2024

### M.Sc. in Mathematics in Data Science

#### Technical University of Munich

📅 Oct 2021 – Mar 2025

### B.Sc. in Computer Science

#### Technical University of Munich

📅 Oct 2018 – Sep 2021

Thesis: An Experimental Evaluation of Heuristics for Constructing Hierarchical Cut Sparsifiers

### B.Sc. in Mathematics

#### Technical University of Munich

📅 Oct 2014 – Sep 2018

Thesis: Mathematical Models of Biological Processes in Nephrons

### University Entrance Qualification

#### Louise-Schoeder-Gymnasium

📅 Sep 2006 – Jul 2014

- implementation of an automated monthly report-generator interacting with large data-sets in Python and MySQL

## Intern & Working Student as Data Scientist in Tax, Technology and Transformation

### Ernst & Young GmbH

📅 Nov 2019 – Jul 2020

- implementation of a web-crawler in Python in the pharmaceutic field for identifying newly added or changed regulations
- developing an automatic email-newsletter, web-based applications in the tax field for calculations and visualizations in Power Bi and a Desktop-application in Python in sharebased salary

## Student Assistant

### Technical University of Munich

📅 Feb 2019 – Jul 2019, Sep-Oct 2018, Sep-Oct 2020, Apr-Sept 2022

- tutor advanced mathematics, introductory mathematics,  $\text{\LaTeX}$ -scripting, and exam-corrector

## Internship Software development

### BMW Group

📅 Mar 2018 – Sep 2018

- implementation of an optimization algorithms for generating realistic load cases for virtual concept-evaluations of the manual montage of flexible car-components

## PUBLICATIONS

(Presented by N. Liang)

### Tools-Athena: A Plug-and-Play Advisor for Retrieval-Augmented Generation using VectorDBs

📅 12-14 Oct 2025

📍 Irvine, IISWC 2025

📄 <https://ieeexplore.ieee.org/document/11241995>

(Presented by F. Wenz)

### BENCHPRESS: An Annotation System for Rapid Text-to-SQL Benchmark Curation

📅 17-19 Jun 2025

📍 Montreal, SEMLA 2025.

📅 18-21 Jan 2026

📍 Santa Cruz, CIDR 2026.

📄 <https://vldb.org/cidrrdb/papers/2026/p16-wenz.pdf>

(Presented by F. Wenz)

### BEAVER: An Enterprise Benchmark for Text-to-SQL

📅 3 Sep 2025

📍 Vienna, ACL 2025.

📄 [arxiv.org/abs/2409.02038](https://arxiv.org/abs/2409.02038)

(Presented by F. Wenz)

### Mind the Data Gap: Bridging LLMs to Enterprise Data Integration

📅 19-22 Jan 2025

📍 Amsterdam, CIDR 2025.

📄 [vldb.org/cidrrdb/papers/2025/p34-kayali.pdf](https://vldb.org/cidrrdb/papers/2025/p34-kayali.pdf)

## COMPUTER-KNOWLEDGE

C++



Haskell,  
Java, MySQL, Python



Matlab, Git, Linux,  $\text{\LaTeX}$



Lua, Adobe  
(Photoshop, Illustrator)



Microsoft Office  
(Word, Excel, PowerPoint)



PowerBi



## REFEREES

**Prof. Jana Giceva**

@ database system lab (TUM)

✉ available upon request

**Dr. Cagatay Demiralp**

@ Computer Science and Artificial Intelligence Laboratory (MIT)  
+ Amzon Web Services & AI

✉ available upon request

**Dr. Nesime Tatbul**

@ Computer Science and Artificial Intelligence Laboratory (MIT)  
+ Intel Labs

✉ available upon request

**Prof. Michael Stonebtkker**

@ Computer Science and Artificial Intelligence Laboratory (MIT)

✉ available upon request

**Felix Weiß**

@ Celonis

✉ available upon request

**Tugrul Kolad**

@ Amundi

✉ available upon request

**David Lappe**

@ (former) Ernst & Young GmbH

✉ available upon request

**Bernhard Wolf**

@ BMW Group

✉ available upon request

**PD. Andreas Johann**

@ Chair for Advanced Mathematics and Analytical Mechanics (TUM)

✉ available upon request