Johannes Michael Tauscher

Data Scientist – Applied AI & Production Systems

Data Scientist and Applied AI Engineer grounded in statistics and optimization, building reliable ML/LLM systems from framing to production.

j-tauscher@hotmail.com • linkedin.com/in/johannes-tauscher • github.com/thejonnyt • thejonnyt.github.io

PROFESSIONAL EXPERIENCE

Co-Founder, Technical Lead (Early Stage)
Nov 2024 – Dec 2025

Raumdeuter GbR
Leipzig, Germany

Building an AI-driven communication and engagement platform for members and fans of clubs and associations to enable participation and engagement — a way to voice your opinion. I led the technical development: from conceptualization and architecture to implementation of the product as a customer-deployed MVP with the help of a 4-person development team and coordinated external IT partners.

- Delivered a customer-deployed MVP $0 \rightarrow 1$ in under 12 months on a €30k budget, targeting capital-efficient scale.
- Led product ideation workshops to align vision and problem-solution fit roadmap.
- Architected ML/NLP pipelines (RAG, sentiment, topic clustering) as core AI/ML solution capabilities.
- Owned infrastructure and DevOps delivery, cutting release friction and improving onboarding for interns, contractors, and external partners.

Technologies: Docker, CI/CD, Bash, Python, NLP/ML, LLMs, SpaCy, FastAPI, gRPC, Qdrant, RAG, pytest

Mathematics Teacher (Substitute)
Nov 2022 – Jun 2023

Rahn Education (Private School)
Leipzig, Germany

Responsible for the supervision of talented pupils (groups of 4 to 8)

- Prepared customized university-level material to challenge and spark interest in mathematics.
- Initiated a pupil-led lecture series about Artificial Intelligence and its use cases.
- Taught basic principles of the mathematics behind 3D printing and coding.

Technologies: PowerPoint, Algebra

Student Assistant
May 2019 – Jan 2021

Child and Adolescent Psychiatry, Philipps University Marburg
Marburg, Germany

Responsible for the validation and reproduction of earlier results of research as well as for programming Machine Learning algorithms in R to solve classification problems in the field of ASD diagnosis.

- Hyperparameter tuning for decision trees/random forests improved diagnostic performance by 10–15 percentage points; results supported 3 peer-reviewed publications (70 citations).
- Worked in a highly interdisciplinary manner with psychologists and medical doctors to understand the clinical background and implications of the data + explain the statistical methods to non-data scientists.
- Commended for a reliable, way of conveying statistical content; invited to stay on as a scientific researcher.

Technologies: Statistics, Machine Learning, Random Forests, Decision Trees, R, ggplot2, tidyverse

Intern, Data Analyst   Aug 2019 – Oct 2019

Westphalia DataLab GmbH   Münster, Germany

As a Data Analyst intern at Westphalia DataLab GmbH, I worked on a medical-diagnostics project combining 3D imaging and clinical data; developed and improved an experiment-tracking R/Shiny frontend and performed clinical-data analysis with Python and R.

- Enhanced experiment-tracking R/Shiny frontend accessing data in MongoDB and SQL to improve researcher workflow.
- Analyzed 3D imaging + clinical datasets in Python/R to surface diagnostic insights.
- Internship evaluation rated performance as outstanding.

Technologies: Python, R, R Shiny, tidyverse, data.table, PyTorch, Convolutional Neural Networks, MongoDB

EDUCATION

M.Sc Data Science, Faculty of Mathematics and Computer Science   Apr 2021 – Oct 2024

Leipzig University   Leipzig, Germany

- Overall Score: 1.7, personal focus on advanced statistics, Machine Learning, data privacy, and LLM applications.
- Master's Thesis (1.1): Neural Machine Translation with Transformers - Leveraging the Pivot Technique for Low-Resource Language Pairs

B.Sc. Data Science, Faculty of Mathematics and Computer Science   Oct 2017 – Mar 2021

Philipps University Marburg   Marburg, Germany

- Bachelor's Thesis (1.7): Multidimensional Data Exploration and Visualization of Membrane Proteins Attributes
- Curriculum focuses on mathematics, statistics, computer science and programming.

B.A. Marketing and Technical Business Administration   Mar 2012 – Oct 2017

HAW Hamburg   Hamburg, Germany

- Bachelor's Thesis (1.7): Use of Linear Discriminant Analysis as an alternative to tree analysis - methodological evaluation and exemplary implementation using a survey on the success of movies.
- Curriculum focuses on marketing, engineering and business administration.

TECHNICAL SKILLS

Programming Languages: Python, Bash, R, Java, JavaScript, TypeScript, SQL

Analysis, ML & AI: LLMs, NLP, Machine Learning, R Shiny, SPSS, Statistics, SpaCy, RAG, Algebra, Decision Trees, Random Forests, PyTorch, pandas, NumPy, scikit-learn, SentencePiece, OpenNMT-py, Convolutional Neural Networks, matplotlib, ggplot2, tidyverse, data.table

Backend & DevOps: noSQL, Docker, gRPC, MongoDB, MariaDB, PostgreSQL, Qdrant, nginx, CI/CD, FastAPI, Docker Compose

Web & Creative: Next.js, React, Adobe Photoshop, HTML, CSS

Tools & Systems: Linux, Markdown, git, GitHub, LaTeX, Excel, PowerPoint, MS Office, MS Teams, Miro, Business Model Canvas, Lean Startup, VSCode, pytest, Protocol Buffers, Docker Hub, gcode