Daniel Fridljand

Munich, Germany danielfridljand.de daniel-fridljand FridljDa

Summary

Software consultant and data scientist with deep expertise in AI/ML systems, statistical methodology, and production software engineering. I combine rigorous mathematical foundations (M.Sc. Mathematics, Yale Exchange) with hands-on experience building LLM-powered applications, data pipelines, and enterprise systems. Published first-author in Nature Medicine, developed novel statistical algorithms at EMBL, and currently architect end-to-end AI solutions at TNG Technology Consulting.

Software Development Experience

Software Consultant - Applied AI *Munich, Germany*
TNG Technology Consulting *Dec 2025 – present*

- **Project: AI-Powered Customer Support Automation**
  - Sole developer architecting end-to-end automation system for cinema ticketing SaaS platform, processing 2000+ real customer tickets from data engineering through intent classification to automated resolution execution
- **Data Engineering**: Built reproducible 22-stage Snakemake pipeline with scatter-gather parallelism, GDPR-compliant PII detection, balanced sampling, and automated LLM labeling for training datasets
- **REST API & Architecture**: Designed 20+ REST endpoints following OpenAPI 3.1 specification with Hexagonal Architecture (Ports & Adapters), ensuring clean separation between business logic and infrastructure
- **LLM Solution Design**: Engineered hybrid automation for 21 customer intent categories with deterministic and agentic resolution strategies, using 77 Microsoft Prompty templates and DSPy prompt optimization
- **Workflow Orchestration**: Implemented durable workflow system using Temporal with 5-stage pipeline including human-in-the-loop approval via Signals, ensuring reliable execution with automatic retries
- **Observability & Testing**: Integrated Langfuse for comprehensive AI monitoring with 5 evaluation types and 8 score metrics; built 97-file test suite covering unit, integration, end-to-end, and AI performance evaluation
- **AI-Assisted Development**: Pioneered spec-driven development workflow with custom Cursor IDE skills for trace analysis and prompt refinement; managed 20+ feature specs with version-controlled project constitution
- **Tech Stack**: FastAPI, Streamlit, Chainlit, Python 3.13+, PydanticAI, Temporal, Langfuse, Snakemake, Docker
- **Project: LLM Document Validation**
  - Core developer in 3-person team building automated proposal review service with .docx processing, evaluation framework using Pydantic models, and Langfuse observability with Grafana dashboards

Software Consultant - Enterprise Modernization *Munich, Germany*
TNG Technology Consulting *Dec 2024 – Dec 2025*

- **DevOps Excellence**: Engineered critical pipeline optimization replacing legacy custom dependency solution with Artifactory proxy, reducing pipeline runtime from 8 hours to 30 seconds and eliminating development bottlenecks
- **DevSecOps Leadership**: Established proactive security culture by integrating OWASP scans and deploying Grafana dashboards for CVE monitoring, providing management visibility and reducing application risk
- **Enterprise Modernization**: Drove Java 8→17 and JBoss→WildFly migration for global supply chain management application, stabilizing bidding and ordering services for 24/7 availability and international logistics partners
- **Containerization & Security**: Architected fully containerized development and build environments using Docker and Podman with daemonless, rootless execution for enhanced security in shared CI environments
- **Agile Leadership**: Selected for Program Increment (PI) planning sessions, contributing to strategic planning and cross-team coordination while enforcing two-week sprint discipline with continuous integration
- **Tech Stack**: React, TypeScript, Java 8/17, Spring Framework, JBoss/WildFly, Oracle DB, Gradle, Docker/Podman, Jenkins, SonarQube, Jfrog Artifactory

Data Science Experience

Research Data Analyst - Computational Oncology *Basel, Switzerland*
ETH Zürich - Department of Biosystems Science and Engineering *Feb 2024 – Sept 2024*

- **Novel Statistical Methods**: Developed Bayesian non-parametric methods for mutational signature estimation, extending Hierarchical Dirichlet Process models to incorporate phylogenetic structures for cancer evolution
- **Single-Cell Genomics**: Analyzed whole-exome sequencing data from Tumor Profiler Study (187 cells, 10 melanoma tumors), successfully identifying eight latent mutational signatures capturing evolutionary patterns
- **Mathematical Innovation**: Implemented hierarchical dependency structure where signature distributions for child nodes are drawn from parent distributions, mathematically enforcing biological inheritance patterns
- **Teaching**: Delivered lecture on Statistical Models in Computational Biology covering hidden Markov models, EM algorithm, and Variational inference

Research Data Analyst - Environmental Epidemiology *Palo Alto, USA*
Stanford University School of Medicine *July 2023 – Dec 2023*

- **Nature Medicine First-Author**: Led end-to-end statistical analysis resulting in first-author publication in Nature Medicine (2024), quantifying air pollution's contribution to racial and socioeconomic mortality gaps
- **Big Data Engineering**: Processed 63+ million death records (1990-2016) across 3,000+ US counties, harmonizing satellite pollution data, census demographics, and CDC vital statistics with multi-scale geographic integration
- **Causal Inference**: Implemented causal models to estimate Attributable Fraction of mortality, revealing that over 50% of Black-White mortality difference is attributable to environmental factors
- **Open Science**: Published complete reproducible analysis pipeline on GitHub (https://github.com/FridljDa/pm25_inequality) and deposited national-level estimates in Zenodo following FAIR principles
- **Interactive R Shiny Application**: Developed analytical web application enabling co-authors to explore 17-dimensional data interactively, visualizing non-linear relationships and detecting outliers

Research Data Analyst - Statistical Genomics (Master's Thesis) *Heidelberg, Germany*
European Molecular Biology Laboratory (EMBL) *Oct 2021 – May 2022*

- **IHW-Forest Algorithm**: Developed novel statistical procedure for multiple testing correction in genomics using Random Forests for hypothesis weighting, increasing discovery power by >30% in 16 billion genetic tests
- **Machine Learning Innovation**: Designed Independent Hypothesis Weighting method leveraging multivariate covariates to optimize statistical power beyond traditional corrections while controlling False Discovery Rate
- **C++ Performance Optimization**: Implemented core algorithms in C++ (via Rcpp) for handling massive genomic datasets, bridging statistical abstraction with low-level memory management
- **Scientific Impact**: Presented research at seven events including Yale University, UNC, and DAGStat 2022; conducted peer reviews for Bioinformatics Advances and Cell Biology
- **Master's Thesis**: Grade 1.0 (highest distinction)
  - 'Better multiple Testing: Using multivariate co-data for hypothesis weighting'

Education

University of Heidelberg *Heidelberg, Germany*
M.Sc. in Mathematics *Oct 2020 – May 2023*

- Grade: 1.0 (summa cum laude)
  - highest distinction
- Yale Exchange Scholar (2022-23): Year-long exchange at Yale's Applied Mathematics Program with honors distinction. Coursework: Deep Learning, Geometric & Topological ML (Prof. Krishnaswamy), Differentiable Manifolds
- Master's Thesis: 'Better multiple Testing: Using multivariate co-data for hypothesis weighting' conducted at EMBL (Supervisors: Prof. Jan Johannes, Dr. Wolfgang Huber)
- Focus: Probability Theory, Machine Learning, High-Dimensional Statistics, Stochastic Processes
- Awards: Gerhard C. Starck Foundation Stipend, Baden-Württemberg Stipend, DAAD Stipend

Hebrew University of Jerusalem *Jerusalem, Israel*
Exchange Student in Mathematics *Sept 2019 – Mar 2020*

- Graduate-level coursework: Functional Analysis, Algebraic Combinatorics, Quantitative Methods
- Awards: PROMOS Stipend (DAAD), Hebrew University Stipend

Publications

Disparities in air pollution attributable mortality in the US population by race, ethnicity and sociodemographic factors July 2024

Pascal Geldsetzer, ***Daniel Fridljand***, Mathew V. Kiang, et al.
www.nature.com/articles/s41591-024-03117-0 (Nature Medicine)

Projects

Agent Olympics Hackathon Munich 2026 *Munich Jan 2026*

- Engineered white-hat voice agent executing automated prompt injection attacks, successfully manipulating customer support agents to access restricted systems
- Built company-agnostic attack system with automated OSINT collection and personalized social engineering campaigns across email, SMS, and WhatsApp channels

Skills

Programming & Frameworks: Python (pandas, PyTorch, FastAPI, Streamlit, PydanticAI, DSPy), R (tidyverse, Shiny, Rcpp, Bioconductor), Java (Spring, JBoss/WildFly), TypeScript/React, C++ (Rcpp), SQL

AI & Machine Learning: LLM Engineering (prompt engineering, DSPy optimization, PydanticAI), Deep Learning (PyTorch, TensorFlow), Statistical ML (Random Forests, Bayesian Non-parametrics, Causal Inference), NLP, Computer Vision

Infrastructure & DevOps: Docker/Podman, Kubernetes, CI/CD (Jenkins, GitHub Actions), AWS (ECR, S3), Monitoring (Prometheus, Grafana, Langfuse), Workflow Orchestration (Temporal, Snakemake)

Data Science & Statistics: Big Data ETL (16B+ records), Bayesian Methods (HDP, Dirichlet Process), Causal Inference, Multiple Testing Correction, Geospatial Analysis, Time-Series Analysis, Visualization (ggplot2, matplotlib)

Specialized Domains: Bioinformatics (single-cell sequencing, cancer genomics, GWAS), Environmental Epidemiology, Customer Support Automation, Enterprise Modernization

Languages: English (C2), German (Native), Russian (Native)