

Rajna Fani

fani@biochem.mpg.de | LinkedIn | Google Scholar | Portfolio

EDUCATION

- **Max Planck Institute of Biochemistry** Munich, Germany
PhD Candidate, Machine Learning for Protein Design Jun 2026 – Present
- **Technical University of Munich (TUM)** Munich, Germany
Master of Science, Computer Science (Major: Machine Learning) Oct 2022 – Oct 2025
 - Master's thesis at MIT's Laboratory for Computational Physiology
 - Erasmus Exchange Semester at Aalto University, Helsinki
- **Ludwig Maximilian University of Munich (LMU)** Munich, Germany
Bachelor of Science, Computer Science (Minor: Computational Linguistics) Apr 2019 – Oct 2022

RESEARCH EXPERIENCE

- **Max Planck Institute of Biochemistry** Munich, Germany
Doctoral Researcher – Machine Learning for Protein Design Jun 2026 – Present
 - Developing machine learning methods for protein design, combining generative modelling and representation learning to engineer functional proteins.
- **Massachusetts Institute of Technology (MIT) – Lab for Computational Physiology** Cambridge, USA
Visiting Researcher – Master's Thesis Mar 2025 – Aug 2025
 - Developed machine learning approaches for clinical data imputation and representation learning across laboratory values, focusing on masked modelling techniques for temporal EHR data.
 - Contributed to a conversational data access system enabling natural language querying of clinical databases, now deployed for international datathons and research collaborations.
 - Analysed clinical documentation to explore links between family engagement and decision-making in critical care.

PUBLICATIONS

- **Coefficient of Variation Masking: A Volatility-Aware Strategy for EHR Foundation Models** — *first author*. ML4H 2025 (Proceedings) & NeurIPS 2025 TS4H Workshop.
- **Rethinking Tokenization for Clinical Time Series: When Less is More** — *first co-author*. ML4H 2025 (Findings).
- **M3: Conversational LLMs Simplify Secure Clinical Data Access** — *first co-author*. ML4H 2025 (Demo).
- **Cached Summary Embeddings for Memory-Efficient EHR Inference** — CHIL 2026; earlier version at the ICLR 2026 Workshop on Time Series in the Age of Large Models.
- **DataAtlas: Automatic Generation of Data Dictionaries Using Large Language Models** — JAMIA Open, 9(3), 2026.
- **Croissant Baker: Metadata Generation for Discoverable, Governable, and Reusable ML Datasets** — arXiv preprint, 2026.
- **Towards Optimizing and Evaluating a RAG QA Chatbot Using LLMs with Human in the Loop** — NAACL @DASH 2024. *Best Paper Award*.

MENTORING

- **MIT Critical Data – Global Clinical AI Datathons** Global
Mentor & Workshop Facilitator Jul 2024 – Present
 - Supporting the deployment of the M3 conversational interface across international datathons (Korea, China, Emory University), advising on dataset alignment, workflow integration, and user onboarding.
 - **GenAI Health Hack 2026 (Barcelona)**: Mentored interdisciplinary teams building LLM- and RAG-based prototypes for real challenges of the Catalan health system, at the second edition co-organised by Hospital Clínic de Barcelona and MIT Critical Data.
 - **8th ZIMAM Digital Health Forum 2025 (Dubai, UAE)**: Invited speaker and mentor at the GCC digital health forum on AI, data governance, and workforce development, representing MIT Critical Data's clinical AI initiatives.
 - **Korea Clinical Datathon 2025 (Seoul)**: Mentored teams on developing and evaluating AI agents in healthcare, co-hosted by Seoul National University Hospital and MIT Critical Data, using MIT's M3 agent and the K-MIMIC multi-center ICU dataset.
 - **Future of Smart Health Forum 2025 (Hangzhou, China)**: Invited speaker and mentor at the Zhejiang University–MIT Critical Data joint event; presented “Data Science Workflow in the Era of GenAI” and facilitated workshops on ethical AI.

- **SCCM Datathon 2025 (Chicago)**: Mentored clinicians and data scientists analysing the timing of first antibiotic administration in septic shock and its association with ICU mortality.
- **Duke University 2025 (Durham)**: Mentored a clinician–data scientist team evaluating whether LLMs can generate evidence-based ICU recommendations for fluids, vasopressors, and antibiotics.
- **Brown University 2025 (Providence)**: Led a “Promptathon” workshop introducing high school students to prompt engineering and safety in clinical AI.
- **Harvard Medical School 2025 (Cambridge)**: Co-led a workshop for public health students on critical thinking, bias detection, and responsible AI deployment in healthcare.
- **Hospital Clínic de Barcelona 2024**: Mentored a team building an EHR-based chatbot for chronic pain education, advising on LLM evaluation using clarity, relevance, and user trust metrics.
- **Balgrist Hospital 2024 (Zurich)**: Facilitated sessions on ICU data exploration and clinical dataset bias detection and their impact on ML performance.

INDUSTRY EXPERIENCE

- **Roland Berger** Munich, Germany
AI Engineer Dec 2025 – June 2026
 - Designed and deployed scalable GenAI solutions in client-facing and internal consulting projects, using RAG architectures, LangGraph, Docker, and vector databases across Azure and GCP; contributed to agentic LLM systems for enterprise use cases.
 - Translated GenAI capabilities into actionable insights for C-level stakeholders through LLM evaluation, prompt engineering, and responsible AI frameworks within the Roland Berger AI Lab.
- **Rohde & Schwarz** Munich, Germany
AI/ML Engineer – Working Student Aug 2024 – Mar 2025
 - Developed and optimized a RAG pipeline AI assistant by integrating LLMs, vector databases, and embedding-based retrieval.
 - Automated ML pipelines for large-scale data extraction, preprocessing, and evaluation across structured and unstructured technical documentation.
- **SAP** Munich, Germany
AI Research Intern Sep 2023 – Jul 2024
 - Developed an HR Q&A chatbot in collaboration with TUM, using SAP’s AI environment to reduce internal HR ticket volume; improved ticket resolution performance by 30% through iterative evaluation with advanced NLP metrics.
- **Get-Ikigai** Munich, Germany
NLP Engineer – Working Student Sep 2023 – Aug 2024
 - Built NLP pipelines to analyse AI–user chats using sentiment analysis, topic modelling, and clustering, improving agent feedback loops.
- **CGI DE** Munich, Germany
Software/AI Engineer – Working Student May 2022 – Dec 2023
 - Enhanced *My BMW* and *MINI* apps with real-time vehicle updates (Android); pioneered NLP for log analysis, cutting error identification times by 20%.

COMMUNITY & SCIENTIFIC ENGAGEMENT

- **OlyNet Student Organization** Munich, Germany
Chief Technology Officer Jul 2022 – Oct 2024
 - Led infrastructure and router setup for 100+ students in Olympic Village.
- **TUM.ai Student Initiative** Munich, Germany
Team Lead – Marketing & Alumni Apr 2023 – Present
 - Increased LinkedIn views by 49.7% and Instagram impressions by 43.3%.

AWARDS & RECOGNITION

- Society of Critical Care in Medicine Datathon 2025 – Grant Recipient
- NAACL @DASH 2024 – Best Paper Award
- Palantir Global Women in Tech 2024 Scholar
- Erasmus+ Scholarship – Awarded Twice for Study Abroad in Finland and USA
- Winner – Microsoft GenAI Hackathon 2023

LANGUAGES

Albanian – native	German – fluent (C1)	English – fluent (C1)	Italian – intermediate
-------------------	----------------------	-----------------------	------------------------