Johannes Roth

johannes@roth24.de LinkedIn GitHub jroth.space Leipzig, Germany

EDUCATION

PhD Candidate – Computational Cognitive Neuroscience ***May 2022 – Present***
Max Planck Institute for Human Cognitive and Brain Sciences & University of Gießen

- Built a high-throughput filtering pipeline to curate **LAION-Natural** (120M images) from LAION-2B (two billion images), made publicly available for research.
- Developed **Natural Controversial Stimuli**, an active learning framework that optimizes stimuli to maximize disagreement between model-based similarity judgments.
- Simulated fMRI experimental designs to identify optimal stimulus presentation rates, proving the **viability of fast presentation rates** for condition-rich datasets.

M. Sc. Computer Science ***2017 – 2021***
Leipzig University
Grade **1.2** (Distinction). Focus: Data Analysis, Machine Learning, Medical Image Processing.

- **Master's Thesis (Grade 1.1)** - Used GANs to synthesize stimuli that maximally activate specific, targeted brain regions, recovering known category-selective areas in human brains.

B. Sc. Business Information Systems ***2014 – 2017***
Leipzig University
Grade **1.5**. Focus: Distributed Systems, E-Commerce, Data Management, Economics

WORK EXPERIENCE

Research Assistant – ML in Medicine ***Jun 2021 – May 2022***
ScaDS.AI Dresden/Leipzig

- Engineered a multi-plane **UNet++ ensemble** for Glioblastoma segmentation (BraTS 2021). Integrated Dice and Boundary loss to achieve competitive segmentation performance (Dice score: 0.90 for whole tumor).
- Developed an attention-based mortality prediction model using **FT-Transformer** and **SAINT** architectures. Implemented epistemic uncertainty estimation to improve safety in medical decision support, achieving 0.85 AUC-ROC.

Data Scientist (Working Student) ***Oct 2019 – May 2021***
CHECK24 (Travel Vertical)

- Designed, implemented and deployed an image-processing micro-service (**Flask + Redis**) enabling fast ML inference over **>20 million** images.
- Used Bayesian hyperparameter optimization for improving ranking metrics and conversion rates of the hotel recommendation system.
- Built Grafana dashboards and outlier detection systems to monitor product price stability and API health.

Full-stack Developer (Freelance) ***Oct 2020 – May 2021***
Kimetric UG

- Implemented two academic websites using **Django**, Nginx, and Gunicorn. Configured Linux hosting environments and automated deployment scripts (CI/CD).

Data Scientist (Working Student) ***Oct 2018 – Oct 2019***
Webdata Solutions GmbH (now Vistex)

- Revamped product-matching pipeline by implementing a neural-network based approach, trained on self-collected web-scraped datasets. Increased matching accuracy from **<50% to 92%**.
- Implemented interpretability techniques (Grad-CAM) to visualize and debug model attention maps, ensuring the network focused on relevant product features rather than background noise.

TECHNICAL SKILLS

Core AI/ML	PyTorch, TensorFlow, LLMs, Transformers (CLIP, ViT), Generative AI (GANs), Contrastive Learning, Metric Learning, Tabular Transformers, Recommendation Systems
Engineering	Python, Docker, SQL (Postgres), Git, CI/CD, Flask, Redis, Linux, Bash, SLURM (HPC)
Methodologies	Model Alignment, Explainable AI, Bayesian Uncertainty Estimation, A/B Testing
Languages	German (Native), English (Fluent)

PROJECTS & OPEN SOURCE

- **thingsvision** (Core Contributor) – Modular feature-extraction library for computer vision. Used to extract features from SOTA models (CLIP, ViT, etc.) for research.
- **relaion-2b-natural** – Curated subset of the massive LAION-2B dataset, optimized for naturalistic vision studies.
- **CNNvis** – Interactive web-based visualization of CNN layer activity, developed for public science communication.

SELECTED PUBLICATIONS & AWARDS

- **Award (2025):** CMBB Replication Award for contribution to reliable coding practices in neuroscience.
- **J. Roth**, M. N. Hebart. *How to sample the world for understanding the visual system*. CCN 2025 (Oral Presentation).
- **J. Roth** et al. *Ten principles for reliable, efficient, and adaptable coding*. Communications Psychology (2025, In Press).
- **J. Roth** et al. *Multi-plane UNet++ Ensemble for Glioblastoma Segmentation*. BraTS Challenge 2021.

LEADERSHIP & COMMUNITY

- **PhD Representative (2023–2024):** Elected to represent >180 doctoral researchers at MPI CBS.
- **Mentoring:** Supervised working students and interns in the research group.
- **Talks:** Presented research findings at international conferences (CCN) and internal institute colloquia.