EMIL DMITRUK

Machine Learning & Data Science | PhD, Computer Science | Python · Julia

emildm26@gmail.com 07934 189 664 edd26 edmitruk edmitruk.com London, UK

SUMMARY

Computer scientist (PhD) who turns advanced mathematics into performant, well-tested software. Works from first principles, reducing a problem to its core, designing experiments that build understanding, then delivering correct, reproducible results; effective at model development and finding subtle bugs. Specialist in topological data analysis (persistent homology), with strengths in graph and network analysis, hands-on machine learning and information-theoretic modelling. 7 years in research (PhD + postdoc), building scientific software (analysis pipelines, teaching materials and open-source tooling) in Julia, now working primarily in Python; set up and administer a High-Performance Computing (HPC) cluster. Author of 8 publications, including a first-authored article in *PLOS Computational Biology*.

EXPERIENCE

Research Associate

University of Hertfordshire, Biocomputation Research Group

Oct 2023 – Present Hatfield, UK

- Re-architected core computation pipelines for ∼**10–15**× **throughput** over the original research implementation, via functional refactoring (pure functions, memoization) and algorithmic caching of intermediate results in Julia and Python, cutting multi-day batch jobs to hours
- Set up and administer a **Slurm HPC cluster**, provisioned as code with Ansible; designed fault-tolerant distributed processing for memory-intensive workloads (100+ GB per stage)
- Ran the experiments for an empowerment-based reinforcement-learning framework (IMOL 2025 workshop; extended first-author manuscript in preparation); see *Research Projects*

PhD Researcher

University of Hertfordshire, Biocomputation Research Group

Mar 2019 – May 2025 Hatfield, UK

- Applied algebraic topology to high-dimensional network data (persistent homology, clique filtrations and persistence landscapes) to characterize brain connectivity; first author on four CNS conference abstracts and a preprint
- First-authored in *PLOS Computational Biology*: applied persistent homology (cubical complexes) to quantify human perception of abstract art, comparing artist works against GAN-generated pseudo-art via eye-tracking, EEG and questionnaires (N=60); designed and ran the complete statistical analysis
- Built reproducible analysis pipelines with permutation testing and multiple-comparison correction; authored open-source Julia packages that extract network properties through TDA; design custom scientific visualizations (Julia/Makie) that make high-dimensional, abstract concepts legible; collaborated across four countries (UK, Germany, Poland, Italy)

Visiting Lecturer

University of Hertfordshire

Jan 2020 – Present Hatfield, UK

- Delivered mathematics and data-science modules to 120+ students per cohort over 5 years (interactive live-coding, building and debugging models with students in real time)
- Supervised 15 MSc researchers across 6 dissertation projects in applied ML, computer vision and statistical analysis

Research Intern

Gwangju Institute of Science and Technology (2018) · Pro-PLUS SA (2016)

South Korea · Poland

Built end-to-end signal-processing and classification pipelines for biomedical time-series (EEG, ECG): preprocessing, feature extraction and neural-network classification.

SELECTED RESEARCH PROJECTS

Empowerment for Open-Ended Learning

Information-theoretic RL · Python, JAX

- Research framework testing whether empowerment (mutual information between an agent's goals and outcomes) can unify goal discovery and skill acquisition under a limited budget
- Implemented gridworld and `xland-minigrid` (JAX/JIT) environments and tabular Q-learning / Dyna-Q agents; ran systematic, ablated experiments
- Found silent correctness errors that had invalidated earlier results, and built the diagnostics that isolated and corrected them

MesoSCOUT: Open-Source Topological Network Analysis

Julia

- Framework characterizing large, high-dimensional networks via persistent homology; a rank-based design detects mesoscale structure from the relative strength of connections, not absolute counts
- Demonstrated robustness to acquisition differences, enabling comparison across two independent datasets (COBRE clinical cohort, N=88; HCP), first-authored preprint (2025)

TECHNICAL SKILLS

Languages: Python (primary), Julia, SQL, Bash; C++, MATLAB, R

Core methods: Topological data analysis (persistent homology), graph & network analysis, information theory, reinforcement learning, statistical inference & experimental design, time-series & signal processing, scientific & data visualization

ML & scientific computing: PyTorch, JAX, scikit-learn, NumPy, pandas, SciPy, statsmodels, Makie; deep learning (CNNs, RNNs)

Infrastructure: Linux, HPC (Slurm cluster setup & administration), Ansible (IaC), Docker, Git/GitHub, distributed computing, CI/CD (GitHub Actions)

GenAI & LLMs: RAG pipelines, vector databases (Qdrant), LLM APIs (OpenAI, Anthropic), multi-agent systems, knowledge graphs (Neo4j), LLM observability (LangFuse)

SELECTED PUBLICATIONS

E. Dmitruk et al., "Art's hidden topology — A window into human perception," *PLOS Computational Biology*, 22(5):e1014156, 2026.

E. Dmitruk, C. Metzner, V. Steuber, S. Kadir, "MesoSCOUT: A novel tool for revealing mesoscale organization in the white-matter connectome," *bioRχiv preprint*, 2025.

N. Catenacci Volpi, K. Archer, **E. Dmitruk** et al., "Empowerment for Goal-driven Open-Ended Learning," *IMOL Workshop*, 2025; extended first-author version in preparation.

M. Yaqoob, **E. Dmitruk**, V. Steuber, S. Kadir, "From evolution to creation of spiking neural networks using graph-based rules," *CNS*2025, J Comput Neurosci*, 2025.

Full list of 8 publications, including four first-authored CNS conference abstracts (2020–2023), available on request.*

EDUCATION

PhD in Computer Science

University of Hertfordshire

2019 – 2025 | Hatfield, UK

Thesis: "Topological insight into neuroscience" (algebraic topology applied to network analysis)

MSc Electronics & Computer Science in Medicine

Warsaw University of Technology

2017 – 2019 | Warsaw, Poland

Includes ERASMUS exchange at KU Leuven, Belgium (Biomedical Engineering, 2017–2018)

BSc Biomedical Engineering (Distinction)

Warsaw University of Technology

2013 – 2017 | Warsaw, Poland

CERTIFICATIONS & APPLIED AI

AI Devs 3 & 2, AI_devs (2023, 2025). Intensive hands-on cohorts on LLM and agent engineering: multi-agent systems, retrieval-augmented generation with vector databases (Qdrant), knowledge graphs, browser automation, LLM observability (LangFuse), and OpenAI/Whisper/DALL-E integration with output-validation guards.

LANGUAGES

Polish: Native speaker | **English**: Professional proficiency (CAE) | **German**: Elementary