

Matei Cananau

ML/AI Research Engineer

Contact

- **Email:** mateicananau@gmail.com
- **Substack (AI Alignment writing):** <https://mateicananau.substack.com>
- **Phone number:** +46 727 064 134

Education

2021 - 2026: KTH Royal Institute of Technology, Sweden

- BSc & MSE in Media Technology (2021-2026)
- MSc in Machine Learning (2024-2026)

Experience

2025 - NOW: Co-Founder at Tellintel. KTH Innovation Launch Batch 22

- Built platform to systemize innovation for enterprises.
- Developed Socratic AI alignment philosophy to make the AI prompt the user rather than outsourcing creativity to the model.

2023 - 2026: Co-Founder & AI Consultant at Servai

- Built LLM-based systems using OpenAI, Gemini, and Azure's AI Hub with real-time integration through clients' APIs.
- Designed pipelines for data analysis through MySQL and PostgreSQL.
- Deployed AI systems with 10k+ chats.

2022 - 2023: TA in Programming & Scientific Computing at KTH

- Taught Python and numerical methods.
- Helped students in their assignment, coding, and problem solving skills.

Projects

MSc Thesis: Hylomorphic AI: Virtue-Based Ethical Alignment for LLMs

- Extracted internal belief representations using RepE and LAT through Aristotelianism.
- Implemented two steering methods as post-training alternatives to RLHF.
- Improved robustness and reduced contradictions.
- Compared base vs aligned model behavior.

Advanced Deep Learning: Label Swapping for Imbalanced Music Datasets

- Implemented LNR method to shift decision boundaries in imbalanced datasets.
- Trained ResNet18 on MFCC inputs, evaluated minority class performance.
- Reproduced prior work and applied method to a new domain. Grade A.

BSc Thesis: Digital Mindfulness and its Impact on Young Adults

- Studied how taking mindful breaks from technology impacts well-being.
- Paper was used as a reading assignment for future Media students.

Skills

- **Machine Learning:** PyTorch, Deep Learning, transformers, CNNs, regression, NLP, model evaluation, computer vision, DGM, diffusion, MLOps
- **AI Systems:** LLM APIs, RAG, prompting, agents, evaluation, deployment, fine-tuning, MCPs
- **Programming:** Python, NumPy, Pandas, Scikit-learn, SQL, JavaScript
- **Misc:** fast learner, communication, presentations, AI philosophy, design, Git, GitHub, Bash, Terminal, Docker, Figma, visualization, web development