Bartłomiej Cupiał

github webpage twitter scholar

Education

University of Warsaw - PhD Student (October, 2023 - July, 2027)
Research in Reinforcement Learning and Large Language Models

Jagiellonian University - Master Student (October, 2021 - July, 2023)
Thesis topic "How to best transfer knowledge in reinforcement learning?"

Wrocław University of Science and Technology - Bachelor Student (October, 2016 - February, 2020)
Thesis topic “Camera recognition from photos”

Research Experience

Sakana AI, supervised by Jeffrey Seely - Research Intern (September, 2025 - January, 2026)
Decentralized multi-agent coordination via Sheaf-Constrained ADMM

UCL Dark, Tim Rocktäschel Group - Visiting Scholar (January, 2025 - July, 2025)
Eliciting dynamic planning for LLM Agents with PPO

IDEAS NCBR - Research Intern (January, 2023 - present)
Research in Reinforcement Learning and Large Language Models

FNP Team-Net Grant - Research Intern (October, 2022 - July, 2023)
Research on GAN-based Plugin Models for Video Generation

Industry Experience

Hemolens Diagnostics - Machine Learning Engineer and Team Lead (November, 2019 - October, 2022)
Research & Development for automation of non-invasive cardiac diagnostics system

Asseco - C++ Developer in starter program (July, 2019 - October, 2019)
Development & Maintenance of national social insurance system (KSI ZUS)

Lectures and presentations

MLinPL Conference Speaker, Warsaw, Poland (November, 2024)

Giving seminar University College London (June, 2024)

Giving seminar Korea Institute For Advanced Study (June, 2024)

Giving seminar Machine Learning and AI Academy, youtube (February, 2024)

Conducting Workshops in Solvro Students' Association (February, 2020)

MLinPL Conference Student Speaker, Warsaw, Poland (November, 2019)

Publications

1. Learning Multi-Agent Coordination via Sheaf-ADMM

ICML 2026
Sakana AI, AKCES NCBR, University of Warsaw (September, 2025 - January, 2026)
Standard neural architectures are monolithic: a single large network processes inputs as a unified entity. By contrast, much of the intelligence observed in nature is collective, emerging from many locally informed agents that coordinate to solve global tasks. Inspired by this, we developed a framework that replaces a monolithic network with a population of interacting neural agents, allowing complex behavior to emerge through communication and coordination.
https://arxiv.org/abs/2605.31005

2. Learning When to Plan: Efficiently Allocating Test-Time Compute for LLM Agents

IDEAS NCBR, University of Warsaw, University College London (January, 2025 - July, 2025)
Planning before every action (e.g. ReAct) has become the standard in agentic pipelines, but we show this is far from optimal for multi-step RL. Allocating reasoning resources should involve planning, where models reason, make a plan, and then stick to it, rather than re-planning at every step. We trained 8B models to follow this principle using a two-stage pipeline: SFT on synthetic data to teach how to make and follow plans, and with RL fine-tuning to teach when to do so. Beyond improved performance in long-horizon tasks like Crafter, this method also enabled effective human-agent collaboration, where high-level human plans allowed the agent to mine a diamond.
https://arxiv.org/abs/2509.03581

3. BALROG: Benchmarking Agentic LLM and VLM Reasoning On Games

ICLR 2025
IDEAS NCBR, University of Warsaw (June, 2024 - October, 2024)
Benchmark designed to evaluate the agentic capabilities of LLMs and VLMs through a diverse set of challenging games, ranging from simple tasks to complex environments like the NetHackLearning Environment. Results reveal that while current models show some success in easier games, they struggle significantly with more challenging tasks, particularly in vision-based decision-making. Work done in collaboration with PhD students from UCL, University of Oxford and NYU
https://arxiv.org/abs/2411.13543

4. Fine-tuning Reinforcement Learning Models is Secretly a Forgetting Mitigation Problem

Spotlight at ICML 2024
IDEAS NCBR, University of Warsaw, Jagiellonian University (October, 2022 - February, 2024)
Addressing challenges of fine-tuning pre-trained RL agents, which tends to be unstable. Experiments suggest that knowledge retention helps stabilize and scale the model both in the simple robotic environment Meta World as well as in the very challenging domain of NetHack where I was able to beat previous SOTA by over 2x.
https://arxiv.org/abs/2402.02868

5. GAN-based Plugin Model for Video Generation with Applications in Colonoscopy

PLOS ONE
Jagiellonian University (May, 2022 - November, 2023)
Video generation in high resolution is a very expensive task. By pre-training GAN on images and using a plugin model which constructed trajectories in the noise space we were able to generate realistic videos in 1024x1024 resolution.
https://arxiv.org/pdf/2311.03884

Patents

1. Non-invasive diagnosis of coronary artery disease (CAD)

Author of the two patents in medical domain

Hemolens Diagnostics (November, 2019 - October, 2022)

Non-invasive diagnosis involves a multi-step process utilizing advanced imaging and computational techniques. The procedure begins with creating a high-resolution CT scan of a patient chest area. Then coronary arteries are extracted with deep learning methods, by performing segmentation and conversion into triangular meshes. Finally computational fluid dynamics (CFD) simulation is performed to simulate blood pressure before and after coronary artery stenosis. The process allows doctors to diagnose CAD accurately without the discomfort and risks associated with invasive diagnosis.

https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2023166330

https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2024134284