

# Piotr Piękos

PhD Candidate

✉ [piotrpiekos@gmail.com](mailto:piotrpiekos@gmail.com)

🌐 [piotrpiekos.github.io](https://piotrpiekos.github.io)

🐙 [github: piotrpiekos](https://github.com/piotrpiekos)

🌐 [LinkedIn](#)

🔍 [Google Scholar](#)

## Education

### KAUST, Saudi Arabia.

- **GPA 4.0/4.0** Computer Science, Ph.D. (08.2022-Present)  
Working on sparse models (Mixture-of-Experts), neural network architectures and reinforcement learning algorithms.  
Supervised by Prof. Jürgen Schmidhuber

### University of Warsaw, Poland.

- Mathematics, M.Sc. (grad. 2021)  
Thesis: Improving mathematical reasoning in BERT (ACL 2021 Oral)
- Mathematics, B.Sc. (grad. 2018)  
Thesis: A survey of machine learning methods in Natural Language Processing
- Computer Science, B.Sc. (grad. 2017)  
Thesis: An implementation of the face recognition algorithm as an application on an android phone.

## Experience

- 05.2024- **Amazon, Amsterdam, Netherlands, Applied Scientist Internship.**
- 10.2024 Creating hierarchical discrete representations from dense vector representations used for representations of product in the Amazon catalog. Lead to a publication about [Hyperbolic Residual Quantization](#) (NEGL NeurIPS 2025 Workshop) apart from being implemented in the Amazon systems.
- 06.2023- **IDSIA, Lugano, Switzerland, Research Internship.**
- 08.2023 Continued previous remote collaboration that resulted in 2 papers: [SwitchHead](#) (NeurIPS 2024) and [Compositional Optimality Equation](#) (GCRL NeurIPS 2023 Workshop).
- 01.2022- **Polish Academy of Sciences, AWARElab group, Researcher.**
- 06.2022 A collaborator in the [Fast and Precise](#) (Oral - Notable Top 5%, ICLR 2023).
- 10.2021- **Allegro.pl, Research Engineer.**
- 01.2022 Improving the quality of the neural search used in the Allegro.pl portal. Improved the speed of the data processing for the training, which allowed to increase the size of the data used for the training and, hence, improved the performance of the model.
- 07.2017- **ITmagination, Machine Learning Engineer / ML Team Leader.**
- 06.2020 Designing and implementing machine learning systems for external companies. Responsibilities: Example projects:
  - Automatic patent assistant that suggests mentors and next steps based on the description of the project. The implementation was based on the BERT language model.
  - System for predicting necessary staff to operate a shop on a given day.
  - Evaluating online influencers for marketing companies with Object detection methods such as Faster-RCNN.
- 07.2016- **Hcore, Python Developer Internship.**
- 09.2016 Backend Software Engineering in Python for a network monitoring product.

---

## Publications

### **Mixture of Sparse Attention: Content-Based Learnable Sparse Attention via Expert-Choice Routing**

- *NeurIPS 2025 (Efficient Reasoning Workshop)*, , .



Piotr Piękos, Róbert Csordás, Jürgen Schmidhuber

### **Huxley-Gödel Machine: Human-Level Coding Agent Development by an Approximation of the Optimal Self-Improving Machine**

, .



Wenyi Wang\*, Piotr Piękos\*, Li Nanbo, Firas Laakom, Yimeng Chen, Mateusz Ostaszewski, Mingchen Zhuge, Jürgen Schmidhuber

### **SwitchHead: Accelerating Transformers with Mixture-of-Experts Attention**

- *NeurIPS 2024*, , .

Róbert Csordás, Piotr Piękos, Kazuki Irie, Jürgen Schmidhuber

### **PhysGym: Benchmarking LLMs in Interactive Physics Discovery with Controlled Priors**

- *NeurIPS 2025*, , .



Yimong Chen, Piotr Piękos, Mateusz Ostaszewski, Firas Laakom, Jürgen Schmidhuber

### **Hyperbolic Residual Quantization: Discrete Representations for Data with Latent Hierarchies**

- *NeurIPS 2025 (NEGEL Workshop)*, .



Piotr Piękos, Subhradeep Kayal, Alexandros Karatzoglou

### **Oral(Notable Top 5%) Fast and Precise: Adjusting Planning Horizon with Adaptive Subgoal Search**

- *ICLR 2023*, , .

Michał Zawalski, Michał Tyrolski, Konrad Czechowski, Tomasz Odrzygóźdź, Damian Stachura, Piotr Piękos, Yuhuai Wu, Łukasz Kuciński, Piotr Miłoś

### **Oral Measuring and Improving BERT's Mathematical Abilities by Predicting the Order of Reasoning**

- *ACL 2021*, , .



Piotr Piękos, Henryk Michalewski, Mateusz Malinowski

### **Best Paper Award Mindstorms in Natural Language-Based Societies of Mind**

- *NeurIPS 2024 (R0-FoMo Workshop) CVMJ 2025*, .

Mingchen Zhuge, Haozhe Liu, Francesco Faccio, Dylan R Ashley, Róbert Csordás, Anand Gopalakrishnan, Abdullah Hamdi, Hasan Abed Al Kader Hammoud, Vincent Herrmann, Kazuki Irie, Louis Kirsch, Bing Li, Guohao Li, Shuming Liu, Jinjie Mai, Piotr Piękos, Aditya Ramesh, Imanol Schlag, Weimin Shi, Aleksandar Stanić, Wenyi Wang, Yuhui Wang, Mengmeng Xu, Deng-Ping Fan, Bernard Ghanem, Jürgen Schmidhuber

### **Utilizing a Malfunctioning 3D Printer by Modeling Its Dynamics with Artificial Intelligence**

- *ICRA 2024*, , .

Renzo Cabalero\*, Piotr Piękos\*, Eric Feron, Jürgen Schmidhuber

### **Efficient Value Propagation with the Compositional Optimality Equation**

- *NeurIPS 2023 (GCRL workshop)*, .

Piotr Piękos, Aditya Ramesh, Francesco Faccio, Jürgen Schmidhuber

---

## Awards

**Best Poster Award** — EEML 2025 Summer School, for *Mixture of Sparse Attention: Content-Based Learnable Sparse Attention via Expert-Choice Routing*

**Oral Presentation** — ICLR 2023, for *Fast and Precise: Adjusting Planning Horizon with Adaptive Subgoal Search*

**Best Paper Award** — R0-FoMo NeurIPS 2023 Workshop, for *Mindstorms in Natural Language-Based Societies of Mind*

**Oral Presentation** — ACL 2021, for *Measuring and Improving BERT's Mathematical Abilities by Predicting the Order of Reasoning*

**Best Poster Award** — EEML 2021 Summer School, for *Measuring and Improving BERT's Mathematical Abilities by Predicting the Order of Reasoning*

## Open-source

**BERT for trax** 🤖: Added BERT model and masked language modelling pipeline preparation to trax.

**EGS** 🤖: Triton kernel for fused operation of gather/scatter with a linear layer.

## Technical skills

Python, PyTorch, [Triton](#),  
Huggingface, Transformers, TRL  
Reinforcement Learning, Deep Learning,  
NLP, LLMs, MoE, Algorithms,  
Linux, Bash, GIT, SQL, Unit tests

## Languages

English	Fluent (111/120 TOEFL score)
Polish	Native

## Interests

Chess, Music, Psychology