

# Piotr Nawrot

☎ (+48) 726 408 661  
✉ [piotr.nawrot@ed.ac.uk](mailto:piotr.nawrot@ed.ac.uk)  
📄 [piotrnawrot.github.io](https://piotrnawrot.github.io)

---

## Education

- 2022 - 2026 **PhD, Natural Language Processing**, *University of Edinburgh*.  
Topic: **Adaptive Computation for Neural Language Models**.  
Advisors: Dr. Edoardo Maria Ponti and Prof. Ivan Titov.
- 2018 - 2021 **Bachelor, Computer Science**, *University of Warsaw*.  
Thesis: **Transformers for Classification and Image Generation**.  
Top grades in Machine Learning, Dean's List for 2 semesters.

---

## Experience

- May 2023 - Present **Research Intern**, *Nvidia*, Remote, UK.  
Pursued research on improving the **efficiency of Large Language Models** during inference with Adrian Łańcucki, Marcin Chochowski, and David Tarjan.
- May 2022 - Aug 2022 **Research Scientist Intern**, *Facebook AI Research*, Paris, France.  
Pursued research on **unsupervised speech representation learning** with Jade Copet, Yossi Adi, Gabriel Synnaeve, and Emmanuel Dupoux.
- Jan 2022 - May 2022 **Research Assistant**, *University of Wrocław*, Wrocław, Poland.  
Pursued research on Dynamic Pooling for Autoregressive Transformer Language Models which resulted in a **publication in Proceedings of ACL 2023**.
- Jul 2021 - Dec 2021 **Deep Learning and Algorithms Intern**, *Nvidia*, Warsaw, Poland.  
Implemented multi-node wav2vec 2.0 inference with two external LMs, observing a 5% relative valid WER improvement over beam decoding with n-gram.  
Proposed and implemented modifications of wav2vec 2.0 that improved relative valid WER by 9%.  
Implemented a Python interface for streaming ASR models in real time.
- Oct 2020 - Oct 2021 **Research Side Project**, *Google Brain*, Warsaw, Poland.  
Pursued research on Hierarchical Language Models which resulted in a **publication in Findings of NAACL 2022**.  
Contributed to Google's Trax library: Hourglass model, Transformer-XL relative attention, Rotary positional embeddings.
- Jun 2020 - Nov 2020 **Deep Learning and Algorithms Intern**, *Nvidia*, Warsaw, Poland.  
Refactored research code with **multi-speaker functionality of TTS model** and contributed to open-source repository: [github.com/NVIDIA/DeepLearningExamples](https://github.com/NVIDIA/DeepLearningExamples).  
Pursued research on **extracting grapheme boundaries** from outputs of ASR models. Introduced new rules to the main text data preprocessing module which is used and shared across Nvidia's speech teams.

---

## Competitive Programming Achievements

- Dec 2019 **Bronze medal in the ACM ICPC Central European Regional Contest**.

- Oct 2019 **5th place in the Polish Collegiate Programming Contest.**
- Sep 2018 **7th place in Microsoft BubbleCup 11.**
- May 2018 **Silver Medal in the 25th Polish Olympiad in Informatics.**
- Jun 2017 **Ranked 219th out of over 25000 participants in Google Code Jam.**
- May 2017 **Bronze Medal in the 24th Polish Olympiad in Informatics.**

---

## Extracurricular Activities

- 2023 **Open Source Contributor**, Released a **nanoT5** GitHub repository (**915★**) for efficient pre-training and fine-tuning of T5-style language models.
- 2023 **Reviewer**, *ICLR 2024, ACL 2023 SRW, Instruction Workshop NeurIPS 2023.*
- 2021 **Panel Discussion Coordinator**, *ML in PL Conference*, Co-organized a discussion in English on the 'Sins and Marvels of AI Research'.
- 2020 **Board Member**, *Machine Learning Society, University of Warsaw*, Organised bi-weekly seminars to discuss recent advances in the AI&ML field.
- 2019 **Volunteer Tutor**, *Meet IT*, Mentored two high-school students to become laureates in the Olympiad in Informatics (Top 10% nationally).
- 2018 **Programming Competition Author**, Prepared 8 algorithmic tasks with tests, wrote model solutions in C++, and organised a lecture afterwards.

---

## Selected Publications

1. **Piotr Nawrot**, Adrian Łańcucki, Marcin Chochowski, David Tarjan, Edoardo M. Ponti. 2024. *Dynamic Memory Compression: Retrofitting LLMs for Accelerated Inference*. arXiv:2403.09636.
2. **Piotr Nawrot**. 2023. *nanoT5: A PyTorch Framework for Pre-training and Fine-tuning T5-style Models with Limited Resources*. In Proceedings of NLP-OSS Workshop at EMNLP 2023.
3. **Piotr Nawrot**, Jan Chorowski, Adrian Łańcucki, Edoardo Maria Ponti. 2022. *Efficient Transformers with Dynamic Token Pooling*. In Proceedings of ACL 2023. **Nominated for Best Paper Award**.
4. **Piotr Nawrot**, Szymon Tworkowski, Michał Tyrolski, Łukasz Kaiser, Yuhuai Wu, Christian Szegedy, Henryk Michalewski. 2021. *Hierarchical Transformers Are More Efficient Language Models*. In Findings of NAACL 2022.