Evgeniia Tokarchuk

PhD Candidate, University of Amsterdam

evgeniia@tokarch.uk • +31 (627) 560-449 • evgeniia.tokarch.uk

SUMMARY

I am a last-year PhD candidate at the University of Amsterdam. I am interested in [text] generative modeling, geometry-aware modeling and representation learning

EDUCATION

University of Amsterdam, Amsterdam, The Netherlands

- PhD Candidate, Language Technology Lab — May 2021 – present
  - Working on continuous text representation for generative models under the supervision of Dr. Vlad Niculae.

RWTH Aachen University, Aachen, Germany

- M.Sc. Software Systems Engineering — Oct 2017 – Mar 2021
  - GPA: 1.7 (1.0 best; 5.0 worst)
  - **Relevant Coursework**: Statistical Classification and Machine Learning, Natural Language Processing, Machine Learning, Advanced Machine Learning, Automatic Speech Recognition (Lab)
  - **Master thesis topic**: Tight model integration for pivot-based Neural Machine Translation

South Ural State University, Chelyabinsk, Russia

- Specialist in Information Security for Automated Systems — Sep 2011 – Jun 2016
  - Graduated with Honors

SKILLS

- **Technical**: machine learning, natural language processing, machine translation, python (8+yrs), pytorch, linux/bash shell, git
- **Languages**: Russian (native), English (C1), Dutch (A1)

WORK EXPERIENCE

eBay, Aachen, Germany

- Working Student, Human Language Tech — Oct 2019 – Mar 2021
  - Worked on end-to-end fine-tuning of pivot-based neural machine translation using non-autoregressive models and reinforcement learning.

i6, RWTH Aachen University, Aachen, Germany

- Student Worker, Human Language Technology and Pattern recognition group, — Feb 2018 – Sep 2019
  - Worked on data engineering for Named Entity Recognition task using Flair (PyTorch-based framework).
  - Experimented with Character-Based Machine Translation using RETURNN (TensorFlow-based framework).

PressFoto, Chelyabinsk, Russia

- Algorithm's developer, R&D department — Oct 2016 – Aug 2017
  - Worked on optimization of image search quality results using text queries. Worked on embeddings similarity search.

The Honeynet Project,

- Google Summer of Code 2016, — May 2016 – Aug 2016
  - Developed from scratch **SNARE** and **TANNER** honeypot project using python3 and aiohttp framework.

PAPERS

- **E. Tokarchuk**, H.C. Bakker, V. Niculae. On the Matter of Embeddings Dispersion on Hyperspheres. Extended abstract, *GRaM workshop*, ICML 2024 (non-archival)
- **E. Tokarchuk**, V. Niculae. The Unreasonable Effectiveness of Random Target Embeddings for Continuous-Output Neural Machine Translation. *Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL)*. June 2024 (oral presentation).
- **E. Tokarchuk**, V. Niculae. On Target Representation in Continuous-output Neural Machine Translation, *RepL4NLP workshop*, collocated with ACL, May 2022.
- **E. Tokarchuk**, J. Rosendahl, W.Wang, P. Petrushkov, T. Lancewicki, S. Khadivi and H. Ney. Integrated Training for Sequence-to-Sequence Models Using Non-Autoregressive Transformer, *IWSLT*, August 2021.
- **E. Tokarchuk**, D. Thulke, W. Wang, C. Dugast, and H. Ney. Investigation on Data Adaptation Techniques for Neural Named Entity Recognition. *ACL Student Research Workshop*, August 2021.

- **E. Tokarchuk**, J. Rosendahl, W. Wang, P. Petrushkov, T. Lancewicki, S. Khadivi and H. Ney. Towards Reinforcement Learning for Pivot-based Neural Machine Translation with Non-autoregressive Transformer, *Reinforcement Learning for Real Life Workshop*, July 2021.

TEACHING	**University of Amsterdam**, Amsterdam, The Netherlands

- TA, Natural Language Processing 2   Apr 2022 – Jun 2023
  - Supervised student's projects on applications of reinforcement learning for non-autoregressive NMT.

University of Amsterdam, Amsterdam, The Netherlands
- TA, Natural Language Processing 2   Apr 2024 – Jun 2024
  - Supervised student's projects on applications of MBR for autoregressive and non-autoregressive models.

SUPERVISION & MENTORSHIP

IAI community mentor, University of Amsterdam   Nov 2023 – present
- Inclusive AI (IAI) community for the Master's programme AI mentor

BSc KI, University of Amsterdam   Apr 2023 – Jul 2023
- Optimising inference speed of beam search continuous output NMT, Isaac Santhagens

BSc KI, University of Amsterdam   Apr 2023 – Jul 2023
- Optimising inference speed of beam search continuous output NMT, Isaac Santhagens

BSc KI, University of Amsterdam   Nov 2022 – Mar 2023
- Analysis of Beam Search for decoding continuous text representations, Kacper Bartosik

BSc KI, University of Amsterdam   Apr 2022 – Jun 2022
- Length Estimation in Non-autoregressive Text Generation Models, Erik Buis

Master Thesis AI, University of Amsterdam   Nov 2021 – Oct 2022
- Recycling English Language Models for Dutch Summarisation, Redouane Dahmani

AWARDS & SCHOLARSHIPS
- **ProbAI 2023 participant**, Trondheim, Norway   Jun 2023 – Jun 2017
- **DAAD Study Scholarships for Graduates of All Disciplines**   2018 – 2019

COMMUNITY SERVICE
- **Conference review**: ACL*, NeurIPS, ICML, ICLR   2021 – present
- **NGO Volunteers with No Limits**: chairwoman   Mar 2022 – present
- **The HoneyNet Project**: Google Summer of Code Mentor   2017 – 2021