Taisiya Glushkova

LinkedIn / Google Scholar / Github / glushkovato@gmail.com

ML/NLP engineer with a PhD in Computer Science, specializing in evaluation methods for production AI systems including Machine Translation (MT) and Large Language Models (LLMs). Experienced in translating research into practical tools, including evaluation pipelines that improved model reliability for real-world applications, and working closely with collaborators, mentoring junior researchers, and ensuring research outcomes aligned with business goals. Passionate about applying AI across domains and driving projects from concept to impact.

Based in Lisbon, Portugal. Hold a valid Portuguese residency permit, allowing work authorization.

Technical Skills

- Deep Learning, Machine Learning, Natural Language Processing.
- Fluent in Python, with minor experience with R and C++. Experience with DL frameworks: PyTorch, Keras.
- Proficient with common data science libraries such as NumPy, Pandas, Matplotlib, NLTK, Scikit-learn, and Transformers.
- Experience with LLM fine-tuning (LoRA, QLoRA), memory optimization, and transformer-based architectures.
- Comfortable with common data science tools, e.g. bash, git, slurm, SQL, Jupyter, AWS (EC2, S3, SageMaker), WandB.

Education

PhD in Computer Science and Engineering, Instituto Superior Técnico — Lisbon, Portugal
Thesis title: Uncertainty Estimation and Robustness in Machine Translation Evaluation — 2020 – 2024
Collaborative projects: MAIA (**Multilingual AI Agent Assistants**), NextGenAI (**Center for Responsible AI**)

MSc in Computer Science, HSE University — 2018 – 2020

BSc in Applied Mathematics and Computer Science, HSE University — 2014 – 2018

Work Experience

Polynome, ML Engineer — Dubai, UAE (remote) | 2024 – 2025
Worked on adapting Large Language Models (LLMs) for high-quality machine translation between low-resource languages, focusing on efficient fine-tuning methods and evaluation approaches.

- Designed a memory-efficient fine-tuning pipeline using LoRA adaptation with custom training infrastructure and evaluation metrics (BLEU, COMET) to enable robust neural machine translation between Russian and Kazakh languages.
- Expanded the scope to develop a general-purpose Kazakh LLM through comprehensive data collection for continuous pretraining and fine-tuning steps. Conducted efficiency optimization research on existing models including MoE pruning techniques and tokenizer optimization to improve performance and resource utilization.

Instituto Superior Técnico, PhD Researcher — Lisbon, Portugal | 2020 – 2024
Led research on uncertainty estimation for efficient machine translation (MT) evaluation, resulting in the development of more robust evaluation models, under the supervision of André F. T. Martins.

- Contributed to extending COMET, state-of-the-art MT evaluation metric, with uncertainty measurement feature.
- Mentored junior PhD researchers and coordinated with partner organizations (Unbabel, Center for Responsible AI).
- Published papers at EMNLP 2021, EMNLP 2022, EAMT 2023.

Carnegie Mellon University, Research Intern — Pittsburgh, USA | Summer 2022
Developed robust and interpretable MT evaluation methods under the supervision of Alon Lavie.

- Collaborated with Unbabel on WMT Shared Tasks, securing second place in Metrics and winning the Quality Estimation task.
- Contributed to strategic planning and execution of research initiatives, aligning research goals with business objectives.

Digital Decisions, ML Engineer — Moscow, Russia | 2019 – 2020
Developed a hierarchical classification system for customer messages at Vkusvill, optimizing routing and response accuracy.

- Managed projects timelines, coordinated with stakeholders, and ensured deliverables met quality standards.

Unbabel, Applied AI Research Intern — Lisbon, Portugal | Summer 2019
Implemented language model-based data filtering for machine translation, improving data quality and translation accuracy.

- Enhanced translation model performance in several language pairs, leading to a higher customer satisfaction.

Diginetica, Junior Data Scientist/Engineer — Moscow, Russia | 2017
Built tools improving search and product discovery for eCommerce clients, driving measurable user and business impact.

- Designed a spell-check and correction tool for retail search queries, resulting in 12% growth in search-driven revenue.
- Delivered concise product review summarization tool, improving customer experience and engagement.
- Applied data-driven prioritization to backlog tasks, ensuring efficient delivery of high-impact features.

Yandex, Software Engineering Intern — Moscow, Russia | 2015
Implemented classification and ranking of multiple object definitions by user relevance.

Selected Publications

Uncertainty-Aware Machine Translation Evaluation [ACL Anthology, slides]
Taisiya Glushkova, Chrysoula Zerva, Ricardo Rei, André F. T. Martins
Presented at the *Sixth Conference on Machine Translation (WMT21) at EMNLP 2021*
EMNLP Findings 2021

Disentangling Uncertainty in Machine Translation Evaluation [ACL Anthology, poster]
Chrysoula Zerva, **Taisiya Glushkova**, Ricardo Rei, André F. T. Martins
EMNLP 2022

BLEU Meets COMET: Combining Lexical and Neural Metrics Towards Robust Machine Translation Evaluation
Taisiya Glushkova, Chrysoula Zerva, André F. T. Martins [ACL Anthology, poster]
EAMT 2023

COMET-22: Unbabel-IST 2022 Submission for the Metrics Shared Task [ACL Anthology, poster]
Ricardo Rei, José G. C. de Souza, Duarte Alves, Chrysoula Zerva, Ana C. Farinha,
Taisiya Glushkova, Alon Lavie, Luisa Coheur, André F. T. Martins
EMNLP WMT 2022, *Second place on the WMT 2022 Metrics shared task*

CometKiwi: IST-Unbabel 2022 Submission for the Quality Estimation Shared Task [ACL Anthology]
Ricardo Rei, Marcos Treviso, Nuno M. Guerreiro, Chrysoula Zerva, Ana C. Farinha, Christine Maroti,
José G. C. de Souza, **Taisiya Glushkova**, Alon Lavie, Luisa Coheur, André F. T. Martins
EMNLP WMT 2022, *Winner of the WMT 2022 QE shared task*

Are References Really Needed? Unbabel-IST 2021 Submission for the Metrics Shared Task [ACL Anthology]
Ricardo Rei, Ana C. Farinha, Chrysoula Zerva, Daan van Stigt, Craig Stewart, Pedro G. Ramos,
Taisiya Glushkova, André F. T. Martins, Alon Lavie
EMNLP WMT 2021

Teaching

Instituto Superior Técnico — Lisbon, Portugal
Teaching Assistant for Deep Learning course — 2021 – 2022

HSE University — Moscow, Russia
Teaching Assistant for Natural Language Processing (NLP) course — 2018 – 2019
Teaching Assistant for Python course — 2016 – 2018

Professional & Academic Activities

Conference Reviewing: NAACL 2024, ACL 2026
Poster presentation at Center for Responsible AI | Demo Day — June 2024
Poster presentation at EAMT 2023, *Tampere, Finland* — June 2023
Poster presentation at EMNLP 2022, *Abu Dhabi, UAE* — December 2022
Participated in a poster session at CMU Portugal 2022 Summit — November 2022
Oral presentation at EMNLP WMT 2021 Workshop, *Punta Cana, Dominican Republic* — November 2021
Participated in LxMLS (8th Lisbon Machine Learning School) — Summer 2018