

Fedor Sizov

+4915141388187 | fedia.sizov@gmail.com | [linkedin.com/in/fedor-sizov](https://www.linkedin.com/in/fedor-sizov) | github.com/sfedia

BIO

I am a Computational Linguistics student at Saarland University and a research assistant at the German Research Center for Artificial Intelligence (DFKI). My focus areas include Natural Language Understanding (NLU) and Machine Translation (MT).

EDUCATION

Saarland University

Bachelor of Science in Computational Linguistics

Saarbrücken, Germany

Oct. 2023 – Jun. 2026

Russian University of Humanities

Bachelor of Arts in Fundamental and Computational Linguistics

Moscow, Russia

Sept. 2019 – Oct. 2023

PUBLICATIONS

1. **Fedor Sizov**, Cristina España-Bonet, Josef van Genabith, Roy Xie, Koel Dutta Chowdhury. Analysing Translation Artifacts: A Comparative Study of LLMs, NMTs, and Human Translations. Accepted to *Ninth Conference on Machine Translation (WMT24)*, *EMNLP 2024*. ([pdf](#))
2. Daria Zhornik, **Fedor Sizov**. Linguistic variation and minor languages corpora: a case study of mansi dialects. In *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2018"*, May 2018. ([pdf](#))

EXPERIENCE

Research Assistant

German Research Center for Artificial Intelligence

June 2024 – *Currently*

Saarbrücken, Germany

- I conduct research on identifying and mitigating translation artifacts in human- or machine-translated text.
- First author of the paper “Analysing Translation Artifacts: A Comparative Study of LLMs, NMTs, and Human Translations”, accepted to EMNLP.

Software Engineer

MillionAgents

July 2023 – Nov 2023

Moscow, Russia

- Developed about 15 parsers in Python for collecting data from different retail store networks: some of parsers were written using requests/lxml, other involved Selenium and DOM scripting
- Introduced unit testing inside the company and added unit tests for majority of company’s parsers
- Analyzed HTTP requests of Android applications using Burp Suite proxying and APK decompilation
- Updated and improved schemas in the company’s PostgreSQL database
- Made updates to company’s Django views

Software Engineer Intern

Yandex

June 2022 – Sep 2022

Moscow, Russia

- As a part of monitoring and automation development group, I automated a range of hardware lifecycle scenarios such as upgrades and migrations

CONFERENCE TALKS

- Analysing Translation Artifacts: A Comparative Study of LLMs, NMTs, and Human Translations. Oral Presentation at WMT24 (EMNLP 2024).
- Revisiting nltk-drt: DRT extension for NLTK updated and improved. Conference talk at TaCoS 2024 (33rd Computational Linguistics Student conference).

TECHNICAL SKILLS

- **Programming Languages:** Python, C++, JavaScript
- **Tools:** Git, Docker, Linux
- **Machine Learning:** PyTorch, NumPy, scikit-learn, pandas