

Daryna Dementieva

daryna.dementieva@tum.de
Personal page: <https://dardem.github.io>
<http://www.linkedin.com/in/daryna-dementieva>

Social Research Computing Group, CIT
Technical University of Munich (TUM)
Garching, Germany

RESEARCH INTERESTS	natural language processing for social good, fake news detection, multilingualism, hate speech combating, text detoxification, text style transfer, explainable AI
EDUCATION	<p>November 2019 – September 2022 Ph.D. in Computer Science Skolkovo Institute of Science and Technology (Skoltech), Moscow <i>Thesis:</i> Methods for Fighting Harmful Multilingual Information <i>Supervisor:</i> Ass. Prof. Alexander Panchenko <i>Links:</i> information and video of the defense</p> <p>September 2017 – August 2019 Master of Science in Applied Math and Physics Moscow Institute of Science and Technology (MIPT), Moscow <i>Thesis:</i> News Aggregation and Personalization for Professional Groups <i>Supervisor:</i> Prof. Konstantin Vorontsov</p> <p>September 2013 – August 2017 Bachelor of Science in Applied Math Taras Shevchenko's Kyiv National University (KNU), Kyiv</p>
EXPERIENCE	<p>Postdoctoral Researcher TUM Munich Jun 2022 - now <i>Host:</i> Prof. Dr. Alexander Fraser <i>Teaching:</i> Lab Courses, Bachelors, and Masters theses supervision; NLP lectures. <i>Research Projects:</i> eXplainable AI (XAI) for NLP models; Safe and Fair Human-written and Machine-generated Texts; Multilingual and Ukrainian NLP.</p> <p>Research Engineer Skoltech Moscow Mar 2021 - Mar 2022 Leading the <i>Text Detoxification</i> project connected with PhD studies. The results of the research were formed as 2 ACL papers, 1 journal paper, RUSSE'22 shared task, and were deployed into MTS.AI chatbot.</p> <p>Data Science Intern Beiersdorf Hamburg Jun 2018 - Aug 2018 Concept design and prototype development of the automatic trends extractor from texts from social networks.</p> <p>Data Scientist Visiology Moscow Aug 2017 - Aug 2019 <i>NLP2SQL:</i> Chat-bot assistant for Business Intelligence. The chat-bot was presented at Business Intelligence 2.0 conference (1.5k attendees) as the first Russian BI chat-bot and was deployed as ViTalk.</p>
SELECTED PUBLICATIONS (GOOGLE SCHOLAR)	<p>D Dementieva, N Babakov, A Panchenko, "MultiParaDetox: Extending Text Detoxification with Parallel Data to New Languages", <i>NAACL</i> 2024</p> <p>D Dementieva, V Khylenko, N Babakov, G Groh, "Toxicity Classification in Ukrainian", <i>WOAH, NACCL</i> 2024</p> <p>E Mosca, D Dementieva, T Ajdari, M Kummeth, K Gringauz, Y Zhou, G Groh, "IFAN:</p>

An Explainability-Focused Interaction Framework for Humans and NLP Models”, *AACL* 2023

D Dementieva, M Kuimov, A Panchenko, “[Multiverse: Multilingual Evidence for Fake News Detection](#)”, *Journal of Imaging*, 2023

V Logacheva*, **D Dementieva***, S Ustyantsev, D Moskovskiy, D Dale, I Krotova, N Semenov, A Panchenko, “[ParaDetox: Detoxification with Parallel Data](#)”, *ACL*, 2022

D Dale, A Voronov, **D Dementieva**, V Logacheva, O Kozlova, N Semenov, A Panchenko, “[Text Detoxification using Large Pre-trained Neural Models](#)”, *EMNLP*, 2021

D Dementieva, D Moskovskiy, V Logacheva, D Dale, O Kozlova, N Semenov, A Panchenko, “[Methods for detoxification of texts for the russian language](#)”, *MDPI*, 2021

E Kasneci, K Seßler, S Küchemann, M Bannert, **D Dementieva**, F Fischer, ... G Kasneci. “[ChatGPT for good? On opportunities and challenges of large language models for education](#)”. *Learning and individual differences*, 2023

TALKS AND POSTERS

Adam-Smith at SemEval-2023 Task 4: Discovering Human Values in Arguments with Ensembles of Transformer-based Models
15 June, 2023, [ACL 2023](#), [SemEval Workshop](#), Toronto, Canada, **talk**

AdamR at SemEval-2023 Task 10: Solving the Class Imbalance Problem in Sexism Detection with Ensemble Learning
16 June, 2023, [ACL 2023](#), [SemEval Workshop](#), Toronto, Canada, **poster**

Methods for Fighting Harmful Multilingual Information
27 June, 2023, [Mannheim NLP Group](#), Mannheim, Germany, **talk**

Methods for Text Style Transfer: Text Detoxification Case
18 April, 2023, [PyCon&PyData](#), Berlin, Germany, **talk**

Methods for Fighting Harmful Multilingual Information
1 November, 2022, [Language Technology Group](#), Hamburg, Germany **talk**

NLP for Social Good: Methods for Fighting Harmful Multilingual Textual Content
26 October, 2022, [MunichNLP](#), Germany **talk**

Methods for Text Style Transfer: Text Detoxification Case
20 October 2022, [MilanoNLP](#), online **talk**

Methods for Fighting Harmful Multilingual Information
15 June 2022, [CIS LMU](#), Munich, Germany **talk**

A Study on Manual and Automatic Evaluation for Text Style Transfer: The Case of Text Detoxification
25 May 2022, [ACL 2022 HumEval Workshop](#), Dublin, Ireland **talk**

ParaDetox: Detoxification with Parallel Data
27 May 2022, [ACL 2022](#), Dublin, Ireland, **poster**

Fake News Detection using Multilingual Evidence
July 2021, [TrustWorthy AI](#), Moscow, Russia **talk**

Virtual Assistant ViTalk or Human Business Intelligence
May 2019, [DataFest](#), Moscow, Russia **talk**

TEACHING EXPERIENCE

NLP Lab Course, XAI for Machine Learning, TUM, 2022-2024: projects supervision.
Natural Language Processing, TUM, 2022-2024: lecturer.
Statistical NLP, Neural NLP, Skoltech, 2019-2021: seminars, projects supervision.

MENTORING

Diego Miguel Lozano (Masters, 2024, TUM) *"Explainable Textual Similarity Metrics via Equivalent and Erroneous Span Detection"*

Arda Andirin (Masters, 2024, TUM) *"Explainable Fact-Checking in News Datasets with Large Language Models"*

Shaghayegh Koli (Masters, 2024, TUM) *"Fake News and Machine-generated Text Detection for Persian and English"* (paper under review)

Valeriia Khylenko (Masters, 2024, TUM) *"Cross-lingual Text Classification Knowledge Transfer to Ukrainian language"* ([paper](#))

Adam Rydelek (Masters, 2023, TUM) *"Sexism in Social Media – Explainable Detection Using Deep Learning Models"* ([paper](#))

Daniel Schroter (Masters, 2023, TUM) *"Discovering Human Values Behind Arguments – A Case for Transformer-based Models and Explainable AI"* ([paper](#), the shared task winner)

Yen-Yu Chang (Masters, 2023, TUM) *"Exploration of Approaches to Counter Hate Speech: The Case of Sexist Speech"*

Yuliana Poliakova (Bachelor, 2023, TUM) *"Detecting the Persuasion Techniques in Online News"*

Daniil Moskovsky (Masters, 2022, Skoltech) *"Supervised Text Detoxification Methods"* ([paper](#))

Mikhail Kuimov (Masters, 2022, Skoltech) *"Multilingual and Cross-lingual Approaches to Processing of News Texts"* ([paper](#))

OPENSOURCE CONTRIBUTION ([GITHUB](#)) ([HUGGINGFACE](#))

One of the admins of [s-nlp](#) with contributions:

- the first of its kind [En-](#) and [Ru-ParaDetox](#) datasets for text detoxification;
- English [roberta.toxicity_classifier](#) and Russian [russian.toxicity_classifier](#);
- SOTA text detoxification models for [English](#) and [Russian](#);
- SOTA multilingual formality [classifier](#).

All research about Text Detoxification project summarized at the [github](#) page.

Establisher of [tum-nlp](#) with contributions:

- SOTA human values [detector](#);
- target-aware counter speech [generator](#);
- sexism speech [detector](#).

Establisher of [textdetox](#) (Multilingual Text Detoxification Project) and [ukr-detect](#) (Ukrainian Texts Classification Project).

All together >500k downloads.

PUBLIC SERVICE

Mentor in [Women In NLP, Munich](#) and [CyberMentor](#) to support young women in STEM education.

REVIEWER	<p>ARR, Area Chair: June 2024</p> <p>ARR: February, 2024, October 2023 (exceptional reviewer), August 2023, June 2023</p> <p>ACL: 2023, 2022, 2021, 2020</p> <p>EMNLP: 2024, 2023, 2022, 2021, 2020</p> <p>LREC: 2024</p> <p>ECAI: 2024</p> <p>ESWA: 2024</p> <p>WOAH: 2024, 2023</p> <p>HCI: 2024, 2023</p> <p>VLDB: 2023, 2021</p> <p>DSAA: 2020</p>
WORKSHOPS & SHARED TASKS ORGANIZER	<p>(<i>upcoming</i>) 3rd Workshop on NLP for Positive Impact, EMNLP 2024</p> <p>Multilingual Text Detoxification, CLEF 2024 (lead)</p> <p>Russian First Text Detoxification Shared Task, Dialogue 2022 (lead)</p>
SUMMER& WINTER SCHOOLS	<p>ELLIS Summer School on Large-Scale AI for Research and Industry, Modena, Italy, September 2023</p> <p>Advanced Language Processing Winter School (ALPS), Grenoble, France, January 2022</p> <p>Advanced Language Processing Winter School (ALPS), Grenoble (online), France, January 2021</p> <p>The Machine Learning Summer School (MLSS), Tübingen (online), Germany, June 2020</p>
SELECTED HONORS & AWARDS	<p>11th Heidelberg Laureate Forum attendee with all costs covered with Abbe Grant from the Carl-Zeiss-Stiftung.</p> <p>Friedrich Schiedel Fellowship for Technology in Society: 20 000\$ from TUM to explore "<i>Harmful Speech Proactive Moderation</i>". (video pitch)</p> <p>Joint India-Germany Research Grant, 2023: topic "<i>Deep neural multilingual models to combat online hate content</i>"; funded stay at Indian Institute of Technology, Kharagpur in Animesh Mukherjee's group.</p> <p>Skoltech Highest Scholarship 2022-2023 for the high research achievements. (top 1%)</p> <p>Junction 2021, Empathy Challenge: <i>third prize</i>. The adaptation of text detoxification research results for a gaming industry use-case.</p> <p>Junction 2019, Ericsson Challenge: <i>second prize</i>. The algorithm development for optimal client-taxi matching and transportation around the city.</p> <p>P&G 2018 IT Business Challenge: grant <i>winner</i> to represent East Europe region at the in-person 5-day challenge in Athens, Greece.</p> <p>Innopolis 2015 Programming Contest: <i>winner</i>; educational grant award. (top 1%)</p>
INTERESTS	athletics, yachting, (wind)surfing, skiing, psychology, feminism
LANGUAGES	<p>Ukrainian: native</p> <p>Russian: native</p> <p>English: professional</p> <p>German: intermediate</p>