Dr. Daryna Dementieva

daryna.dementieva@tum.de, +4915172462384
Personal page: https://dardem.github.io, Google Scholar
http://www.linkedin.com/in/daryna-dementieva

Computation, Information and Technology
Technical University of Munich (TUM)
Garching, Germany

RESEARCH INTERESTS

natural language processing for social good; harmful content mitigation; multilingual, fair, and low-resource AI; explainable AI; responsible AI

EDUCATION

November 2019 – September 2022
Ph.D. in Computer Science
Skolkovo Institute of Science and Technology (Skoltech), Moscow
Thesis: Methods for Fighting Harmful Multilingual Information (*defended online*)
Supervisor: Ass. Prof. Alexander Panchenko
Links: information and video of the defense

September 2017 – August 2019
Master of Science in Applied Math and Physics
Moscow Institute of Science and Technology (MIPT), Moscow
Thesis: News Aggregation and Personalization for Professional Groups
Supervisor: Prof. Konstantin Vorontsov

September 2013 – August 2017
Bachelor of Science in Applied Math
Taras Shevchenko's Kyiv National University (KNU), Kyiv

EXPERIENCE

Postdoctoral Researcher — TUM
Jun 2022 - now — Munich
Host: Prof. Dr. Alexander Fraser (*since June 2022, Prof. Dr. Georg Groh*)
Teaching: Lab Courses for Natural Language Processing and Large Language Models; Responsible AI and Language Technologies Seminar; Bachelors, and Masters theses supervision; NLP lectures.
Research Projects: Safe and Fair Human-written and Machine-generated Texts; Hate and Toxic Speech Mitigation; eXplainable AI (XAI); Multilingual, Low-resource, and Ukrainian NLP.

Research Engineer — Skoltech
Mar 2021 - Mar 2022 — Moscow
Leading the *Text Detoxification* 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.

Data Science Intern — Beiersdorf
Jun 2018 - Aug 2018 — Hamburg
Concept design and prototype development of the automatic trends extractor from texts from social networks.

Data Scientist — Visiology
Aug 2017 - Aug 2019 — Moscow
NLP2SQL: 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.

TALKS AND POSTERS

Multilingual and Responsible Language Technologies for Social Good
Future of NLP, HITS, Heidelberg, Germany **talk**

EmoBench-UA: A Benchmark Dataset for Emotion Detection in Ukrainian

November, 2025, EMNLP Main Conference 2025, Suzhou, China **poster**

HATEPRISM: Policies, platforms, and research integration. Advancing NLP for hate speech proactive mitigation
August, 2025, ACL Main Conference 2025, Vienna, Austria **poster**

Methods for Multilingual Harmful Speech Mitigation
June, 2025, 2025 EALM workshop - Ethic and Alignment of (Large) Language Models, Marseille, France **keynote**

Methods for Multilingual Harmful Speech Mitigation
March, 2025, The Alan Turing Institute Seminar, London, UK, talk

Language Technologies for All, UNESCO
February, 2025, LT4ALL 2025, Paris, France, poster

Multilingual and Explainable Text Detoxification with Parallel Data
January, 2025, COLING 2025, Abu-Dhabi, UAE, talk

Cross-lingual Text Classification Transfer: The Case of Ukrainian
January, 2025, COLING 2025, Abu-Dhabi, UAE, talk

The 3rd Workshop on NLP for Positive Impact
November, 2024, EMNLP 2024, Miami, USA, panel

MultiParaDetox: Extending Text Detoxification with Parallel Data to New Languages
June, 2024, NAACL 2024, Mexico City, Mexico, poster

How to Do Monolingual, Multilingual, and Cross-lingual Text Classification in April, 2024
April, 2024, PyCon&PyData, Berlin, Germany, talk

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, Iteland **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

Bachelors NLP Lab Course, TUM, 2024-2025: organizer.
Masters 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.

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 $>2M$ downloads.

PUBLIC SERVICE

Organizer of [MunichNLP]() Community and Meetups. Mentor in [Women In NLP, Munich]() and [CyberMentor]() to support young women in STEM education.

REVIEWER

ARR, Area Chair: since June 2024
ARR: February, 2024, October 2023 ([exceptional reviewer]()), August 2023, June 2023
ACL: 2023, 2022, 2021, 2020
EMNLP: 2024, 2023, 2022, 2021, 2020
LREC: 2024
ECAI: 2024
ESWA: 2024
WOAH: 2024, 2023
HCI: 2024, 2023
VLDB: 2023, 2021
DSAA: 2020

WORKSHOPS & SHARED TASKS ORGANIZER

[LLMs with Limited Resources for Slavic Languages](), WMT, EMNLP 2025 (organizer)

[4rd Workshop on NLP for Positive Impact](), ACL 2025 (lead)

[3rd Workshop on NLP for Positive Impact](#), EMNLP 2024 (organizer)

[Multilingual Text Detoxification Shared Task](#), CLEF 2024 & 2025 (lead)

[Russian First Text Detoxification Shared Task](#), Dialogue 2022 (lead)

SUMMER&
WINTER
SCHOOLS
[ELLIS Summer School on Large-Scale AI for Research and Industry](#), Modena, Italy, September 2023
[Advanced Language Processing Winter School](#) (ALPS), Grenoble, France, January 2022
[Advanced Language Processing Winter School](#) (ALPS), Grenoble (online), France, January 2021
[The Machine Learning Summer School](#) (MLSS), Tübingen (online), Germany, June 2020

SELECTED
HONORS
&AWARDS
[Featured](#) in the Women in AI series (February 2025) for contributions to research on combating fake news and mitigating hate speech.

11th Heidelberg Laureate Forum attendee with all costs covered with Abbe Grant from the [Carl-Zeiss-Stiftung](#). (top 1%)

Friedrich Schiedel Fellowship for Technology in Society: 20 000$ from TUM to explore "*Harmful Speech Proactive Moderation*". ([video pitch](#))

Joint India-Germany Research Grant, 2023: topic *"Deep neural multilingual models to combat online hate content"*; funded stay at [Indian Institute of Technology, Kharagpur](#) in [Animesh Mukherjee](#)'s group.

Toloka.ai Research Grant: 3 000$ for various multilingual and cross-lingual annotation projects.

Skoltech Highest Scholarship 2022-2023 for the high research achievements. (top 1%)

Junction 2021, Empathy Challenge: *third prize*. The adaptation of text detoxification research results for a gaming industry use-case.

Junction 2019, Ericsson Challenge: *second prize*. The algorithm development for optimal client-taxi matching and transportation around the city.

P&G 2018 IT Business Challenge: grant *winner* to represent East Europe region at the in-person 5-day challenge in Athens, Greece.

Innopolis 2015 Programming Contest: *winner*; educational grant award. (top 1%)

INTERESTS
athletics, yachting, (wind)surfing, skiing, psychology, feminism

LANGUAGES
Ukrainian: native
Russian: native
English: professional
German: intermediate

Interests
Reading, Pilates&Yoga, Swimming, Windsurfing, Politics&Feminism,
Traveling (Countries visited: 31)