PhD in Computational and Data Science and Engineering



Hamburg, Germany



+49 15203257163



Universität Hamburg webpage



Google Scholar



Personal page



irina. nikishina @uni-hamburg.de

Research interests —

Question Answering, Argument Mining, Retrieval Augmented Generation, Lexical Semantics, Taxonomies, Taxonomy Enrichment, Detoxification, Relation Extraction, Entity Linking, Entity Disambiguation, Linguistics

Languages -

Russian (native)

English (C1)

French (C1)

German (A2)

Computer skills —

Python, PyTorch, HuggingFace, Docker, LaTeX, Java, SQL

Personal interests —

traveling, museums, art galleries, painting, handmade beaded jewelry, France, French culture, foreign languages, gym, ping-pong, snowboarding

Education

2019-2022 **PhD**

Skolkovo Institute of Science and Technology, Moscow, Russia

Computational and Data Science and Engineering

thesis: "Learning Linguistic Tree Structures with Text and Graph Methods"

advisor: Alexander Panchenko

links: text

2017-2019 Master's degree (with honors)

NRU HSE, Moscow, Russia

Computational linguistics

thesis: "Entity Disambiguation to Wikipedia for Languages with Different Corpora

Volumes"

advisor: Anastasiya A. Bonch-Osmolovskaya

Java and Oracle courses

NetCracker, Nizhny Novgorod, Russia

Oracle, Java, OOP, MapReduce, Postgres, Spring, ONOS, GWT, Ajax

final project: "Developing a T-SDN controller tester (using selenium, requests,

openpyxl, pyyaml Python libraries" advisor: Dmitry Amosov

2016 Exchange Program

Franche-Comte University, France

5th semester spent at Franche-Comte University, courses language: French internship project: "Developing a Python tool for interpreting queries in Relational

Algebra"

advisor: Frederic Dadeau

2013-2017 Bachelor's degree (with honors)

NRU HSE, Nizhny Novgorod, Russia

Fundamental and applied linguistics

thesis: "Automatic Information Extraction about Persons from News Articles"

advisors: Irina Utkina and Alexey Malafeev

links: text

Work Experience

Aug 2022-Feb 2025

Universität Hamburg

Hamburg, Germany

postdoc at Language Technology group on the ACQuA-2.0 project: "Development of the Comparative Question Answering system"

hosted by Chris Biemann

2021-2022

Skoltech

Moscow, Russia

Research Engineer (Python)

project: "Text Detoxification" (dataset annotation and annotation verification; technical support for the RUSSE'2022 shared task; help with writing research articles on the shared task)

Jan-May 2021

Innoscripta

St. Petersburg, Russia

Data Scientist, Python NLP engineer

Developing system for keyword search, improve NER for extracting keywords and German product names, management of keyword annotation in Doccano for the

German language

Sep 2017-2019

ISP RAS

Moscow, Russia

research assistant in Python and Java in the "Texterra" team

NLP research in Entity Linking, writing research articles

supervisors: Andrey Sysoev and Ivan Andrianov

Aug-Sep 2018

MrBot

Moscow, Russia

summer internship

project: "developing dialogue system architecture in Python"

Franche-Comte University

Jun-Aug 2017

NetCracker

Nizhny Novgorod, Russia

junior software engineer

developing a T-SDN controller automatic testing system (Python: selenium, requests,

openpyxl, pyyaml)

May 2016

Besancon, France

summer internship

project: "Developing a Python tool for interpreting queries in Relational Algebra"

advisor: Frederic Dadeau

PhD in Computational and Data Science and Engineering



Hamburg, Germany



+49 15203257163



Universität Hamburg webpage



Personal page



Google Scholar



irina.nikishina@uni-hamburg.de

Student Supervision

12. Viktor Moskvoretskii (Apr 2023—now, HSE, online)

Project on enhancing LLM with Taxonomic structures

led to the paper accepted to LREC-COLING 2024 "Are Large Language Models Good at Lexical Semantics? A Case of Taxonomy Learning", to ACL 2024 "TaxoLLaMA: WordNet-based Model for Solving Multiple Lexical Sematic Tasks" and to "Low-Resource Machine Translation through the Lens of Personalized Federated Learning"

11. **Ahmad Shallouf** (Jan 2024—now, Universität Hamburg, GE)

MSc thesis supervision: "CompGLUE: Datasets and Evaluation System for Comparative Question Answering Tasks"

led to the paper accepted to LREC-COLING 2024 : "CAM 2.0: End-to-End Open Domain Comparative Question Answering System"

10. Angel Ontiveros (Jan-July 2024)

MSc thesis supervision: "Development and Evaluation of an IT Tool for ESG Compliance Consulting for Companies Using LLMs"

9. **Jan Strich** (Jan 2024—now, Universität Hamburg, GE)

MSc thesis co-supervision with Florian Schneider: "Improving Large Language Models in Repository Level Programming through Self-Alignment and Retrieval-Augmented Generation"

8. Maria Manina, Daria Ignatenko, Nikolay Dolgov (Sep 2023—now, HSE, online)

MSc student project "Benchmark system for Comparative Question Answering"

7. Alina Rechina (Sep 2023—now, HSE, online)

consultant for the BSc student thesis with Viktor Moskvoretskii: "Multilingual Large Language Models for predicting IS-A relationships"

led to preprint: "TaxoLLaMA: WordNet-based Model for Solving Multiple Lexical Sematic Tasks"

6. Ekaterina Neminova (Sep 2023—now, HSE, online)

consultant for the BSc student thesis Viktor Moskvoretskii: "Investigation of Large Language Models for Taxonomy-related Tasks"

led to preprint: "TaxoLLaMA: WordNet-based Model for Solving Multiple Lexical Sematic Tasks"

 Ahmad Shallouf, Hannah Herasimchyk, Rudy Garrido Veliz, Natia Mestvirishvii (Summer 2023, Universität Hamburg, GE)

MSc student project supervision

led to the paper accepted to LREC-COLING 2024 : "CAM 2.0: End-to-End Open Domain Comparative Question Answering System"

4. Maria Maslova, Stefan Rebrikov (2022-2023, HSE, online)

MSc student project supervision "Russian Comparative Argumentative Machine" led to paper accepted to AIST-2023: "RuCAM: Russian Comparative Argumentative Machine"

3. Polina Chernomorchenko (2022-2023, HSE, online)

BSc student thesis supervision: "Automatic Taxonomy Enrichment Using Generative Models" led to 2 papers published at AACL-2023 and AIST-2023 conferences: "Predicting Terms in IS-A Relations with Pre-trained Transformers" and "Leveraging Taxonomic Information from Large Language Models for Hyponymy Prediction"

2. Alsu Vakhitova (2021-2022, Skoltech, RU)

mentoring MSc student thesis "Candidate-free taxonomy enrichment with pre-trained neural language models"

led to 2 papers published on AACL-2022 and TextGraph-16 Workshop: "TaxFree: a Visualization Tool for Candidate-free Taxonomy Enrichment" and "Cross-modal contextualized hidden state projection method for expanding of taxonomic graphs"

1. Anna Safaryan, Petr Filchenkov, Weijia Yan (2020-2021, HSE, online)

MSc student project co-supervision with Andrey Kutuzov

led to a publication on AIST-2020 "Semantic Recommendation System for Bilingual Corpus of Academic Papers"

PhD in Computational and Data Science and Engineering



Hamburg, Germany



+49 15203257163



Universität Hamburg webpage



Personal page



Google Scholar



irina.nikishina@uni-hamburg.de

Teaching experience

Dec 2023 Lecture on Encoder and Decoder Transformers (en)

role: online lecture recording for the AI & DT school on Transformers,

online

duration: 1 hour

audience: Generative AI school participants

Summer 2023 Web interfaces for Language Technology Systems Hamburg, GE

role: Teaching Assistant duration: 3 months

audience: MSc students in Computer Science

2022-2024 Course on Transformers (en & ru) online

role: Lecturer and Seminarist

duration: 2 months

audience: MSc students & Data Scientists in the Industry

repeated: 6 times links: Github

Neural Natural Language Processing (en)

NRU HSE, online

role: Lecturer and Seminarist duration: Terms 1 and 2

audience: 4th year BSc students in Fundamental and Applied Linguistics

links: Github

2021-2022 Machine Learning 2 (en) NRU HSE, RU

role: Seminarist

duration: Terms 2 and 3

audience: 4th year BSc students in Data Science and Business Analytics

Neural Natural Language Processing (en) Skoltech, RU

role: Teaching Assistant duration: Term 4

audience: MSc students in Computer Scince

repeated: 2 runs

2020-2022 Statistical Natural Language Processing (en) Skoltech, RU

Teaching Assistant duration: Term 2

audience: MSc students in Computer Scince

repeated: 3 runs

2020-2022 Applied Research in Culture (ru) NRU HSE, RU

role: Lecturer and Seminarist

duration: Term 1

audience: MSc students in Applied Cultural Studies

repeated: 2 runs links: Github

2020-2021 Programming and Linguistic Data (ru) NRU HSE, RU

role: Lecturer and Seminarist

duration: Terms 1-4

audience: 1st year BSc students in Applied Linguistics

links: Github

2019-2020 Digital Literacy (ru) NRU HSE, RU

role: Lecturer and Seminarist

duration: Terms 3-4

audience: 1st year BSc students in Applied Cultural Studies

repeated: 2 runs links: slides

PhD in Computational and Data Science and Engineering



Hamburg, Germany +49 15203257163



Universität Hamburg webpage



Personal page



Google Scholar



irina.nikishina@uni-hamburg.de

Publications |

- 28. Moskvoretskii, V., Tupitsa, N., Biemann, C., Horvath, S., Gorbunov, E., Nikishina, I. (2024). Low-Resource Machine Translation through the Lens of Personalized Federated Learning. arXiv preprint arXiv:2406.12564. (pdf)
- 27. Nikishina, I., Tikhonova, M., Chekalina, V., Zaytsev, A., Vazhentsev, A., Panchenko, A. (2024). Industry vs Academia: Running a Course on Transformers in Two Setups in Proceedings of the 6th Workshop on Teaching NLP, Bangkok. ACL 2024. (pdf)
- 26. Günzler, R., Sevgili, Ö., Remus, S., Biemann, C., Nikishina, I. (2024). Sövereign at PerpectiveArg2024: Using LLMs with Argument Mining in Proceedings of the 11th Workshop on Argument Mining, Bangkok. ACL 2024. (pdf)
- 25. Rouhizadeh, H., Nikishina, I., Yazdani, A., Bornet, A., Zhang, B., Ehrsam, J., Gaudet-Blavignac, C., Naderi, N., Teodoro D. (2023) BioWiC: An Evaluation Benchmark for Biomedical Concept Representation. Scientific Data 11, 455 (2024). (pdf) (Q1)
- 24. Moskvoretskii, V., Neminova, E., Lobanova, A., Panchenko, A., Nikishina, I. (2024) TaxoLLaMA: WordNet-based Model for Solving Multiple Lexical Semantic Tasks. (ACL 2024) . (pdf)
- 23. Nikishina, I., Bondarenko, A., Zaczek, S., Haag, O., Hagen, M., Biemann C. (2024) Extending the Comparative Argumentative Machine: Multilingualism and Stance Detection. RATIO 2024, Bielefeld, Germany. (pdf)
- 22. Shallouf, A., Herasimchyk, H., Salnikov, M., Veliz, R., Mestvirishvili N., Panchenko, A., Biemann, C., Nikishina, I. (2024) CAM 2.0: End-to-End Open Domain Comparative Question Answering System. The 2024 Joint International Conference on Computational Linguistics, Language and Evaluation (LREC-COLING 2024) (pdf).
- 21. Moskvoretskii, V., Panchenko, A., Nikishina, I. (2024) Are Large Language Models Good at Lexical Semantics? A Case of Taxonomy Learning. The 2024 Joint International Conference on Computational Linguistics, Language and Evaluation (LREC-COLING 2024) . (pdf)
- 20. Maslova, M., Rebrikov, S., Artsishevski, A., Zaczek, S., Biemann C., Nikishina I. (2023) RuCAM: Comparative Argumentative Machine for the Russian Language. In Proceedings of Analysis of Images, Social Networks and Texts. (AIST 2023) (pdf)
- 19. Chernomorchenko, P., Panchenko, A., Nikishina I. (2023) Leveraging Taxonomic Information from Large Language Models for Hyponymy Prediction. In Proceedings of Analysis of Images, Social Networks, and Texts. (AIST 2023) (pdf)
- 18. Salnikov, M., Le, H., Rajput, P., Nikishina, I., Braslavski, P., Malykh, V., Panchenko, A. (2023) Large Language Models Meet Knowledge Graphs to Answer Factoid Questions. Pacific Asia Conference on Language, Information and Computation (PACLIC 2023), pp. 635-644. (pdf)
- 17. Nikishina, I., Chernomorchenko, P., Demidova, A., Panchenko, A., Biemann, C. (2023) Predicting Terms in IS-A Relations with Pre-trained Transformers. In Proceedings of the 13th IJCNLP and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL Findings 2023), pp. 134-148. (pdf)
- 16. Nikishina, I., Adrianov, I., Vakhitova, A., Panchenko, A. (2022) TaxFree: a Visualization Tool for Candidate-free Taxonomy Enrichment. In: The 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing, AACL/IJCNLP (System Demonstrations), (AACL Demo 2022) pp. 39-47. (pdf)
- 15. Nikishina, I., Vakhitova, A., Tutubalina, E., Panchenko, A. (2022) Cross-Modal Contextualized Hidden State Projection Method for Expanding of Taxonomic Graphs. In: Proceedings of TextGraphs-16: Graph-based Methods for Natural Language Processing, pp. 11–24. (TextGraphs-16) . (pdf)
- 14. Kotelnikov, E., Loukashevich, N., Nikishina, I., Panchenko, A. (2022) RuArg-2022: Argument Mining Evaluation. In: Computational linguistics and intellectual technologies: papers from the annual conference "Dialogue". (Dialogue 2022) (pdf)

PhD in Computational and Data Science and Engineering



Hamburg, Germany



+49 15203257163



Universität Hamburg webpage



Personal page



Google Scholar



irina.nikishina@uni-hamburg.de

- Dementieva D., Logacheva, V., Nikishina, I., Fenogenova A., Dale, D., Krotova I., Semenov, N., Shavrina T., Panchenko, A. (2022) RUSSE-2022: Findings of the First Russian Detoxification Shared Task Based on Parallel Corpora. In: Computational linguistics and intellectual technologies: papers from the annual conference "Dialogue". (Dialogue 2022) (pdf)
- 12. Logacheva, V., Dementieva D., Krotova I., Fenogenova A., **Nikishina, I.**, Shavrina T., Panchenko, A. (2022) *A Study on Manual and Automatic Evaluation for Text Style Transfer.* In Proceedings of the 2nd Workshop on Human Evaluation of NLP Systems (HumEval 2022), pp. 90–101. (pdf)
- 11. **Nikishina, I.**, Tikhomirov M., Logacheva, V., Nazarov Yu., Panchenko, A., Loukashevich, N. (2022) *Taxonomy Enrichment with Text and Graph Vector Representations*. Semantic Web, vol. 13, no. 3, pp. 1-35. (pdf) (Q1)
- 10. Safaryan, A., Filchenkov, P., Yan, W., Kutuzov A., **Nikishina I.** (2021) *Semantic Recommendation System for Bilingual Corpus of Academic Papers*. In Proceedings of Analysis of Images, Social Networks and Texts. (AIST 2021) (pdf)
- 9. **Nikishina, I.**, Loukashevich, N., Logacheva, V., Panchenko, A. (2021) *Exploring Graph-based Representations for Taxonomy Enrichment for the Russian Language*. South African Centre for Digital Language Resources (SADiLaR) Potchefstroom, South Africa: 126. (11th Global Wordnet Conference) (pdf)
- 8. **Nikishina, I.**, Logacheva, V., Panchenko, A., Loukashevich, N. (2020) *Studying Taxonomy Enrichment on Diachronic WordNet Versions*. Proceedings of the 28th International Conference on Computational Linguistics, pp. 3095–3106. (COLING 2020) (pdf) (core A)
- 7. **Nikishina, I.**, Logacheva, V., Panchenko, A., Loukashevich, N. (2020) *RUSSE'2020: Findings of the First Taxonomy Enrichment Task for the Russian Language*. In Computational Linguistics and Intellectual Technologies: Papers from the Annual Conference "Dialogue". (Dialogue 2020) (pdf)
- Sysoev A., Nikishina I. (2019) Deep JEDi: Deep Joint Entity Disambiguation to Wikipedia. In Proceedings of Analysis of Images, Social Networks, and Texts, pp. 230–241. (AIST 2019) (pdf)
- 5. Kutuzov A., **Nikishina I.** (2019) *Double-Blind Peer-Reviewing and Inclusiveness in Russian NLP Conferences*. In Proceedings of Analysis of Images, Social Networks and Texts, pp. 3-8. (AIST 2019) (pdf)
- 4. **Nikishina I.**, Sokolova I., Tikhomirov D., Bonch-Osmolovskaya A. (2019) *Automatic Direct Speech Tagging in Russian Prose: Markup and Parser.* Dialogue Conference on Computational Linguistics and Intelligent Technologies. (Dialogue 2019) (pdf)
- 3. Sysoev A., Nikishina I. (2018) *Smart Context Generation for Disambiguation to Wikipedia*. Conference on Artificial Intelligence and Natural Language. Springer, Cham, 2018. pp. 11-22. (AIST 2018) (pdf)
- 2. **Nikishina I.**, Bakarov A., Kutuzov A. (2018) *RusNLP: Semantic Search Engine for Russian NLP Conference Papers*. In Analysis of Images, Social Networks and Texts. Lecture Notes in Computer Science, vol 11179, Springer, Cham, 2018, pp. 111-120. (AIST 2018) (pdf)
- 1. Bakarov A., Kutuzov A., **Nikishina I.** (2018) *Russian Computational Linguistics: Topical Structure in 2007-2017 Conference Papers.* Dialogue Conference on Computational Linguistics and Intelligent Technologies. (Dialogue 2017) (pdf)

Reviewing

I participated in the following reviewing processes:

- ACL (4 papers in 2024, 3 papers in 2023, 1 paper in 2020 as subreviewer)
- EMNLP (4 papers 2024, 3 papers in 2023)
- LREC-COLING 2024 (3 papers)
- HPCMASC24 (1 paper)

PhD in Computational and Data Science and Engineering



Hamburg, Germany



+49 15203257163



Universität Hamburg webpage



Personal page



Google Scholar



irina.nikishina@uni-hamburg.de

- KONVENS 2023 (1 paper)
- ICLR 2021 (1 paper as sub-reviewer)
- J Machine Learning 2021 (1 paper as sub-reviewer)
- LREC 2020 (1 paper as subreviewer)
- COLING 2020 (1 paper as subreviewer)
- NLP and Computational Social Science Workshop at EMNLP 2022 (1 paper)
- The Perspective Argument Retrieval Shared Task 2024 (1 paper)
- TextGraphs Workshop: Graph-based Algorithms for Natural Language Processing (1 paper as subreviewer in 2019, 3 papers in 2024)
- Dialogue (2 papers in 2023, 2 papers in 2020)
- AIST: International Conference on Analysis of Images, Social Networks and Texts (3 papers in 2023, 1 paper 2021)

I also reviewed BSc and MSc thesis on NLP in the following university programs:

- Universität Hamburg (3 MSc students in 2024)
- Skoltech (1 MSc student in 2024)
- Higher School of Economics (2 MSc students in 2023, 1 BSc student in 2023 1 BSc student in 2022)

Organizing experience

May-Sep 2024

AIST-2024

Bishkek, Kyrgyzstan

role: a member of AIST organizing committee

tasks: conference management, contacting authors/participants, call for papers design and distribution, sending invitations to the reviewers, website design and publishing updates on the website.

Jan-Aug 2024

TextGraphs-17

Bangkok, Thailand

role: member of TextGraphs-17 organizing committee "Graph-based Methods for Natural Language Processing" collocated with ACL-2024 in Bangkok, Thailand on August 11–16, 2024

tasks: organizing shared task in Codalab.

Sep 2023

AIST-2023

Yerevan, Armenia

role: a member of AIST organizing committee

tasks: conference management, contacting authors/participants, call for papers design and distribution, sending invitations to the reviewers, website design and publishing updates on the website.

Jan-Mar 2022

RuArg-2022 shared task

Moscow, Russia

role: organizer of Dialogue 2022 shared task on Argument Mining for the Russian Language

links: Github, CodaLab

Jan-Mar 2022

RUSSE'2022 shared task

Moscow, Russia

role: organiser of the shared task on Russian Text Detoxification Based on Parallel Corpora for the Dialogue-2022 conference

links: Github, CodaLab

Jul-Dec 2021

AIST-2021

Tbilisi, Georgia

role: member of AIST-2021 organizing committee: conference management, contacting authors/participants, call for papers design and distribution, sending invitations to the reviewers, website design and publishing updates on the website, conference local management (offline).

Jul-Oct 2020

AIST-2020

Moscow Russia

role: member of AIST organizing committee: conference management, contacting authors/participants, call for papers design and distribution, sending invitations to the reviewers, website design and publishing updates on the website.

Jan-Mar 2020

RUSSE'2020 shared task

Moscow, Russia

role: organiser of the shared task on Taxonomy enrichment for the Dialogue-2020 conference

links: Github, CodaLab

PhD in Computational and Data Science and Engineering



Hamburg, Germany



+49 15203257163



Universität Hamburg webpage



Personal page



Google Scholar



irina.nikishina@uni-hamburg.de

Invited talks

May 2024	Higher School of Economics online Guest seminar on "Taxonomy Enrichment with LLMs"							
Mar 2024	ISLE Expert Online Forum Panelist on the topic "Linguistics versus AI" links: video							
Apr 2023	NLP department, MBZUAI Guest talk on "Comparative Question Answering" and "Taxonomy Enrichment"							
Feb 2023	NLP group, University of Oslo Oslo, Norway Guest seminar on "Taxonomy Enrichment" (based on PhD thesis)							
Oct 2022	Higher School of Economics Moscow, Russia Guest seminar on "Taxonomy Enrichment" (based on the PhD thesis)							
Sep 2022	Computational Linguistics, Indiana University Indiana, USA Guest seminar on Taxonomy Enrichment (based on the PhD thesis)							
Dec 2021	AIST-2021 Tbilisi, Georgia Guest talk on "Taxonomy Enrichment with Texts and Graph Vector Representations"							
C								

(Student Conference)

Her	e are li	sted	Summer	School	s and	Co	onferences	without	papers,	posters,	or talks:
-----	----------	------	--------	--------	-------	----	------------	---------	---------	----------	-----------

Summer S	School and Conference participation				
Here are listed	Summer Schools and Conferences without papers, pos	sters, or talks:			
Feb 2023	HPLT & NLPL Winter School on Large-Scale Language Modeling and Neural Machine Translation with Web Data Skeikampen, Norwa				
Aug 2021	Oxford Machine Learning Summer School 2021	online			
Aug 2021	32st European Summer School in Logic, Language and Information	online			
Jan 2021	Advanced Language Processing Winter School 2021	online			
Aug 2020	The Summer School of Machine Learning at Skoltech (SMILES)	Moscow, Russia			
Aug 2019	31st European Summer School in Logic, Language and Information	Riga, Latvia			
July 2019	8th International Conference — Analysis of Images, Social Networks and Texts, AIST 2019	Kazan, Russia			
Oct 2017	Summer School in Digital Humanities	NRU HSE, Moscow, Russia			
Apr 2017	Student Conference of the Institute of Linguistics (SKIL)	RSUH, Moscow, Russia			
Oct 2016	"ConCorT" (Student Session)	Nizhny Novgorod, Russia			
Sep 2014	Indian Summer School in Linguistics	Moscow, Russia			
Apr 2014	"ConCorT Junior" (Student Conference)	Nizhny Novgorod, Russia			
Feb 2014	Literature: histoiry, theories, comparativistics				

NRU HSE, Moscow, Russia