

IRINA SAPARINA

Edinburgh, United Kingdom ◊ I.Saparina@sms.ed.ac.uk ◊ saparina.github.io

PROFILE

- Final-year Ph.D. student at the University of Edinburgh supervised by Prof. [Mirella Lapata](#); viva early 2026.
- Expertise in AI and NLP: code generation, reasoning, structured prediction. Developed test-time search techniques for improving code generation [7] and methods for handling ambiguity in structured tasks [1], [4].
- First-author publications include NeurIPS Spotlight [5] and EACL Oral [6]. Author of benchmarks requiring complex reasoning and multi-step problem solving [2], [5].
- Seeking a research role to develop advanced reasoning and planning capabilities in AI systems.

EDUCATION

Ph.D. student in Informatics, The University of Edinburgh Advisor: Mirella Lapata Thesis: <i>Linguistic Variability and Ambiguity in Semantic Parsing</i>	Oct 2022 – present <i>Edinburgh, UK</i>
Ph.D. student in Computer Science, HSE University (GPA: 9.5/10.0) Advisor: Anton Osokin Thesis: <i>Augmenting Semantic Parsing with Intermediate Representations and Search</i>	Nov 2020 – Aug 2022 <i>Moscow, Russia</i>
Master of Computer Science, HSE University (GPA: 9.5/10.0) Advisor: Anton Osokin Thesis: <i>Cost-sensitive training for autoregressive models</i>	Sep 2018 – Jun 2020 <i>Moscow, Russia</i>
Bachelor of Computer Science, Moscow State University (GPA: 4.9/5.0) Thesis: <i>Language-Independent Methods of Extracting Keywords in Web Texts</i>	Sep 2014 – Jun 2018 <i>Moscow, Russia</i>

WORK EXPERIENCE

Yandex Research <i>Research Assistant</i> Worked on semantic parsing and program synthesis with a focus on text-to-SQL: • Developed test-time search techniques for generating better database queries guided by external criteria [7]. • Proposed training with intermediate question representations for translating questions into database queries. Achieved SOTA execution accuracy without costly expert annotation [8].	Jul 2020 – Sep 2022 <i>Moscow, Russia</i>
Samsung Laboratory at HSE University <i>Research Assistant</i> Worked on reinforcement learning for sequence-to-sequence models: • Investigated imitation learning and cost-sensitive training to directly optimize task-specific metrics. • Designed alternative learning objectives, improving performance in machine translation and code generation [9].	Jun 2018 – Jul 2020 <i>Moscow, Russia</i>
Laboratory of Programming Technologies, Moscow State University <i>Intern</i> Developed a Python tool for analyzing social media posts: • Implemented multilingual keyword extraction using matrix factorization (SVD, ONMF). • Improved ROUGE metric by 5% on real-world data.	Jun 2017 – Sep 2017 <i>Moscow, Russia</i>

TECHNICAL SKILLS

Languages: Python (expert), SQL, C/C++
Frameworks: PyTorch (expert), HuggingFace Transformers, vLLM
Tools & Infra: Git, Docker, Linux, Slurm, Kubernetes, L^AT_EX

PUBLICATIONS

- [1] **Irina Saparina** and Mirella Lapata, “[Reasoning About Intent for Ambiguous Requests](#),” Under review, 2025.
- [2] Miao Li, Alexander Gurung, **Irina Saparina**, and Mirella Lapata, “[Who Gets Cited Most? Benchmarking Long-Context Language Models on Scientific Articles](#),” Under review, 2025.
- [3] Sergey Troshin, **Irina Saparina**, Antske Fokkens, and Vlad Niculae, “[Asking a Language Model for Diverse Responses](#),” in Workshop on Uncertainty-Aware NLP at EMNLP, 2025.
- [4] **Irina Saparina** and Mirella Lapata, “[Disambiguate First, Parse Later: Generating Interpretations for Ambiguity Resolution in Semantic Parsing](#),” in Findings of ACL, 2025.
- [5] **Irina Saparina** and Mirella Lapata, “[AMBROSIA: A Benchmark for Parsing Ambiguous Questions into Database Queries](#),” in NeurIPS Datasets and Benchmarks Track (Spotlight), 2024.
- [6] **Irina Saparina** and Mirella Lapata, “[Improving Generalization in Semantic Parsing by Increasing Natural Language Variation](#),” in EACL (Oral), 2024.
- [7] Anton Osokin, **Irina Saparina**, and Ramil Yarullin, “[Searching for Better Database Queries](#),” in Findings of EACL, 2023.
- [8] **Irina Saparina** and Anton Osokin, “[SPARQLing Database Queries from Intermediate Question Decompositions](#),” in EMNLP, 2021.
- [9] **Irina Saparina** and Anton Osokin, “[Cost-Sensitive Training for Autoregressive Models](#),” in Women in Machine Learning Workshop at NeurIPS, 2019.

AWARDS

Outstanding Reviewer at EMNLP 2024	2024
Scholarship of Advanced Doctoral Programme at HSE University	2020 - 2022
EMNLP Student Travel Grant	2021
NeurIPS Travel Grant	2019
Russian Academic Excellence Scholarship	2014-2020
Winner of SAS Data Analysis Hackathon	2017

TEACHING

Foundations of Natural Language Processing , University of Edinburgh	
Teaching Assistant	Spring 2025
Deep Learning , HSE University	
Lecturer	Spring 2022
Seminar Teacher	Fall 2020
Teaching Assistant	Spring and Fall 2019
Bayesian Methods of Machine Learning , HSE University and Moscow State University	
Teaching Assistant	Fall 2018
Summer School on Deep Learning and Bayesian Methods (Deep Bayes)	
Assistant	Summer 2018, Summer 2019

SERVICE

Reviewer: WiML @ NeurIPS 2019; EMNLP 2023; ACL ARR 2024 Feb (ACL 2024), Jun (EMNLP 2024), Feb 2025 (ACL 2025); NeurIPS 2025.

Mentor: Women in Machine Learning 2024-2025.