

CONTACT INFORMATION	<div>Vienna, Austria</div> <div>+43 681 84495288</div> <div>anastasiia.sedova@univie.ac.at</div>	<div> <a href="https://anasedova.github.io">https://anasedova.github.io</a></div> <div> @anasedova</div> <div> @anastasiia-sedova</div>
	<p>I am a fourth-year PhD student with a strong foundation in mathematics and linguistics, as well as expertise in machine learning and programming. My past projects have focused on noisy and resource-constrained learning, and analyzing behavior of language models in these settings. My current research interests include, but are not limited to, model explainability and interpretability, with the goal of improving the trustworthiness and reliability of AI systems.</p>	
EDUCATION	<div>University of Vienna</div> <div>Ph.D. in Computer Science advised by <i>Benjamin Roth</i></div> <div>University of Munich (LMU), CIS</div> <div>M.Sc. in Computational Linguistics (minor: Computer Science) <i>Cum Laude</i></div> <div><ul style="list-style-type: none"><li>GPA: 3,7 / 4,0. Thesis: <i>“Bias Correction for Weak Supervision”</i></li></ul></div> <div>University of Potsdam</div> <div>B.Sc. in Computational Linguistics, Visiting Student</div> <div><ul style="list-style-type: none"><li>GPA: 3,8 / 4,0.</li></ul></div> <div>Saint Petersburg State University</div> <div>B.Sc. in Computational Linguistics <i>Cum Laude</i></div> <div><ul style="list-style-type: none"><li>GPA: 4,0 / 4,0. Thesis: <i>“Topic Modeling based on Lemata and Lexical Constructions”</i>.</li></ul></div>	<div>10/2020 - 09/2025</div> <div>Vienna, Austria</div> <div>10/2018 – 09/2020</div> <div>Munich, Germany</div> <div>04/2016 – 09/2016</div> <div>Potsdam, Germany</div> <div>09/2013 – 08/2017</div> <div>Saint Petersburg, Russia</div>
AWARDS & SCHOLARSHIPS	<div>Full DAAD Scholarship for Excellently Qualified Graduates (30.000 €)</div> <div>DAAD; University of Munich. <i>Ca. 650 applicants yearly, 20% acceptance rate.</i></div> <div>Top-10 Outstanding Students Award (20.000 €)</div> <div>Saint Petersburg State University</div> <div>Scholarship for Outstanding Foreign Students (4.000 €)</div> <div>University of Potsdam</div>	<div>10/2018 – 09/2020</div> <div>09/2014 – 08/2017</div> <div>04/2016 – 09/2016</div>
ACADEMIC EXPERIENCE	<div>University of Vienna</div> <div><i>I conduct research as a member of the Data Mining and Machine Learning research group, supervise BSc/MSc theses, and teach Bachelor &amp; Master courses.</i></div> <div>Reviewer</div> <div>ACL, EMNLP (2024: outstanding reviewer), NAACL, EACL, Coling, ECML, IEEE PAMI, etc.</div>	<div>04/2021 - present</div> <div>09/2021 - present</div>
PROFESSIONAL EXPERIENCE	<div>Google Summer of Code</div> <div>Intern @DeepPavlov</div> <div><i>I implemented document-level relation extraction with automatic entity recognition and triplet extraction using Wikidata for English and Russian languages in DeepPavlov, an open-source conversational AI library. [GitHub Gist]</i></div> <div>Center for Information and Language Processing (LMU Munich)</div> <div>Student Research Assistant</div> <div><i>I was working in Hinrich Schütze NLP research group on developing methods for learning high-quality embeddings for rare words considering not only a word's surface form, but also relevant contexts.</i></div> <div>Deutsche Telekom</div> <div>Regression Test Lead</div> <div><i>Regression team management. Software Acceptance Testing, Functional, UI, Manual &amp; Automated Testing. High- and low-level test cases analysis &amp; planning, risk analysis. SQL, XML, XSLT, Jenkins, TestBench, Jira, Confluence, Agile.</i></div> <div>ABBY</div> <div>Research Intern</div> <div><i>I participated in Open Corpora project - the most extensive annotated corpora for the Russian language (at that time).</i></div> <div>Institute for Linguistics</div> <div>Research Assistant</div> <div><i>I explored how the tradition of Russian dictionaries and scientific works have been establishing since the 18th century.</i></div> <div>Lexicography Lab (Saint Petersburg State University)</div> <div>Research Assistant</div> <div><i>I participated in editing the regional dialects dictionaries.</i></div>	<div>05/2021 - 08/2021</div> <div>06/2019 – 12/2019</div> <div>08/2017 – 09/2018</div> <div>04/2015 – 09/2015</div> <div>01/2014 – 01/2016</div> <div>10/2013 – 04/2016</div>
INVITED TALKS & RESEARCH VISITS	<div>University of Munich</div> <div><i>Invited Tutorial: “Weak Supervision for Information Extraction from Text” @ Department of Statistics.</i></div> <div>University of Munich</div> <div><i>Research Visit to the research group of Hinrich Schütze @ Center for Information and Language Processing.</i></div> <div>Aalborg University (Copenhagen)</div> <div><i>Invited Tutorial: “Weak Supervision for Information Extraction from Text” @ Seminar “Machine Learning with Limited Data”.</i></div>	<div>Germany, 04/2024</div> <div>Germany, 02/2024</div> <div>Denmark, 09/2023</div>

**Technical University of Vienna**

*Invited Talk: “Weak Supervision: Possible Applications and Future Directions” @ NLP Seminar.*

Austria, 01/2023

**European Language Resource Coordination (ELRC)**

*Invited Presentation of Knodle framework @ ELRC Workshop.*

Austria, 11/2021

**Munich NLP Meetup**

*Invited Talk: “Knowledge-supervised Deep Learning” @ Knowledge Extraction & Knowledge-Supervised DL Meetup.*

Germany, 07/2021

**Bosch**

*Invited Talk: “Knodle: Modular Weakly Supervised Learning with PyTorch” @ NLP and Semantic Reasoning Lab.*

Germany, 05/2021

SELECTED  
PUBLICATIONS

**Sedova A.**, Litschko R., Frassinelli D., Roth B., Plank B. (2024) *To Know or Not To Know? Analyzing Self-Consistency of Large Language Models under Ambiguity* (EMNLP’24 Findings).

Kougia V., **Sedova A.**, Stephan A., Roth B. (2024) *Analysing zero-shot temporal relation extraction on clinical notes using temporal consistency*. BioNLP Workshop @ The 62nd Annual Meeting of the Association for Computational Linguistics (ACL’24).

Xia Y., **Sedova A.**, Araujo PHL, Kougia V., Nußbaumer L., Roth B. (2024). *Exploring prompts to elicit memorization in masked language model-based named entity recognition*. ArXiv preprint.

**Sedova A.**, Roth B. (2023). *ULF: Unsupervised Labeling Function Correction using Cross-Validation for Weak Supervision*. The 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP’23).

**Sedova A.**, Zellinger L., Roth B. (2023). *Learning with Noisy Labels by Adaptive Gradient-Based Outlier Removal*. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD’23).

**Sedova A.**, Roth B. (2023). *ACTC: Active Threshold Calibration for Cold-Start Knowledge Graph Completion*. The 61st Annual Meeting of the Association for Computational Linguistics (ACL’23).

**Sedova A.**, Stephan A., Speranskaya M., Roth B. (2021). *Knodle: Modular Weakly Supervised Learning with PyTorch*. Workshop on Representation Learning for NLP (RepL4NLP) @ The 59th Annual Meeting of the Association for Computational Linguistics (ACL’21).

Mitrofanova O., **Sedova A.** (2017). *Topic Modelling in Parallel and Comparable Fiction Texts (the case study of English and Russian prose)*. XIX International Conference “Internet and Modern Society” (IMS’17).

**Sedova A.**, Mitrofanova O. (2017). *Topic Modelling of Russian Texts based on Lemmata and Lexical Constructions*. “Information Technology and Computational Linguistics” (ITCL’17).

TEACHING  
EXPERIENCE

*In the **Data Analysis Project**, I lead students working on semester projects (e.g., proposing topics, drafting proposals, supplying literature, designing experiments, preparing possible follow-up works, etc.).*

**Winter Term 2024/2025**

- Database Systems (MSc).
- Open-Source Large Language Models (MSc).

**Summer Term 2024**

- Deep Learning for Natural Language Processing (MSc).
- Data Analysis Project (MSc): (1) Multilingual Parser for Under-resourced Languages; (2) Language Similarity by Multilingual Parsing Performance.
- Introduction to programming (BSc).

**Winter Term 2023/2024**

- Capabilities and Limitations of language-based AI models (MSc).
- Introduction to programming (BSc).
- Data Analysis Project (MSc): (1) A blessing or a curse? Analysis of scholarly citations, (2) Cross-metric evaluation of fairness optimization.

**Summer Term 2023**

- Data Analysis Project (MSc): (1) Recent Approaches to Authorship Attribution; (2) Authorship Attribution Dataset based on the Reddit Threads; (3) Counterfactual Reasoning with KG & Box Embeddings; (4) Prompt-based Temporal Relation Extraction.

**Winter Term 2022/2023**

- Computational Linguistics (MSc)
- Data Analysis Project (MSc): (1) Weak Supervision in Medical Domain; (2) Incomplete Schema Clustering for the Relation Extraction.

**Summer Term 2022**

- Topics in Deep Learning and Natural Language Processing (MSc)
- Scientific Data Management (MSc)

**Winter Term 2021/2022**

- Text Mining for Non-Computer Scientists on the Example of Discourse Analysis (MSc)
- Data Analysis Project (MSc): (1) Extraction of Structured Information using Weak Supervision; (2) Wiener Diarium: Finding Toponyms in Obituaries; (3) Automatic Identifying Civilians Killed by Police.

NON-DEGREE  
STUDIES

**Advanced Language Processing School (ALPS'22)**

01/2022

*ALPS is co-organized by LIG (Univ. Grenoble Alpes), Naver Labs Europe, and Cohere, and consists of a week-long series of NLP lectures, seminars, student projects, etc. 50 participants were selected on a competitive basis.*

**Oxford Machine Learning Summer School (OxML'21)**

07/2021 - 08/2021

*OxML is co-organized by AI for Global Goals, CIFAR and the University of Oxford and includes tracks on representation learning, NLP, statistical ML, and ML in healthcare. 400 participants selected on a competitive basis; 6.5% acceptance rate.*

**Summer School in Informational Retrieval (RussIR'17)**

08/2017

*The program covered various topics in information retrieval and cognate disciplines focusing on deep learning applied to texts documents, conversational data, images, and video. 91 students were selected on a competitive basis.*

TECHNICAL  
SKILLS

- *Programming Languages:* Python (excellent) / C#, R, SQL, Bash (familiar with).
- *Toolkits:* PyTorch, HuggingFace, TensorFlow, Keras, Conda, Pandas, Matplotlib, SciPy, Spacy, NLTK, Sklearn, OpenCV.
- *Others:* HTML/XML, UNIX, Jupyter Notebook, Slurm, Git, GitHub, GitHub Actions.

LANGUAGES

English (fluent; IELTS - 7.5), German (fluent; TestDaF - C1), French (conversational), Latin & Ancient Greek (with dictionary), Russian (native).