

CONTACT INFORMATION	Kolingasse 14-16, room 05.18 1090 Vienna, Austria <a href="tel:+4368184495288">+43 681 84495288</a>	✉ <a href="mailto:anastasiia.sedova@univie.ac.at">anastasiia.sedova@univie.ac.at</a> 👉 <a href="https://anasedova.github.io">https://anasedova.github.io</a> 🐙 @anasedova 🌐 @anastasiia-sedova-82340118b
EDUCATION	<p><b>Ph.D. in Computer Science</b> <span style="float: right;"><b>10/2020 - now</b></span> Data Mining &amp; Machine Learning Research Group, University of Vienna <span style="float: right;">Vienna, Austria</span></p> <ul style="list-style-type: none"> <li>• Advised by <i>Prof.Dr.Benjamin Roth</i>.</li> <li>• Research interests: data-centric approaches in NLP, noisy &amp; low-resource learning, model interpretability, model trustworthiness.</li> </ul> <p><b>M.Sc. in Computational Linguistics (minor: CS)</b> <span style="float: right;"><b>10/2018 – 09/2020</b></span> CIS, University of Munich (LMU) <span style="float: right;">Munich, Germany</span></p> <ul style="list-style-type: none"> <li>• GPA: 3,7 / 4,0. Thesis: <i>“Bias Correction for Weak Supervision”</i>.</li> <li>• Selected Courses: Knowledge Discovery and Data Mining, Machine Learning, Deep Learning for NLP, Machine Translation, Finite Automata.</li> </ul> <p><b>B.Sc. in Computational Linguistics, Visiting Student</b> <span style="float: right;"><b>04/2016 – 09/2016</b></span> University of Potsdam <span style="float: right;">Potsdam, Germany</span></p> <p><b>B.Sc. in Computational Linguistics <i>Cum Laude</i></b> <span style="float: right;"><b>09/2013 – 08/2017</b></span> Saint Petersburg State University <span style="float: right;">Saint Petersburg, Russia</span></p> <ul style="list-style-type: none"> <li>• GPA: 4,0 / 4,0. Thesis: <i>“Topic Modeling based on Lemata and Lexical Constructions”</i>.</li> <li>• Selected Courses: Mathematical Models of Language, Constructive Syntax, Structural Linguistics, Formal Semantics, Higher Mathematics, Probability Theory, Advanced Programming.</li> </ul>	
AWARDS & SCHOLARSHIPS	<p><b>Full DAAD Scholarship for Master Study</b> <span style="float: right;">10/2018 – 09/2020</span> <i>A full scholarship for excellently qualified graduates awarded by the association of German institutions of higher education (DAAD); ca. 650 applicants yearly, 20% acceptance rate.</i></p> <p><b>Top-10 Outstanding Students Award</b> (Saint Petersburg State University) <span style="float: right;">09/2014 – 08/2017</span></p> <p><b>Scholarship for Outstanding Foreign Students</b> (University of Potsdam) <span style="float: right;">04/2016 – 09/2016</span></p>	
ACADEMIC EXPERIENCE	<p><b>University of Vienna</b> <span style="float: right;">04/2021 - present</span> <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></p> <p><b>Reviewer</b> <span style="float: right;">09/2021 - present</span> ACL, EMNLP, NAACL, EACL, ECML, IEEE PAMI, etc.</p>	
PROFESSIONAL EXPERIENCE	<p><b>Google Summer of Code</b> <span style="float: right;">05/2021 - 08/2021</span> Intern @DeepPavlov <i>DeepPavlov is an open-source conversational AI library. I implemented document-level relation extraction with automatic entity recognition and triplet extraction using Wikidata for English and Russian languages. [GitHub Gist]</i></p> <p><b>Center for Information and Language Processing (LMU Munich)</b> <span style="float: right;">06/2019 – 12/2019</span> Student Research Assistant <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></p> <p><b>Deutsche Telekom</b> <span style="float: right;">08/2017 – 09/2018</span> Regression Test Lead <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></p> <p><b>ABBY</b> <span style="float: right;">04/2015 – 09/2015</span> Research Intern <i>I participated in Open Corpora project - the most extensive annotated corpora for the Russian language (at that time).</i></p> <p><b>Institute for Linguistics (Russian Academy of Sciences)</b> <span style="float: right;">01/2014 – 01/2016</span> Research Assistant <i>I explored how the tradition of Russian dictionaries and scientific works have been establishing since the 18th century.</i></p> <p><b>Lexicography Lab (Saint Petersburg State University)</b> <span style="float: right;">10/2013 – 04/2016</span> Research Assistant <i>I participated in editing the regional dialects dictionaries.</i></p>	
INVITED TALKS & RESEARCH VISITS	<p><b>University of Munich</b> <span style="float: right;">Germany, 04/2024</span> <i>Invited Tutorial: “Weak Supervision for Information Extraction from Text” @ Department of Statistics.</i></p> <p><b>University of Munich</b> <span style="float: right;">Germany, 02/2024</span> <i>Research Visit to the research group of Hinrich Schütze and Barbara Plank @ Center for Information and Language Processing.</i></p>	

**Aalborg University (Copenhagen)** Denmark, 09/2023  
*Invited Tutorial: "Weak Supervision for Information Extraction from Text" @ Seminar "Machine Learning with Limited Data".*

**Technical University of Vienna** Austria, 01/2023  
*Invited Talk: "Weak Supervision: Possible Applications and Future Directions" @ NLP Seminar.*

**European Language Resource Coordination (ELRC)** Austria, 11/2021  
*Invited Presentation of Knodle framework @ ELRC Workshop.*

**Munich NLP Meetup** Germany, 07/2021  
*Invited Talk: "Knowledge-supervised Deep Learning" @ Knowledge Extraction & Knowledge-Supervised DL Meetup.*

**Bosch** Germany, 05/2021  
*Invited Talk: "Knodle: Modular Weakly Supervised Learning with PyTorch" @ NLP and Semantic Reasoning Lab.*

SELECTED  
PUBLICATIONS

**Sedova A.**, Litschko R., Frassinelli D., Roth B., Plank B. (2024) *Behavioral Testing: Can Large Language Models Implicitly Resolve Ambiguous Entities?* ArXiv preprint.

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 led groups of students working on semester Project (e.g., proposing topics, drafting proposals, supplying literature, providing supervision, designing experiments, considering evaluation strategies, preparing possible follow-up works, etc.).*

**Summer Term 2024**

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

**Winter Term 2023/2024**

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

**Summer Term 2023**

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

**Winter Term 2022/2023**

Computational Linguistics (MSc).  
Data Analysis Project (MSc): Weak Supervision in Medical Domain.  
Data Analysis Project (MSc): 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): Extraction of Structured Information using Weak Supervision.

Data Analysis Project (MSc): Wiener Diarium: Finding Toponyms in Obituaries.

Data Analysis Project (MSc): 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).