

Anna MOSOLOVA

+33 07 83 69 83 80 [✉ a.mosolova333@gmail.com](mailto:a.mosolova333@gmail.com) [🔗 Google Scholar](#) [🌐 anya-bel](#)

Education

Université Paris Cité October 2022 – February 2026
PhD – Natural language processing Paris, France

- Thesis title: Semi-supervised word sense and frame induction [🔗 link](#)
- Supervisors: Marie Candito (LLF, CNRS, Paris, France) and Carlos Ramisch (LIS, CNRS, Marseille, France)
- Jury: Prof. Laura Kallmeyer (Heinrich Heine Univ Düsseldorf), Prof. Emmanuel Morin (LS2N, CNRS), Prof. Benoît Crabbé (LLF, CNRS), Prof. Benoît Sagot (ALManACH, INRIA)

Université de Lorraine September 2020 – September 2022
MSc – Natural language processing (Average: 17.2/20.0) Nancy, France

- Thesis title: Machine translation for under-resourced languages of Russia
- Supervisor: Kamel Smaïli (SMarT, LORIA, Nancy, France)

Novosibirsk State University September 2016 - August 2020
BA – Fundamental and applied linguistics (Average: 5.0/5.0) Novosibirsk, Russia

- Thesis title: Reconstruction of historical word forms using deep learning
- Supervisor: Andrey Kutuzov (Language Technology Group, University of Oslo, Oslo, Norway)

Experience

ALManACH team, INRIA January 2026 – Present
Postdoctoral researcher Paris, France

- Working on evaluation and post-training of Large Language Models for specialized domains in French

LLF laboratory, Université Paris Cité October 2022 – February 2026
PhD researcher in Natural Language Processing Paris, France

- Established new SoTA result (71.4%) on unsupervised Word-in-Context task by improving token-level representations of language models using contrastive fine-tuning
- Established new SoTA result (76.0%) on a newly proposed more natural evaluation framework for Word Sense Induction
- Published papers in top-tier conferences (ACL, EACL, TALN)

Data Monsters April 2019 – September 2022
NLP Engineer Palo Alto, USA (remote, part time)

- Built an algorithm for children's speech recognition and error detection, along with a text simplification system for reading education
- Developed a natural language understanding module for a dialogue system
- Built an algorithm to compute cost-effective delivery itineraries
- Developed a recommendation system for social events
- Constructed a named entity recognition algorithm with custom entities
- Developed an automatic content moderation system
- Implemented an intent classification algorithm using active learning for unlabeled data
- Blogposts on Medium: [🔗 link](#) , [🔗 link](#) , [🔗 link](#) , [🔗 link](#)

SMarT team, LORIA March 2022 – August 2022
Intern Nancy, France

- Improved multi-task machine translation from Russian to low-resource languages of Russia using GPT-generated data
- Published a paper at the COLING 2022 workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT)

Institute of computational technologies SB RAS May 2019 – December 2019
Junior developer Novosibirsk, Russia

- Developed an automatic accentuation system for Russian with regard to individual features of the speaker within the grant for "Development and implementation of a system for multi-level analysis of poetic texts".
- Published a paper at the Data Analytics and Management in Data Intensive Domains (DAMDID) conference.

Teaching

- Introduction à la programmation en Python** | *Université Paris Cité, L3 (in French)* 2022–2024
- 30 hours of practical sessions with a group of 20 students
 - Course given with Timothée Bernard (LLF, Université Paris Cité) as lecturer
 - Course content: data structures, loops, conditional statements, debugging, main Python libraries, file manipulation, classical algorithmic tasks
- NLP projects** | *Université Paris Cité, L3 (in French) and M1 (in English)* 2024–2026
- 30 hours of supervision of 6-10 student groups working on NLP projects
 - Course given jointly with Marie Candito (LLF, Université Paris Cité), who supervised the remaining groups
 - Project topics: word sense disambiguation, auto-completion and next-word prediction, retrofitting word embeddings, search engine development, language detection, custom topics
- Exploration informatique de corpus annotés** | *Université Sorbonne Nouvelle, L3 and M1 (in French)* 2023
- 24 hours of lectures and practical sessions with a group of 20 students
 - Course given jointly with David Kletz (then at Lattice, Université Sorbonne Nouvelle), who taught the second group
 - Course content: corpus types and annotation, grammars for corpus exploration, Turing machines, finite automata, regular expressions

Additional information

Training

- École d'été en Traitement Automatique des Langues (June 2023, Marseille, France)
- Comptech Winter School in Machine Learning (January 2019, Novosibirsk, Russia)

Reviewing

- Served as a secondary reviewer at ACL 2025

Presentations

- Presented talks at industrial and scientific conferences: *DataFest Siberia 2018, CodeFest 2019, DevFest 2019, AINL 2019*

Awards

- Mistral AI hackathon, 2nd place in fine-tuning track
- Ilya Segalovich science award 2019, finalist
- HSE international contest in computational linguistics, 3rd place
- Neurohive hackathon, 4th place
- State scholarship for achievements in study & research (6 times)

Skills

Languages: French (Advanced), English (Advanced), Russian (Native)

Programming Languages: Python

Libraries: pytorch, transformers, trl, accelerate, scikit-learn, scipy, numpy, pandas, spacy, NLTK, beautifulsoup, streamlit

Development Tools: Git, WandB, MLFlow

Databases: SQL

Publications

- When Tables Go Crazy: Evaluating Multimodal Models on French Financial Documents** 2026
Virginie Moulleron, Théo Lasnier, Anna Mosolova, and Djamé Seddah
Proceedings of the 7th Financial Narrative Processing (FNP) workshop
[link](#)
- In the LLM era, Word Sense Induction remains unsolved** 2025
Anna Mosolova, Marie Candito, and Carlos Ramisch
Findings of the Association for Computational Linguistics: ACL 2025
[link](#)
- Affinement des représentations des tokens dans les modèles de langue pré-entraînés avec l'apprentissage contrastif : une étude entre modèles et entre langues** 2025
Anna Mosolova, Marie Candito, and Carlos Ramisch
Actes de la 32ème Conférence sur le Traitement Automatique des Langues Naturelles
[link](#)
- Injecting Wiktionary to improve token-level contextual representations using contrastive learning** 2024
Anna Mosolova, Marie Candito, and Carlos Ramisch
Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Short Papers)
[link](#)
- The Only Chance to Understand: Machine Translation of the Severely Endangered Low-resource Languages of Eurasia** 2022
Anna Mosolova, Kamel Smaïli
Proceedings of the Fifth Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT 2022)
[link](#)
- Algorithm of Automatic Accentuation with Respect to the Speaking Norm of a Given Author** 2020
Mosolova A.
Supplementary Proceedings of the XXII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2020)
[link in Russian](#)
- Reconstruction of historical word forms using Deep Learning** 2020
Mosolova A. V., Kutuzov A. B.
Computational Linguistics and Intellectual Technologies, Annual International Conference "Dialogue", Student session
- Component-Based Approach to Automatic Poetry Generation** 2019
Mosolova A. V., Bondarenko I. Yu., Gusev P. A., Malysheva A. D., Vodolazsky D. I., Borovikova M. N.
Computational Linguistics and Intellectual Technologies, Annual International Conference "Dialogue", Supplementary volume
- An automated poet for Russian language** 2019
Mosolova A.V., Borovikova M.N., Vodolazsky D.I.
Proceedings of 57th International scientific student conference ISSC-2019: Applied Linguistics
[link in Russian](#)
- A benchmark for estimating the quality of sentence embeddings for Russian language** 2019
Pauls A.E., Mosolova A.V., Stroganov M.S., Timasova E.K., Shugalevskaya N.B.
Proceedings of 57th International scientific student conference ISSC-2019: Informational technologies
[link in Russian](#)

Conditional Random Fields for Metaphor Detection

2018

Mosolova A., Bondarenko I., Fomin V.

Proceedings of the Workshop on Figurative Language Processing, NAACL

[link](#)

Text Augmentation for Neural Networks

2018

Mosolova A., Bondarenko I., Fomin V.

Supplementary Proceedings of the Seventh International Conference on Analysis of Images, Social Networks and Texts

[link](#)