Lena Shakurova | NLP Engineer

✉ lenashakurova.work@gmail.com   ●   🌐 https://github.com/Shakurova

NLP Engineer with a background in computational linguistics and AI with around 3 years of experience applying NLP methods to real-world challenges across different domains.

Education

Radboud University   **Nijmegen**
Master, Artificial Intelligence   *2017–present*
Advanced Research Methods · Text Mining · Statistical Machine Learning · Computer Assisted Language Learning · Information Retrieval

National Research University Higher School of Economics (NRU HSE)   **Moscow**
Bachelor, Computational Linguistics,   *2013–2017*
GPA: 8.7/10
Courses: Discrete Mathematics · Programming (Python) · Linear Algebra and Mathematical Analysis · Web-Design · Computer Tools for Linguistic Research · Programming (R) · Theory of Algorithms · Automatic Natural Language Processing · Databases · Information Retrieval · Neurolinguistics

University of Helsinki   **Helsinki**
Exchange student, Computational Linguistics,   *2015–2016*
GPA: 3.48/4.00
Courses: Programming and Natural Language Processing (NLTK) · Morphological Processing Tools (HFST)

Coursera
Additional Courses   *2015–2017*
Probability (Penn University) · Machine Leaning (Stanford University) · Divide and Conquer, Sorting and Searching, and Randomized Algorithms (Stanford University)

Technical and Personal skills

- **Programming Languages:** Python, R, Matlab
- **Web-desigh Skills:** HTML, CSS, JavaScript, Flask, Scrapy
- **Databases:** MySQL, SQLite, MongoDB, PostgreSQL, Elasticsearch
- **Machine Learning:** Sklearn, Tenserflow, WEKA
- **NLP instruments:** PRAAT, ELAN, Tomita Parser, Toloka, NLTK, LUIS
- **Other:** LaTeX, Jupyter Notebook, git, bash, Linux.
- Sentimental Analysis, Chatbots, Word sense disambiguation (WSD), Topic modeling, Named Entity Recognition (NER), Web-crawling, Text classification and clusterisation, Dupplicates detection, Sequence labelling, Information Retrieval, Automatic summarisation

Projects and Professional Experience

Professional Experience

Intern NLP researcher at TextKernel   **Amsterdam**
python ● tensorflow ● sequence labelling ● cross-lingual embeddings   *November 2018–present*
Doing research on cross-lingual word embeddings for improving the performance of CV parsing.

Freelance Chatbot developer at United Wardrobe   **Utrecht**
dialogflow ● designing conversations   *April 2018–October 2018*
Developing a chatbot for custom support.

○ **Computational Linguist at NCEM** **Moscow**
python • Sklearn • ML • NER • web-crawling • databases *August 2016–August 2017*
Developing social media monitoring system and being a core member of the artificial intelligence and computational linguistics department.

○ **Junior Web Developer at TT-Consult** **Freelance, Moscow**
HTML • PHP • SEO • web crawlers *June 2014–August 2016*
Editing content, debugging code, re-designing web page and doing Search Engine Optimisation.

Projects

○ **Automatic language test generator**
Web app that automatically generates language tests for practicing Russian.
○ **Telegram chatbot for data annotation**
The chatbot interacts with users and by asking simple questions collects data annotations that can later be used to train machine learning models for text mining.

Additional Information

Language Skills

○ **Native:** Russian
○ **Advanced:** English (C1 CEFR Level, CAE)
○ **Intermediate:** German, Dutch
○ **Basic:** Spanish, Finnish, French

Awards

○ **Orange Tulip Scholarship Russia** 2017 awardee
○ **T500 2018** - list of 500 most talented young people in the Dutch digital scene

Conferences

○ **PyData Amsterdam** **Amsterdam**
"How to expand your NLP solution to new languages using tranfer learning" *May 2019*

○ **CLIN29 Conference** **Groningen**
"Best practices for learning domain specific cross-lingual embeddings" *January 2019*

○ **Conference ConCort on Digital Humanities** **Nizhnyj Novgorod**
"Instruments of Quantitative Typological Research" *October 2016*

Publications

○ **Universal Dependencies for Russian: A New Syntactic Dependencies Tagset**, Lyashevkaya O., Droganova K., Zeman D., Alexeeva M., Gavrilova T., Mustafina N. and Shakurova E., 2016
○ **Best practices for learning domain-specific cross-lingual embeddings**, Lena Shakurova, Beata Nyari, Chao Li, Mihai Rotaru, Proceedings of Repl4NLP workshop, ACL, 2019

Volunteering

○ **President of Language Exchange Committee** at ISN Nijmegen, 2017 - 2018

Interests

Programming · Web-design · Travelling · Music · Psychology · Literature · Coffee