

Lena Shakurova | NLP Engineer

📧 github.com/Shakurova • ✉ lena@shakurova.io • 🌐 shakurova.io

Experienced NLP Engineer with a background in Computational Linguistics and AI and 3+ years of professional experience applying NLP methods to real-world challenges across different domains.

Work Experience

- **Self-employed NLP Specialist** **Netherlands**
dialogflow • designing conversations • deep learning • machine learning *April 2018–present*
Working on long-term projects as an NLP specialist, mainly doing chatbot development.
 - Did research on intent detection and named entity recognition approaches for an HR chatbot
 - Developed a prototype of a Dutch-speaking custom support chatbot, using DialogFlow and Python
- **NLP Research Engineer intern** **TextKernel, Amsterdam**
python • tensorflow • sequence labelling • transfer learning *November 2018–July 2019*
cross-lingual embeddings • deep learning
Working full-time in a document understanding team, doing research and development on multilingual CV parsing for low-resource languages.
 - Developed a universal Slavic CV parsing pipeline for 5 languages, using transfer learning and cross-lingual embeddings
 - Built a cross-lingual model that performs well in zero-shot CV parsing scenario
 - Developed guidelines and best practices for learning cross-lingual embeddings and doing transfer learning for our task of CV parsing
 - Created new intrinsic metrics for evaluating the quality of cross-lingual word embeddings
 - Improved error analysis work flow
 - Presented at two conferences and published a paper
- **NLP Engineer** **NCEM, Moscow**
python • machine learning • Sklearn • NER • web-crawling • databases *August 2016–August 2017*
Core member of AI development team working full-time on building a social media monitoring system.
 - Built several domain-specific text classifiers
 - Developed an annotation work flow and guidelines
 - Developed news content-based clusterisation system
 - Improved Named Entity Recognition
 - Worked on key word extraction
- **Junior Web Developer** **TT-Consult, Remote**
HTML • PHP • SEO • web crawlers *June 2014–August 2016*
Editing content, debugging code, re-designing the website and doing Search Engine Optimisation.

Education

- **Masters of Artificial Intelligence** **Nijmegen**
Radboud University *2017–2019*
- **Bachelor of Computational Linguistics** **Moscow**
National Research University Higher School of Economics *2013–2017*
- **Exchange student, Computational Linguistics** **Helsinki**
University of Helsinki *2015–2016*

Skills

- **Programming Languages:** Python, R, Matlab
- **ML and NLP:** NLTK, pandas, gensim, SpaCy, Sklearn, Tensorflow, WEKA, padatious
- **Word embeddings:** Word2Vec, FastText, MUSE, VecMap

- **Chatbots:** LUIS, DialogFlow
- **Data annotation:** Toloka, Tagtog
- **Web-development:** HTML, CSS, JavaScript, Flask, Scrapy
- **Databases:** MySQL, SQLite, MongoDB, PostgreSQL
- **Other:** \LaTeX , Jupyter Notebook, git, bash, Linux, PRAAT, ELAN
- Sentimental Analysis, Chatbots, Word sense disambiguation (WSD), Topic modelling, Named Entity Recognition (NER), Web-crawling, Text classification and clusterisation, Duplicates detection, Sequence labelling, Information Retrieval, Automatic text summarisation

Projects

- **Resume Chatbot** [\[Link\]](#) The chatbot answers simple interview-like questions about me. Built with Python, Flask and DialogFlow.
- **Telegram Chatbot for data annotation** [\[Link\]](#) 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. Built with Python, Telegram and SQLite3.
- **Web app for practising Russian** Web app that automatically generates language tests for practising Russian. Built with Python and Flask.

Academic experience

Conferences

- **PyData Amsterdam** [\[Video\]](#) **Amsterdam**
May 2019
"How to expand your NLP solution to new languages using transfer learning"
- **CLIN29 Conference** **Groningen**
January 2019
"Best practices for learning domain specific cross-lingual embeddings"
- **Conference ConCort on Digital Humanities** **Nizhnyj Novgorod**
October 2016
"Instruments of Quantitative Typological Research"

Publications

- **Best practices for learning domain-specific cross-lingual embeddings, 2019** [\[Link\]](#)
Shakurova L., Nyari B., Li C., Rotaru M.
Proceedings of Repl4NLP workshop, ACL conference
- **Universal Dependencies for Russian: A New Syntactic Dependencies Tagset, 2016** [\[Link\]](#)
Lyashevskaya O., Drogonova K., Zeman D., Alexeeva M., Gavrilova T., Mustafina N. and Shakurova E.

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

Extra activities

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

Interests

Programming · Travelling · Music · Psychology · Literature