Maria Alexeeva
University of Arizona

maxaalexeeva@gmail.com
maxaalexeeva.github.io/

Education

The University of Arizona   Tucson, AZ
- Ph.D. in Linguistics (anticipated 2024)
  Dissertation: *Automatic Extraction and Application of Population Subjective Views*
- M.S. in Human Language Technology (2019)

The University of Utah   Salt Lake City, UT
- M.A. in Linguistics (2015)
- Graduate Teaching English to Speakers of Other Languages (TESOL) Certificate

Far Eastern University of Humanities   Khabarovsk, Russia
- Specialist in Linguistics and Translation, English and Mandarin Chinese (2010)

Relevant Experience

Graduate Researcher (NLP) | CLU Lab, University of Arizona   (August 2018–present)
Supervised by Dr. Mihai Surdeanu, Dr. Rebecca Sharp, and Dr. Clayton Morrison
Present
- Serving as a key researcher and developer for the *HEURISTICS* project under the DARPA *Habitus* program, working on neurosymbolic information extraction systems, with the tasks including data collection and pre-processing, rule-based extraction, model training, and large-language model (LLM) fine-tuning and output evaluation
- Developing datasets: defining tasks, preparing annotation guidelines and training materials, collecting annotations using crowdsourcing and in-house annotation, and conducting quality control
- Providing onboarding and mentorship support to junior researchers

Past
- Led the development of an automatic machine reading pipeline for the *AutoMATES* project (part of the DARPA *Automating Scientific Knowledge Extraction (ASKE)* program)
- Translated information extraction rules from Stanford Dependencies to Universal Dependencies for a biomedical information extraction system (REACH)

NLP Engineer Consultant | Lum AI   Tucson, AZ (2020–July 2023)
- Worked on a variety of NLP-related tasks, which included proposing solutions to achieve project goals; using, modifying, and evaluating in-house tools; data annotation; and error analysis
- Latest project: rule-based information extraction from spoken interview transcripts in the agriculture domain

NLP Intern | Crane AI   New York City, NY (Summer 2018)
- Implemented intent classifiers and trained entity extractors using RASA NLP
- Researched algorithms and tools and evaluated their success potential given the resources available
- Gathered, preprocessed, and analyzed data using various NLP toolkits

Technical Experience

Programming:	Python, Scala
Toolkits/Libraries:	Hugging Face Transformers, Scikit-Learn, NLTK, spaCy, pandas, NetworkX, neo4j
General:	Linux, MacOS, Bash, Jupyter Notebook, git, Docker, Amazon Mechanical Turk

Publications

Dumitru, Razvan-Gabriel*, **Maria Alexeeva***[1], Keith Alcock, Nargiza Ludgate, Cheonkam Jeong, Zara Fatima Abdurahaman, Prateek Puri, Brian Kirchhoff, Santadarshan Sadhu, and Mihai Surdeanu. "Retrieval Augmented Generation of Subjective Explanations for Socioeconomic Scenarios." *Sixth Workshop on NLP and Computational Social Science (at NAACL). 2024.* (accepted, peer-reviewed)

Alexeeva, Maria, Caroline Campbell Hyland, Keith Alcock, Allegra A. Beal Cohen, Hubert Kanyamahanga, Isaac Kobby Anni, and Mihai Surdeanu. "Annotating and Training for Population Subjective Views." *13th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis. 2023.* `https://aclanthology.org/2023.wassa-1.36/` (peer-reviewed)

---

[1]equal contribution

Alexeeva, Maria, Allegra A. Beal Cohen, and Mihai Surdeanu. "Combining Extraction and Generation for Constructing Belief-Consequence Causal Links." *Proceedings of the Third Workshop on Insights from Negative Results in NLP. 2022.* `https://aclanthology.org/2022.insights-1.22/` (peer-reviewed)

Alexeeva, Maria, Rebecca Sharp, Marco A. Valenzuela-Escárcega, Jennifer Kadowaki, Adarsh Pyarelal, and Clayton Morrison. "Mathalign: Linking formula identifiers to their contextual natural language descriptions." *In Proceedings of The 12th Language Resources and Evaluation Conference*, pp. 2204-2212. 2020. `https://aclanthology.org/2020.lrec-1.269/` (peer-reviewed)

Zupon, Andrew, **Maria Alexeeva**, Marco A. Valenzuela-Escárcega, Ajay Nagesh, and Mihai Surdeanu. (02019). "Lightly Supervised Representation Learning with Global Interpretability." *Proceedings of the 3rd Workshop on Structured Prediction for NLP.* 18–28. `http://dx.doi.org/10.18653/v1/W19-1504` (peer-reviewed)

Christison, MaryAnn, and **Maria Alexeeva**. "Formal Characteristics of Learner Language." The TESOL Encyclopedia of English Language Teaching (2018): 1-7.