Maria Alexeeva

University of Arizona
maxaalexeeva@gmail.com

Summary: I am a PhD candidate in Linguistics focusing on Natural Language Processing. I serve as a Research Associate in the CLU Lab (University of Arizona) working primarily on rule-based information extraction and providing technical support and analysis for various DARPA-funded projects. I also have industry experience having done an internship at Crane AI (New York City, NY), and currently working as an NLP engineer consultant for Lum AI (Tucson, AZ).

Education

The University of Arizona Tucson, AZ

- Ph.D. in Linguistics (anticipated 2023)
- M.S. in Human Language Technology (2019)

The University of Utah Salt Lake City, UT

- M.A. in Linguistics (2015)
- Graduate TESOL Certificate

Far Eastern University of Humanities Khabarovsk, Russia

- Specialist in Linguistics and Translation, English and Mandarin Chinese (2010)

Relevant Experience

Research Associate | Dept. of Computer Science, University of Arizona (Dec 2021–present)
Supervised by Dr. M. Surdeanu

- Working on extraction/generation of belief-consequence influence links (the *HEURISTICS* project, the DARPA *Habitus* program), including data collection, model training, and rule-based extraction
- Contributing to the development of a rule-based event extraction pipeline in the agricultural domain
- Providing support to junior researchers (an intern, an undergraduate student)

NLP Engineer Consultant | Lum AI Tucson, AZ (2020–present)

- Working on a variety of NLP-related tasks, which include using, modifying, and evaluating in-house tools, data annotation, and error analysis

Research Assistant | School of Information, University of Arizona (2019–2021)
Supervised by Dr. C. Morrison, Dr. R. Sharp

- Led the development of the automatic text reading pipeline (the *AutoMATES* project, the DARPA *Automating Scientific Knowledge Extraction (ASKE)* program)
- Contributed to planning and conducting annotation exercises
- Provided mentoring to a junior RA

Research Assistant | Dept. of Computer Science, University of Arizona (Fall 2018)
Supervised by Dr. M. Surdeanu

- Translated the information extraction rules for the REACH system from Stanford Dependencies to Universal Dependencies
- Modified existing rules and code based on users' feedback

NLP Intern | Crane AI New York City, NY (Summer 2018)

- Implemented intent classifiers and trained entity extractors using RASA NLP
- Researched algorithms and tools to solve problems at hand and evaluated their success potential given the resources available
- Searched for, preprocessed, and analyzed data using various NLP toolkits

Technical Experience

Programming:	Scala, python
Toolkits/Libraries:	Scikit-Learn, NLTK, NumPy, pandas, NetworkX, neo4j, transformers
General:	Linux, MacOS, Bash/Shell, Jupyter Notebook, git, regular expressions

Publications

Alexeeva, Maria, Allegra A. Beal, 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.

Other Projects

Impact of Clause Boundary Information on the Performance of a Syntactic Dependency Parser
Using clause boundary information, we significantly improved performance of the head identification component of a syntactic dependency parser compared to the baseline in several conditions.

Transformer Models for RST-style Discourse Parsing
We compared the performance of our shift-reduce discourse parser with several transformer models used for encoding elementary discourse units and experiment with small variations in the architecture. Based on our experiments, transformer models by themselves are not able to ensure that a discourse parser will reach state of the art performance even though they may have done so for other natural language processing tasks.

Analyzing the data from Lifestyle Intervention for Ovarian Cancer Enhanced Survival (LIVES) Study
The work was part of a broader effort to explore the possibility of using machine learning techniques in the process of evaluating motivational counseling calls in terms of their ability to improve lifestyle behavior outcomes of cancer survivors. Given the data for two arms of the study, we analyzed the data for whether or not the two arms are sufficiently distinct to be automatically classified; we also explored the possibility of automatically aligning participants' responses to compare reactions of interviewers to similar utterances between the two arms of the experiment.