

# ELIZABETH SPAULDING, PH.D. CANDIDATE

email: elizabeth.spaulding@colorado.edu web: e-spaulding.github.io

## EDUCATION

---

### University of Colorado Boulder

Ph.D. in Computer Science

(Expected April 2025)

### University of Nebraska-Lincoln

B.S. in Mathematics (Minors: Computer Science, Spanish GPA 4.0/4.0)

May 2019

## RESEARCH & INDUSTRY EXPERIENCE

---

### Ph.D. Research in NLP & Machine Learning

University of Colorado Boulder

Aug 2019 – April 2025 (expected)

Boulder, CO

- Developed and fine-tuned advanced neural network models (GPT-4, Llama, BERT) for semantic representation tasks, employing deep learning frameworks such as PyTorch, HuggingFace Transformers, and Langchain.
- Explored statistical techniques to measure how large language models capture meaning in unstructured text data.
- **Key Achievements:** Developed a novel textual data processing pipeline, encompassing several text classification and extraction tasks, enabling visualization and interpretation of LLM meta-linguistic reasoning.

### Research Assistant (DARPA AIDA and KAIROS Projects)

University of Colorado Boulder

June 2020 – May 2024

Boulder, CO

*AIDA and KAIROS were ambitious multi-university DARPA projects centered around extracting and clustering events from news articles and creating knowledge graphs to synthesize and visualize competing narratives in the news.*

- Designed, built, and maintained an ontology for explainable, scalable event extraction from newstext
- Led a diverse, geographically-distributed team of 7 researchers, including faculty, industry scientists, and graduate students, on two peer-reviewed publications on the design of ontological frameworks for explainable NLP.
- Developed event matching code in Java for team's entity and event clustering pipeline.
- **Key Achievements:** Implemented the first large-scale analysis of 9,000+ missing data types in Wikidata (publications [1] and [3]) and created a proposal directed to diverse stakeholders to integrate the missing types en masse.

### Research Intern, CTO Office

Bloomberg, L.P.

June 2022 – Dec 2022

New York, NY

- Designed, built, and evaluated predictive models for semantic proto-role labeling, leveraging advanced deep learning (Pytorch Lightning) and high-performance computing.
- Visualized and communicated machine learning analytics to business stakeholders and research collaborators via TensorBoard.
- Delivered an end-to-end pipeline that streamlined unstructured data ingestion, analysis, and evaluation.
- **Key Achievements:** Delivered state-of-the-art results on semantic proto-role labeling (ACL publication [2]), a large-scale text extraction and analysis task.

### Software Engineer Intern

Atilika

June 2021 – Aug 2021

Tokyo, Japan (Remote)

- Engineered a prototype Chrome extension to integrate English-to-Japanese vocabulary substitution for language learners, using multilingual embeddings for sense disambiguation.
- Implemented data preprocessing and real-time text analysis in JavaScript and Python, ensuring minimal latency and robust dictionary handling.

## TECHNICAL SKILLS

---

**Languages & Tools:** Python, C++, Java, JavaScript, R, PyTorch, HuggingFace, Langchain, Tensorflow, TensorBoard, scikit-learn, Pandas, spaCy, NLTK, Git

**Methods:** Deep learning, neural networks, large language models, clustering, regression, classification, data wrangling, predictive models, knowledge graphs

## CLASSES TAUGHT

---

**CSCI 5832, Natural Language Processing**, CU Boulder, TA, *S24, F24, Su24, F21*  
**CSCI 2824, Discrete Structures**, CU Boulder, Instructor, *Su23*  
**CSCI 1300, Starting Computing (C++)**, CU Boulder, Lead TA, *F19, F20, S21*  
**CSCI 2270, Data Structures (C++)**, CU Boulder, TA, *S20*

## PUBLICATIONS

---

- [1] **Elizabeth Spaulding**, Kathryn Conger, Anatole Gershman, Mahir Morshed, Susan Windisch Brown, James Pustejovsky, Rosario Uceda-Sosa, Sijia Ge and Martha Palmer. 2024. PropBank goes Public: Incorporation into Wikidata. In *Proceedings of the 18th Linguistic Annotation Workshop (LAW-XVIII)*, St. Julians, Malta. European Association for Computational Linguistics.
- [2] **Elizabeth Spaulding**, Gary Kazantsev, and Mark Dredze. 2023. Joint End-to-end Semantic Proto-role Labeling. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 723736, Toronto, Canada. Association for Computational Linguistics.
- [3] **Elizabeth Spaulding**, Kathryn Conger, Anatole Gershman, Rosario Uceda-Sosa, Susan Windisch Brown, James Pustejovsky, Peter Anick, and Martha Palmer. 2023, June 20. The DARPA Wikidata Overlay: Wikidata as an ontology for natural language processing. In *Proceedings of the 15th International Conference on Computational Semantics (ICWS)*. The 2023 Joint ACL - ISO Workshop on Interoperable Semantic Annotation, Nancy, France.
- [4] Sarah Beemer, Zak Boston, April Bukoski, Daniel Chen, Princess Dickens, Andrew Gerlach, Torin Hopkins, Parth Anand Jawale, Chris Koski, Akanksha Malhotra, Piyush Mishra, Saliha Muradoglu, Lan Sang, Tyler Short, Sagarika Shreevastava, **Elizabeth Spaulding**, Testumichi Umada, Beilei Xiang, Changbing Yang, and Mans Hulden. 2020. Linguist vs. Machine: Rapid Development of Finite-State Morphological Grammars. In *Proceedings of the 17th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 162170, Online. Association for Computational Linguistics.

## AWARDS & LEADERSHIP

---

**Summer Graduate School Fellowship** (2024)  
**Nelson A. Prager Family and J. H. Martin Endowed Grad Fellowship** (2023)  
**CU Boulder CS Pedagogy Committee Grad Student Representative** (2023 – 2025)  
**Lead Graduate Student Fellowship** (2022)  
**David T. Spalding Graduate Teaching Fund Fellowship** (2022)  
**CU Boulder Center for Teaching & Learning (CTL) Lead TA** (2022 – 2024)  
**Dean's Summer Research Fellowship** (2020)  
**University of Nebraska Regents Scholarship** (2015)