***Curriculum Vitae* of Anna Martin-Boyle**
Department of Computer Science and Engineering
University of Minnesota
(408) 677-9951
mart5877@umn.edu

EDUCATION
University of Minnesota, Minneapolis, MN
Department of Computer Science and Engineering, first year Ph.D. student
- 2022-present
- Advised by Dr. Dongyeop Kang, member of the Minnesota NLP Lab
- Honors: UMN College of Science and Engineering Fellowship
- Relevant coursework: Deep Learning; Introduction to Data Mining; Software Engineering; Human Centered AI (Fall 2023).

University of Minnesota, Duluth, MN
M.S. in Computer Science, 2020-2022.
- Thesis: *Extracting Task Descriptions from Shared Task Overview Papers*
- Advised by Dr. Ted Pedersen
- Relevant coursework: Intro to Natural Language Processing; Advanced Natural Language Processing; Intro to Artificial Intelligence; Intro to Machine Learning and Data Mining; Sensors and IoT Devices; Modeling with Dynamic Systems.

Northern Michigan University, Marquette, MI
B.S. in Computer Science, summa cum laude, 2017-2020

PRESENTATIONS AND PUBLICATIONS

Ryan Koo, Anna Martin-Boyle, Linghe Wang, and Dongyeop Kang. Decoding the End-to-end Writing Process in Scholarly Manuscript via Writer-action Taxonomy. Accepted at the CHI 2023 In2WritingWorkshop.

Anna Martin-Boyle, Andrew Head, Kyle Lo, Risham Sidhu, Marti Hearst, and Dongyeop Kang. Complex Mathematical Symbol Definition Structures: A Dataset and Model for Coordination Resolution in Definition Extraction. arXiv:2305.14660.

Anna Martin, Ted Pedersen, and Jennifer D'Souza. NLPSharedTasks: A Corpus of Shared Task Overview Papers in Natural Language Processing Domains. In *Proceedings of the first Workshop on Information Extraction from Scientific Publications,* pages 105-120, Online. Association for Computational Linguistics.

Anna Martin, 2022. *Annotating and Automatically Extracting Task Descriptions from Shared Task Overview Papers in Natural Language Processing Domains* (Access No. 2022. 29213770) [Master's Thesis], University of Minnesota, Duluth]. ProQuest Dissertations Publishing.

Anna Martin, Jennifer D'Souza, and Ted Pedersen. Annotation Natural Language Processing Shared Task Descriptions. Poster presented at: *SciNLP 2021: 2nd Workshop on Natural Language Processing for Scientific Text* at AKBC 2021. October 2021. Online.

Anna Martin and Ted Pedersen. 2021. Duluth at SemEval-2021 Task 11: Applying DeBERTa to Contributing Sentence Selection and Dependency Parsing for Entity Extraction. In *Proceedings of the Fifteenth*

Workshop on Semantic Evaluation, Bangkok (online). ACL.

AWARDS

UMN College of Science and Engineering Fellowship (2022-Present)

Seebach Fellowship (2022-2023)

UMD Computer Science Graduate Research Fellowship (2021)

NMU Department of Mathematics and Computer Science Outstanding Graduate (2020)

TECHNICAL SKILLS
Languages: Python, Java, C++. **Machine learning:** PyTorch, HuggingFace, BERT fine-tuning, GPT prompt design. **Data processing:** Pyspark, Pandas, spaCy, Stanford Core NLP, Microsoft Excel. **Communication tools:** LaTeX, Microsoft Powerpoint, G Suite

---

RESEARCH EXPERIENCE
Related Works Composition with LLMs June 2023 - Present
Literature discovery has become a particularly challenging aspect of research due to the rapid growth rate of scientific output. We are examining how LLMs might be used to assist with this process. In particular, we are developing a taxonomy of Related Works composition errors in order to better evaluate GPT4's output when tasked with composing a related works section given a set of references. We are also looking at how LLMs can be used to group references in order to support the author's framing of their work.

Does the Structural Simplification of Sentences Improve Downstream IE Tasks? January 2023 - Present
We are investigating whether prompting ChatGPT to structurally simplify sentences from biomedical texts improves performance of downstream information extraction systems, including current state-of-the-art BERT-like models and ChatGPT in a few-shot setting.

ScholaWrite: Recording Writer Actions for Rhetorical Adjustments July 2022 - Present
We are running a study collecting keystroke data from scholarly writers to better understand the scientific writing process. This data will be annotated using our hierarchical taxonomy of scholarly writing processes.

Coordination Resolution in Mathematic Definition Recognition June 2022 - January 2023
Under the supervision of Dr. Dongyeop Kang, assisted with a dataset and model for extracting definitions of mathematical symbols organized in complex coordination structures.

Extracting Task Descriptions from Shared Task Overview Papers (master's thesis) May 2021-May 2022
With Dr. Jennifer D'Souza, developed a gold-standard corpus of 254 shared task overview papers published by the ACL with annotated task description phrases. Used the corpus to train a machine reader to automatically extract natural language task descriptions that provide a concise description of the task to be performed that makes sense out of the context of the paper.

Bold Ideas: Design of a Dementia Friendly Living Space (study assistant) December 2020-May 2021
  Assisted the head researcher Dr. Arshia Khan by supporting the human participants during study
  sessions and analyzing the electrodermal activity, blood pressure, heart rate, and participant survey
  response data.

RELEVANT WORK EXPERIENCE

Teaching Assistant at University of Minnesota, Duluth August 2020-May 2022

Computer Science II	Fall 2021 - Spring 2022	Instructed by Professor Steven Holtz
Natural Language Processing	Fall 2021	instructed by Dr. Ted Pedersen
Computer Architecture	Spring 2021	instructed by Dr. Ted Pedersen
Software Analysis and Design	Fall 2020 - Spring 2021	instructed by Dr. Timothy Colburn
Computer Ethics	Fall 2020	instructed by Dr. Ted Pedersen

Software Engineering Intern at Datastax, Santa Clara, CA May-October 2019
  Built a system health and performance reporting system in Java that collected throughput and
  latency metrics from multiple components of the code base. Wrote a program to send statistics to a
  dashboard and organized data so that the system's performance could be observed at a glance.

SERVICE

Minnesota NLP Server Admin Fall 2022 - Present
  Requires adding users, updating system and software, and troubleshooting errors as they arise.

NLP Seminar Organization Committee Fall 2022
  Assist in organizing Minnesota Natural Language Processing Group's NLP Seminar by finding and
  reaching out to potential speakers, organizing PR for the events, and hosting speakers virtually.