

Clara Na

csna@cs.cmu.edu | [Google Scholar](#) | [LinkedIn: clara-na](#) | [GitHub: clarana](#) | clarasna.com

EDUCATION

- Carnegie Mellon University** Aug. 2021 – Present
Pittsburgh, PA *GPA: 4.02*
PhD in Language and Information Technology, School of Computer Science
- University of Virginia** Aug. 2017 – May 2021
Charlottesville, VA *GPA: 3.98*
BA Computer Science, BA Mathematics

RESEARCH EXPERIENCE

- AllenNLP Research Intern → External Collaborator** Summer 2023 – Present
Allen Institute for AI, w/ Jesse Dodge & Pradeep Dasigi *Seattle, WA*
 - Exploring the limits of modularity in parallel pre-training of language models (ongoing)
- Research Assistant in Natural Language Processing** Aug. 2021 – Present
Language Technologies Institute @ CMU, w/ Prof. Emma Strubell *Pittsburgh, PA*
 - Studying the current state and historical development of NLP research and LLMs in the historical context of incentives, institutions, and infrastructure (w/ Amanda Bertsch and Sireesh Gururaja, EMNLP '23 and ongoing)
 - Characterizing task- and data-specific factors that influence transformer language model fine-tuning energy intensity (w/ undergrad mentee Xiaorong Wang, EMNLP Findings '23)
 - Exploring the effect of local sharpness-aware optimization on transformer model compression and evaluation thereof (EMNLP Findings '22)
- Research Assistant in Computational Social Science** Sep. 2020 – Feb. 2021
University of Virginia Computer Science Department, w/ Prof. Yangfeng Ji & Prof. Lanfei Shi *Charlottesville, VA*
 - Constructed an academic authorship network and performing network analysis in order to understand how researchers collaborate and contribute to publications
 - Incorporating NLP and machine learning techniques into analyses to provide a framework for enriched network analysis
- Research Assistant in Medical AI** Jan. 2020 – May 2020
University of Virginia, w/ Josephine Lamp & Prof. Lu Feng *Charlottesville, VA*
 - Predicted outcomes from ICU patient medical records using signal temporal logic (STL) learning
- Data Science/Research Intern** June 2018 – June 2019
Convergent Behavioral Science Initiative, w/ Katelyn Stenger & Prof. Leidy Klotz *Charlottesville, VA*
 - Built lexicons of additive and subtractive language to examine design intent from millions of USPTO patents
 - Analyzed patents and design trends over time, across patent classifications, and through references

PUBLICATIONS

Pre-training and Fine-tuning BERT: Energy and Carbon Considerations

Xiaorong Wang*, **Clara Na***, Emma Strubell, Sorelle Friedler, Sasha Luccioni

To appear in Findings of the 2023 Conference on Empirical Methods in Natural Language Processing

To Build Our Future, We Must Know Our Past: Contextualizing Paradigm Shifts in Natural Language Processing

Sireesh Gururaja*, Amanda Bertsch*, **Clara Na***, David Gray Widder, Emma Strubell

To appear in The 2023 Conference on Empirical Methods in Natural Language Processing

The Framework Tax: Disparities Between Inference Efficiency in NLP Research and Deployment

Jared Fernandez, Jacob Kahn, **Clara Na**, Yonatan Bisk, Emma Strubell

To appear in The 2023 Conference on Empirical Methods in Natural Language Processing

Train Flat, Then Compress: Sharpness-Aware Minimization Learns More Compressible Models

Clara Na, Sanket Vaibhav Mehta, Emma Strubell

Findings of the 2022 Conference on Empirical Methods in Natural Language Processing

Less is more? In patents, design transformations that add occur more often than those that subtract.

Katelyn Stenger, **Clara Na**, Leidy Klotz

Ninth International Conference on Design Computing and Cognition (DCC20)

MANUSCRIPTS

Virtual Task Selection in Meta-Learning for Domain Generalization in Semantic Parsing

Cathy Jiao*, Xuecong (Cindy) Fu*, **Clara Na***, Emma Strubell

Less is rare: Analysing categories of design transformations described in patents

Katelyn Stenger, **Clara Na**, Leidy Klotz, Patrick Hancock

TEACHING EXPERIENCE

Teaching Assistant for 11-767 (On-Device Machine Learning)

Fall 2023

Carnegie Mellon University

Pittsburgh, PA

- Office hours, rubrics, grading, writing and testing assignments

Teaching Assistant for CS 1110 (Introduction to Programming)

Fall 2018 – Spring 2021

University of Virginia

Charlottesville, VA

- **Head Teaching Assistant**, Fall 2019 – May 2021
- Held office hours, graded assignments and exams, wrote assignments and exam questions
- Worked with professors and three other head TAs, managing 50+ TAs and 600+ students

Teaching Assistant for CS 4774 (Machine Learning)

Fall 2019

University of Virginia

Charlottesville, VA

- Held office hours, contributed to curriculum development

OTHER WORK EXPERIENCE

Software Engineering Intern

Summer 2020 & Summer 2021

Microsoft (R&D Data Cosmos Team)

Redmond, WA (Remote)

- **2021:** Designed and implemented a Data Mining-based representative benchmark suite for Cosmos perf validation
- **2020:** Implemented smart job selection by relaxed hash value of optimized job graph for Scope Flighting – project voted “**Most Creative**” at **2020 Microsoft Azure Data Intern Showcase**, out of 59 intern projects

Data Science Intern

Summer 2019

The Washington Post

Washington, DC

- Investigated website user activity and subscription patterns to propose new signals for propensity modeling
- Adapted Amazon Sagemaker’s hyperparameter optimization feature to tune a preprocessing pipeline
- Proposed and prototyped a web application for enhancing the digital reading experience for users by automatically generating internal contextual links and summaries for events and named entities

AWARDS & HONORS

Best Paper | *LTI Student Research Symposium*

August 2023

Best Novel Work | *LTI Student Research Symposium*

August 2022

XSEDE Startup Allocation | *Computational resources valued at \$1,630*

February 2022

NSF Graduate Research Fellowship | *National Science Foundation*

2021 – Present

Echols Scholar | *University of Virginia College of Arts and Sciences*

Fall 2017 – Spring 2021

Dean’s List | *University of Virginia*

Spring 2018 – Fall 2019

National Merit Scholar | *National Merit Scholarship Corporation*

2017

President’s Volunteer Service Award | *Corporation for National and Community Service*

2016

SERVICE

LTI Mentoring Program Committee <i>LTI @ CMU SCS</i>	Summer 2022 - Present
Reviewer <i>NeurIPS</i>	2023
Reviewer <i>ICML</i>	2023
Organized LTI-internal {faculty, student} panels on LLMs <i>LTI @ CMU SCS</i>	Spring 2023
LTI Student Executive Committee <i>LTI @ CMU SCS</i>	Fall 2022 - Spring 2023
LTI PhD Representative <i>Graduate Student Assembly @ CMU</i>	Fall 2022 - Spring 2023
Undergraduate AI Mentoring Program Mentor <i>CMU SCS</i>	Fall 2021, 2022
Graduate Application Support Program Mentor <i>LTI @ CMU SCS</i>	Fall 2021, 2022
Graduate Application Support Coordinator <i>LTI @ CMU SCS</i>	Fall 2021

SKILLS & MISC.

Languages: Python, R, C, C++, C#, Java, Intermediate Korean

Libraries: *Python:* pytorch, pandas, numpy, sklearn, pyspark, matplotlib, flask. *R:* dplyr, caret, quanteda

Technologies: Git/GitHub, AWS, Azure (Microsoft Azure Fundamentals certified)

Relevant Coursework: Natural language processing, Vision & language, Machine learning, On device machine learning, Computational ethics, Information retrieval, Algorithms, Operating systems, Artificial intelligence, Cloud computing, Theory of computation, Computer architecture, Data structures, Software development methods, Cybersecurity, Discrete mathematics, Linear algebra, Differential equations, Probability, Abstract algebra, Real analysis, Number theory

Other Activities/Interests/Service: chamber ensemble, piano, musical instruments, mentoring, poetry, running, climbing, organizing, language learning, tutoring