# Madeleine N. van Zuylen

madeleine.vanzuylen@gmail.com | 1-206-915-8875 | Seattle, WA
LinkedIn, Google Scholar, Personal Website

---

## EDUCATION

**Northeastern University |** Boston, MA — Expected: May 2023
MS in Computer Science — GPA: 4.0/4.0

**University of Notre Dame** | Notre Dame, IN — May 2017
BS in Biochemistry and Applied Computational Mathematics and Statistics

## SKILLS

**Technical:** Java, Python, Git, Terraform, Atlantis, C, SQL, R, S3, Samza, Grafana, Mockito, Kafka

## COURSEWORK

**Programming:** Intensive Foundations of CS (Python), Object-Oriented Design (Java), Data Str, Algo & App in Cmp Sys (C), Algorithms (Java), Database Management Sys (R, SQL), Machine Learning (Python), Deep Learning, Web Dev

**Math:** Discrete Structures, Calc III, Linear Algebra, Differential Equations, Intro to Probability, Numerical Analysis, Stat Methods & Data Analysis, Mathematical/Comp Modeling

## WORK EXPERIENCE

**Software Development Engineering Intern | Amazon | AWS | File Storage Gateway**
September 2022 - December 2022

- Provided insight into unresponsive gateways by designing metrics to track gateway availability
- Automated core dumps when a gateway became unavailable to decide when an automatic restart of a gateway process is needed

**Software Development Engineering Intern | Redfin | Listing Ingestion Platform**
May 2022 - August 2022

- Created Samza app to store house listing records in AWS S3 with versioning to provide insight into historical changes in the raw records for 665,000 listings, 62,000 agents, and 14,000 Brokers
- Added 13,500 listings from West Alabama Multiple Listing Service onto redfin.com representing an .8% increase in Redfin's total coverage

**Data Science Analyst II | Allen Institute for AI | Semantic Scholar**
June 2017- Present | June 2017 - May 2022 Full time | September 2020 - July 2021 20 hours/week

- Curated large novel training datasets for machine learning research projects including verifying scientific claims, citation intent, bias in medical studies, information extraction, PDF accessibility, and document summarization
- Designed and built annotation tasks launched to crowdsource workers
- Recruited, trained, and managed hundreds of crowdsource workers to annotate and build datasets for machine learning mode
- Analyzed and iterated on a field of study classifier to label scientific papers by domain

## PUBLICATIONS

**MS^2: Multi-Document Summarization of Medical Studies**
*EMNLP, 2021*

**Extracting a Knowledge Base of Mechanisms from COVID-19 Papers**
*NAACL, Human Language Technologies*, 2021

**Improving the accessibility of scientific documents: Current state, user needs, and a system solution to enhance scientific PDF accessibility for blind and low vision users**
*ArXiv, 2021*

**MedICaT: A Dataset of Medical Images, Captions, and Textual References**
*EMNLP,* 2020

**SCIREX: A Challenge Dataset for Document-Level Information Extraction**
*ACL,* 2020

**Fact or Fiction: Verifying Scientific Claims**
*EMNLP,* 2020

**Quantifying Sex Bias in Clinical Studies at Scale With Automated Data Extraction**
*JAMA Network Open,* 2019

**Structural Scaffolds for Citation Intent Classification in Scientific Publications**
*NAACL, Human Language Technologies*, 2019

**Construction of the Literature Graph in Semantic Scholar**
*NAACL, Human Language Technologies*, 2018

**A Dataset of Peer Reviews (PeerReaD): Collection, Insights and NLP Applications**
*NAACL, Human Language Technologies,* 2018

**Apoptosis-related Genes Control Autophagy and Influence DENV-2 Infection in the Mosquito Vector, Aedes Aegypti**
*Insect Biochemistry & Molecular Biology*, September 2016

## POSTERS AND DEMONSTRATIONS

**SciA11y: Converting Scientific Papers to Accessible HTML**
*ASSETS 2021 Posters and Demonstrations, Artifact Award 1st Place*

---

## OTHER WORK EXPERIENCE AND ACTIVITIES

**Research Intern | Harvard University |** Paulson School of Engineering and Applied Sciences | Weitz Lab
June 2016-August 2016 | Full time

- Researched and built a novel production method for micro-hydrogels carrying DNA primer for droplet barcoding and large scale sequencing projects using glass capillary devices and Polydimethylsiloxane (PDMS) devices

**Research Intern | Harvard University |** Paulson School of Engineering and Applied Sciences | Brenner Lab
June 2015-August 2015 | Full time

- Studied and developed DNA and molecular modeling programs to investigate DNA self-assembly mechanisms and to simulate thermodynamic and mechanical properties of polymers using Monte Carlo and Molecular Dynamics

**Research Intern | University of Notre Dame |** Biology Department | Severson Lab
September 2014-May 2015 | Part time

- Studied Dengue Fever transmission in mosquitoes through analysis of autophagy in the mid-gut, ovaries, and fat bodies to determine possible methods of Dengue Fever prevention. Presented at the College of Science Joint Annual Meeting

**Research Intern | University of Washington** | Genome Sciences Department | Fields Lab
June 2014-August 2014 | Full time

- Researched synonymous codon changes in yeast to determine effect of synonymous codons on GFP production and expression

## BIOCHEMISTRY SKILLS

**Lab-Based:** Site Directed Mutagenesis, *E. coli* and *S. cerevisiae* plasmid transformation, DNA sequencing, comparative growth analysis, Atomic Force Microscopy, Scanning Electron Microscope, mosquito dissection, microfluidic device construction

## VOLUNTEER ACTIVITIES

**King County Family Law Court Appointed Special Advocate (FL CASA)**
January 2020 - December 2020 | 6 hrs/week

- Gather information in high conflict child custody cases and report back to the court recommendations for a parenting plan and useful services to the parties
- Interview parents, collaterals, and child, tour the home, write two reports, and testify in court