

Amandalynne Paullada | Curriculum Vitae

✉ paullada@uw.edu • 🌐 www.amandalynne.pw

RESEARCH INTERESTS

Natural language processing, Language justice, Critical data studies, Digital health equity

EDUCATION

Ph.D. in Computational Linguistics, University of Washington <i>Advisor:</i> Fei Xia	2016 – 2021
M.A. in Computational Linguistics, Brandeis University	2013 – 2015
B.A. in Linguistics, University of California, Santa Cruz <i>Study abroad, full immersion:</i> Universitat de Barcelona – 2011	2008 – 2012
B.A. in Economics, University of California, Santa Cruz	2008 – 2012

PROFESSIONAL EXPERIENCE

Postdoctoral Scholar-Fellow, Biomedical Informatics and Medical Education University of Washington School of Medicine <i>Funding:</i> National Library of Medicine Biomedical Informatics Research Trainee Program	2022 – present
Research Assistant, Biomedical Informatics and Medical Education University of Washington School of Medicine <i>PI:</i> Trevor Cohen	2019 – 2021
Research Intern, Semantic Scholar Allen Institute for Artificial Intelligence <i>Mentors:</i> Sergey Feldman, Kyle Lo	Fall 2018
Data Science for Social Good Student Fellow University of Washington eScience Institute <i>Team:</i> Seattle Mobility Index	Summer 2018
Data Science & NLP Intern Cambia Health Solutions, Seattle WA	Summer 2017
Researcher Smart Information Flow Technologies, Lexington MA	2015 – 2016
Research Intern Smart Information Flow Technologies, Lexington MA	2014 – 2015
Lexical Data Processing Intern PANLEX Project, San Francisco CA	Summer 2013

PUBLICATIONS

Angelina McMillan-Major, **Amandalynne Paullada**, and Yacine Jernite. An interactive exploratory tool for the task of hate speech detection. *HCI+ NLP 2022*, page 11, 2022.

Courtney Mansfield, **Amandalynne Paullada**, and Kristen Howell. Behind the mask: Demographic bias in name detection for PII masking. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 76–89, 2022.

Amandalynne Paullada, Inioluwa Deborah Raji, Emily M Bender, Emily Denton, and Alex Hanna. Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Patterns*, 2(11):100336, 2021.

Inioluwa Deborah Raji, Emily M Bender, **Amandalynne Paullada**, Emily Denton, and Alex Hanna. AI and the everything in the whole wide world benchmark. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.

Naomi Shapiro, Amandalynne Paullada, and Shane Steinert-Threlkeld. A multilabel approach to morphosyntactic probing. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 4486–4524, 2021.

Amandalynne Paullada, Inioluwa Deborah Raji, Emily M Bender, Emily Denton, and Alex Hanna. Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Workshop on Machine Learning Retrospectives, Surveys & Meta-Analyses, NeurIPS*, 2020. **Best Paper Award Recipient**.

Amandalynne Paullada, Bethany Percha, and Trevor Cohen. Improving Biomedical Analogical Retrieval with Embedding of Structural Dependencies. In *Proceedings of the 19th SIGBioMed Workshop on Biomedical Language Processing*, pages 38–48. Association for Computational Linguistics, 2020.

Ian Kennedy, Christian Hess, **Amandalynne Paullada**, and Sarah Chasins. Racialized Discourse in Seattle Rental Ad Texts. *Social Forces*, 08 2020.

Rachael Tatman, Leo Stewart, **Amandalynne Paullada**, and Emma Spiro. Non-lexical Features Encode Political Affiliation on Twitter. In *Proceedings of the Second Workshop on NLP and Computational Social Science*, pages 63–67. Association for Computational Linguistics, 2017.

Rachael Tatman and **Amandalynne Paullada**. Social Identity and Punctuation Variation in the # BlueLivesMatter and # BlackLivesMatter Twitter Communities [poster]. In *33rd Northwest Linguistics Conference, Vancouver, British Columbia*, 2017.

David McDonald, Scott Friedman, **Amandalynne Paullada**, Rusty Bobrow, and Mark Burstein. Extending Biology Models with Deep NLP over Scientific Articles. In *AAAI-16 Workshop on Knowledge Extraction from Text*, 2016.

David McDonald, Scott Friedman, Mark Burstein, Alex Plotnick, **Amandalynne Paullada**, Rusty Bobrow, and Brent Cochran. Learning by Reading: Extending & Localizing Against a Model. In *Fourth Annual Conference on Advances in Cognitive Systems*, 2016.

Peggy Wu, Tammy Ott, **Amandalynne Paullada**, Derek Mayer, Jeremy Gottlieb, and Peter Wall. Cognitive Workload Assessment through Zero Intrusion Detection Methods. In *3rd International Conference on Safety Management and Human Factors*, 2016.

Peggy Wu, Tammy Ott, **Amandalynne Paullada**, Derek Mayer, Jeremy Gottlieb, and Peter Wall. ATHENA – a zero-intrusion no contact method for workload detection using linguistics, keyboard dynamics, and computer vision [poster]. In *18th International Conference on Human-Computer Interaction*, 2016. Best Poster Award.

MANUSCRIPTS UNDER REVIEW

Amandalynne Paullada, Language justice and machine translation on social media, *Modern Languages Open*

Serena Xie, **Amandalynne Paullada**, Yinzhou Wang, Diana Loudon, Magaly Ramirez, Trevor Cohen, and Weichao Yuwen. Can Natural Language Processing Support Cultural Adaptation of Internet- and Mobile-based Health Interventions? A Scoping Review. *Journal of Biomedical Informatics*

Oliver Bear Don't Walk IV, **Amandalynne Paullada**, Avery Everheart, Reggie Casanova-Perez, Trevor Cohen, and Tiffany Veinot. Opportunities for Incorporating Intersectionality into Health Informatics. *Journal of Biomedical Informatics*

PEER-REVIEWED ABSTRACTS AND POSTERS

Can Natural Language Processing Support Cultural Adaptation of Digital Health Interventions? A Scoping Review (Talk)

American Medical Informatics Association (AMIA) Symposium

November 2023

Ethical considerations for secondary use of medical data (Poster)

UW Medical Data Science Symposium

February 2023

Multi-label Element Extraction for Evidence Based Medicine

Second Widening NLP Workshop (talk)

June 2018

12th Women in Machine Learning Workshop (poster)

December 2017

Totes Adorbs: Identity Effects in Colloquial English Truncation (Talk)

USC Undergraduate Linguistics Conference

October 2012

INVITED TALKS

The 'Real World' of Language Technology: A Case Study of Machine Translation

UW Biomedical and Health Informatics Seminar Series (BIME 590)

November 2022

Content moderation has a language problem (Panelist)

RightsCon

June 2022

Pervasive Data Ethics for Computational Research (PERVADE) Webinar Series

Host: Katie Shilton (University of Maryland)

July 2022

Who and what gets counted? Contextual requirements for datasets (Panelist)

Building the NIST AI Risk Management Framework: Workshop #2

March 2022

FATE concerns in NLP (Panelist)

MSR & MILA Workshop

April 2021

Digital Humanities Day (Panelist)

University of Washington

February 2021

Data and its (dis)contents

UW Computational Linguistics Lab (Treehouse)

February 2021

Using Datasets Wisely

Cisco UX & Metrics Reading Group

February 2021

Drift AI

January 2021

Rsqr AI

October 2020

TEACHING

Instructor, University of Washington

Seminar: Biomedical Applications of Language Technology (BIME 591 C) Autumn 2023

Instructor, University of Washington

Seminar: Biomedical Applications of Language Technology (BIME 591 B) Spring 2023

Recipient of BIME Best Colloquium Award

Teaching Assistant, University of Washington

Introduction to Language and Society (LING 233) Spring 2018

Statistical Approaches to Natural Language Processing (LING 572) Winter 2018

Introduction to Linguistic Thought (LING 200) Fall 2016 – Fall 2017

Teaching Assistant, Brandeis University

Critical Reading and Analytical Writing for ESL Students (ESL 1a) Fall 2014

Analytical Writing for ESL Students (ESL 1a) Summer 2014

Morphology (LING 115a) Spring 2014

Syntactic Theory (LING 120) Fall 2013

Tutor, Brandeis University

English Language Program 1-on-1 Writing Tutorial 2013 – 2014

GUEST LECTURES

Strategies for Thinking About Research Ethics

Professional Development in Biomedical Informatics (BIME 586) Winter 2022, 2023

Hosts: Trevor Cohen, John Gennari

Data and its (dis)contents

Value Sensitive Data Processing (LING 575) Spring 2022

Host: Margaret Mitchell

Examining Data and Data Practices in Machine Learning

NLP and the Real World (LING 575) Winter 2021

Host: Fei Xia

SERVICE

Workshop Co-chair

ACL Student Research Workshop @ ACL-IJCNLP 2021

Conference and Workshop Organizing

NLM T15 Training Conference 2023 Scientific Committee, Stanford University

NWLC 2020, Virtual

AFLA 2017, University of Washington

WCCFL 2012, University of California, Santa Cruz

Journal Club Organizing

UW Medical AI Journal Club (2023)

Mentorship

‘Should I do a PhD?’ Mentorship Panel, UW Computational Linguistics Lab (2021)

Graduate School in Linguistics Mentorship Panel, UW Linguistics Department (2021)

Outreaching Grads Program, University of Washington (2020–2021)

Surviving Graduate School Group Mentoring Session, ACL 2020

Journal Reviewer

Frontiers in Psychology
Information, Communication and Society
Journal of Biomedical Informatics
Nature Machine Intelligence

Conference Reviewer

The Web Conference 2023
NeurIPS 2022 (Ethics reviewer)
FAccT 2021, 2022
ACL 2019, 2020 (Secondary reviewer)
CoNLL 2019, 2020

Workshop Reviewer

BioNLP 2023
Workshop on Speech and Language Technologies for Dravidian Languages, 2022
Workshop for Natural Language Processing Open Source Software, 2020
Widening Natural Language Processing (WiNLP), 2019, 2020

Student Governance

Senator, UW Graduate and Professional Senate (2019–2020)
Secretary, Linguistic Society of UW (2017–2019)

SCHOLARSHIPS AND AWARDS

Excellence in Linguistic Research Graduate Fellowship

UW Department of Linguistics, Fall 2021

Travel Grant

Widening NLP Workshop, June 2018
Women in Machine Learning Workshop, December 2017

Diversity Scholarship

PyData Seattle, June 2017

PUBLIC SPEAKING & TUTORIALS

AI THAT WRITES 101

Authored by AI Podcast February 2023
Hosts: Stephen Follows, Eliel Camargo-Molina

Lost in Translation: When Machines Learn Language

TEC Talks Podcast July 2022
Host: Kirsten Martin, Notre Dame Technology Ethics Center (ND TEC)

Coding Best Practices

UW Computational Linguistics Lab (Treehouse) March 2021, Sept. 2018, February 2018

Introduction to Praat Scripting

University of Washington Phonetics Lab October 2017

Introduction to IPython

Boston PyLadies Meetup Group November 2015

Introduction to NLTK

Boston Python User Group October 2015

Computing Linguistic Similarity with Python

Boston PyLadies Meetup Group April 2015

WRITING

‘Machine Translation Shifts Power’

The Gradient magazine

July 2021

Runner-up for Gradient Prize

SKILLS

Programming

Languages: Python, Java

Version control: Git

NLP & Machine Learning

PyTorch, NLTK

Languages

English (Native)

Spanish (Fluent)

French (Intermediate)

Russian (Intermediate)

Korean (Heritage speaker)

PROFESSIONAL DEVELOPMENT

UW Biomedical Research Integrity Program (Virtual)

Summer 2022

FAccT Doctoral Consortium (Virtual)

March 2021

Rising Stars in EECS, UC Berkeley (Virtual)

November 2020

Machine Learning Summer School (MLSS), Madrid, Spain

August 2018

Python Conference (PyCon)

Portland, OR

May 2017

Montréal, QC

April 2015

Montréal, QC

April 2014

VOLUNTEERING

Ravenna Pop-up Kitchen, Seattle WA

2018 – present

Green Seattle Day, Seattle WA

2017, 2019

Big Brothers Big Sisters of Massachusetts Bay, Boston MA

2015 – 2016

Harvard Square Homeless Shelter, Cambridge MA

2016