

Michael Miller Yoder

mmyoder@pitt.edu

<https://michaelmilleryoder.github.io>

EDUCATION	Carnegie Mellon University , Pittsburgh, PA	
	<i>Ph.D.</i> , Language and Information Technologies	2021
	Advisor: Carolyn P. Rosé	
	Thesis: “Computational models of identity presentation in language” PDF	
	M.S. , Language Technologies	2017
	Goshen College , Goshen, IN	
	<i>B.A.</i> , Computer Science	2013
	Minors: English, International Studies	
PROFESSIONAL EXPERIENCE	Teaching Assistant Professor	University of Pittsburgh Pittsburgh, PA
	2024–present	
	Teaching-track faculty member in the School of Computing and Information. Focusing on data science instruction to students from a wide variety of backgrounds in both in-person and online formats.	
	Postdoctoral Research Associate	Carnegie Mellon University Pittsburgh, PA
	2022–2024	
	Researcher with the Center for Applied Research on Targeted Violence (formerly the Collaboratory Against Hate), an interdisciplinary initiative of Carnegie Mellon University and the University of Pittsburgh, as well as the Center for Informed Democracy and Social Cybersecurity at Carnegie Mellon University. Developed new methods and analyses for tracking and detecting hate speech in online contexts, incorporating insights from the social sciences. Advised by Prof. Kathleen M. Carley in the Software and Societal Systems Department and Prof. David West Brown in the Department of English.	
	Part-Time Instructor	University of Pittsburgh Pittsburgh, PA
	2023–2024	
	Instructor for graduate-level NLP course in the Department of Computer Science.	
	Graduate Research Assistant	Carnegie Mellon University Pittsburgh, PA
	2015–2021	
	Applied computational models of discourse to study social meaning in language, largely using social media datasets. Advised by Prof. Carolyn P. Rosé in the Language Technologies Institute.	
	Software Engineering Intern	Yahoo Research/Tumblr New York, NY
	2019	
	Detected communities of Tumblr users from social network structure. Explored community-specific language use. Advised by Dr. Kapil Thadani.	
	Research Programmer/Analyst	Carnegie Mellon University Pittsburgh, PA
	2015	
	Designed and implemented neuroimaging database to support machine learning research. Advised by Prof. Timothy Verstynen in the Department of Psychology.	
	Software Developer & IT Administrator	GSAREH Austin, TX area
	2014	
	Developed applications to recognize and map named entities in English and Arabic articles and social media. Managed software development and IT administration of a small GIS startup named Geospatial Applied Research Experts House.	

TEACHING	Human Language Technologies <small>HTML</small> University of Pittsburgh Spring 2025 Instructor for undergraduate-level NLP course. Developed lectures, in-class activities and coding notebooks, quizzes, homework assignments, and an exam. Guided open-ended student-led projects.
	Applied Predictive Modeling University of Pittsburgh Spring 2025 Online Instructor for graduate-level online course on machine learning methods and applied, practical modeling skills, part of the Applied Data-Driven Methods graduate certificate for working professionals. Guided course projects in which students analyzed their own datasets.
	Introduction to Natural Language Processing <small>HTML</small> University of Pittsburgh Fall 2023, Spring 2024, Fall 2024, Fall 2025 Pittsburgh, PA Instructor for graduate-level NLP course. Developed coursework, including lectures, quizzes, and homework assignments. Guided hands-on, student-led projects.
	Introduction to Data-Centric Computing University of Pittsburgh Fall 2024, Fall 2025 Online Instructor for graduate-level online course developed as part of the Applied Data-Driven Methods graduate certificate for working professionals. Guided student technical projects.
	Computational Ethics for Natural Language Processing <small>HTML</small> Carnegie Mellon University Spring 2020 Pittsburgh, PA Teaching assistant for newly developed NLP ethics course with Profs. Yulia Tsvetkov and Alan Black. Gave guest lectures, developed homeworks, and graded coursework.
GRANTS	Applied Machine Learning Carnegie Mellon University Spring 2019 Pittsburgh, PA Teaching assistant with Prof. Carolyn P. Rosé for a hands-on machine learning course developed for a broad range of students, including those without computer science backgrounds. Gave guest lectures, advised students in office hours, and graded coursework.
	Introduction to Artificial Intelligence Goshen College Spring 2013 Goshen, IN Selected as a tutor for introductory artificial intelligence course. Assisted students one-on-one with AI concepts and coursework.
	Linear Algebra and its Machine Learning Applications Marina Barsky, Michael Miller Yoder Nancy Tannery Grant for Open Educational Resources, University of Pittsburgh, 2025 \$5000
RESEARCH INTERESTS	Identifying and Analyzing Influencer Marketing of Pharmaceutical Products on Social Media Ryan Shi, Michael Miller Yoder, Olga Kravchenko Pitt Cyber, University of Pittsburgh, 2024 \$5100
	Natural language processing, computational social science, data science, AI ethics, digital humanities, sociolinguistics, hate speech, language and identity

Measuring Dimensions of Self-Presentation in Twitter Bios and their Links to Misinformation Sharing

Navid Madani, Rabiraj Bandyopadhyay, Michael Miller Yoder, Stephan D. McCabe, Briony Swire-Thompson, Kenneth Joseph
Proceedings of the International AAAI Conference on Web and Social Media (ICWSM). 2025. PDF

A Weakly Supervised Classifier and Dataset of White Supremacist Language

Michael Miller Yoder, Ahmad Diab, David West Brown, Kathleen M. Carley
Proceedings of the Association for Computational Linguistics (ACL). 2023. PDF

How Hate Speech Varies by Target Identity: A Computational Analysis

Michael Miller Yoder, Lynnette Hui Xian Ng, David West Brown, Kathleen M. Carley
Proceedings of the Conference on Computational Natural Language Learning (CoNLL), 2022. PDF.

Modeling Unstructured Data: Teachers as Learners and Designers of Technology-enhanced Artificial Intelligence Curriculum

Cansu Tatar, Michael Miller Yoder, Madeleine Coven, Kenia Wiedemann, Jie Chao, William Finzer, Shiyan Jiang, Carolyn P. Rosé
Proceedings of the International Society of the Learning Sciences, 2021. PDF.

Phans, Stans and Cishets: Self-presentation Effects on Content Propagation in Tumblr

Michael Miller Yoder, Qinlan Shen, Yansen Wang, Alex Coda, Yunseok Jang, Yale Song, Kapil Thadani, Carolyn P. Rosé
Proceedings of the ACM Conference on Web Science (WebSci), 2020. PDF.

Perceptions of Censorship and Moderation Bias in Political Debate Forums

Qinlan Shen, Michael Miller Yoder, Yohan Jo, Carolyn P. Rosé
Proceedings of the International AAAI Conference on Web and Social Media (ICWSM), 2018. PDF.

Roles and Success in Wikipedia Talk Pages: Identifying Latent Patterns of Behavior

Keith Maki, Michael Miller Yoder, Yohan Jo, Carolyn P. Rosé
Proceedings of the International Joint Conference on Natural Language Processing (IJCNLP), 2017. PDF.

Best Student Paper Award.

Modeling Dialogue Acts with Content Word Filtering and Speaker Preferences

Yohan Jo, Michael Miller Yoder, Hyeju Jang, Carolyn P. Rosé
Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2017. PDF.

Metaphor Detection with Topic Transition, Emotion and Cognition in Context

Hyeju Jang, Yohan Jo, Qinlan Shen, Michael Miller, Shane Moon, Carolyn P. Rosé
Proceedings of the Association for Computational Linguistics (ACL), 2016. PDF.

JOURNAL
ARTICLES

Modes of toxic behavior and game design considerations in online multi-player games

Lynnette Hui Xian Ng, Adrian Xuan Wei Lim, Michael Miller Yoder
First Monday, 30(6), 2025.

The “chanification” of white supremacist extremism

Michael Miller Yoder, David West Brown, Kathleen M. Carley
Computational and Mathematical Organization Theory, 2024. PDF

Implicit Antisemitism and COVID-19 Conspiracy Theory on Twitter: Linking Narratives in a Mixed-Methods Study

Grace Garner, Madeleine McGrann, Maja Lynn, Daniel Klug, Rachel Kranson, Michael Miller Yoder
Jewish Film & New Media: An International Journal, 11(2), 183–213, 2023. HTML

An empirical analysis of high school students’ practices of modelling with unstructured data

Shiyan Jiang, Amato Nocera, Cansu Tatar, Michael Miller Yoder, Jie Chao, Kenia Wiedemann, William Finzer, Carolyn P. Rosé
British Journal of Educational Technology, 1–20, 2022. HTML.

Unpacking a political icon: “Bike lanes” and orders of indexicality

Michael Miller Yoder, Barbara Johnstone
Discourse & Communication, 12(2), 192–208, 2018.

WORKSHOP
PUBLICATIONS

A Fairness Analysis of Human and AI-Generated Student Reflection Summaries

Bhiman Kumar Baghel, Arun Balajiee Lekshmi Narayanan, Michael Miller Yoder
Proceedings of the Workshop on Gender Bias in Natural Language Processing (GeBNLP), 2024. PDF

Identity Construction in a Misogynist Incels Forum

Michael Miller Yoder, Chloe Perry, David West Brown, Kathleen M. Carley, Meredith L. Pruden
Proceedings of the Workshop on Online Abuse and Harms (WOAH), 2023. PDF

Online to Offline Crossover of White Supremacist Propaganda

Ahmad Diab, Bolor-Erdene Jagdagdorj, Lynnette Hui Xian Ng, Yu-Ru Lin, Michael Miller Yoder
International Workshop on Cyber Social Threats (CySoc), Companion Proceedings of the ACM Web Conference (WWW), 2023.

FanfictionNLP: A Text Processing Pipeline for Fanfiction

Michael Miller Yoder, Sopan Khosla, Qinlan Shen, Aakanksha Naik, Huiming Jin, Hariharan Muralidharan, Carolyn P. Rosé
Proceedings of the Workshop on Narrative Understanding (WNU), 2021. PDF.

Code-Switching as a Social Act: The Case of Arabic Wikipedia Talk Pages

Michael Miller Yoder, Shruti Rijhwani, Carolyn P. Rosé, Lori Levin
Proceedings of the Workshop on Natural Language Processing and Computational Social Science (NLP+CSS), 2017. PDF.

WHITEPAPERS &
DEMONSTRATIONS

Misogyny Surfaces as Top Issue in Survey of In-game Experiences

Michael Miller Yoder, Eleanor Young, Jasmin Narisetty, Sam Kulikowski, Will Flanagan, Claire Schuler, Caroline Guo, Zachary Yang
Whitepaper, Collaboratory Against Hate, 2024. PDF.

Research Needs for Countering Extremist Hate

Michael Miller Yoder, Hana Habib

Whitepaper, Collaboratory Against Hate, 2022. PDF.

StoryQ: A Web-Based Machine Learning and Text Mining Tool for K-12 Students

Jie Chao, Bill Finzer, Carolyn P. Rosé, Shiyan Jiang, Michael Miller Yoder, James Fiocco, Chas Murray, Cansu Tatar, Kenia Wiedemann

SIGCSE 2022: Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 2, 1178, 2022.

CONFERENCE PRESENTATIONS

Challenges for Real-Time Toxicity Detection in Online Games

Lynnette Hui Xian Ng, Adrian Xuan Wei Lim, Michael Miller Yoder

Ethical Games Conference, virtual, 2024.

The “chanification” of white supremacist extremism

Michael Miller Yoder, Kathleen M. Carley

Conference of the Center for Informed Democracy and Social Cybersecurity (IDeaS), Pittsburgh, PA, 2023.

Digital Chameleons: Identity Signal Changes in COVID-19 Discourse on Twitter

Samantha C. Phillips, Scott Leo Renshaw, Michael Miller Yoder, Kathleen M. Carley

Conference of the Center for Informed Democracy and Social Cybersecurity (IDeaS), Pittsburgh, PA, 2023.

The Relationship Between Antisemitism and COVID-19 Conspiracies on Twitter

Grace Garner, Madeleine McGrann, Daniel Klug, Rachel Kranson,

Michael Miller Yoder

International Communication Association Conference, Toronto, Canada, 2023.

The Spatiotemporal Reach of Hate Narratives

Ahmad Diab, Michael Miller Yoder, Yu-Ru Lin

Eradicate Hate Global Summit 2022 Poster Session, Pittsburgh, PA, 2022.

Research Needs for Countering Extremist Hate

Michael Miller Yoder, Hana Habib

Eradicate Hate Global Summit 2022 Poster Session, Pittsburgh, PA, 2022.

“Everyone is gay and nothing hurts”: the presentation of queer relationships in fanfiction

Michael Miller Yoder, Luke Breitfeller, Carolyn P. Rosé

Association for Computers and the Humanities, Pittsburgh, PA, 2019.

Ron the Death Eater: plotting characterization from canon to fanfiction

Michael Miller Yoder, Qinlan Shen, James Fiocco, Carolyn P. Rosé

Society for Text & Discourse, New York, NY, 2019.

Identity and interaction on Tumblr

Michael Miller Yoder, Qinlan Shen, Yansen Wang, Alex Coda, Yunseok Jang, Yale Song, Kapil Thadani, Carolyn P. Rosé

Student Research Symposium

Language Technologies Institute, Carnegie Mellon University, 2019.

Best Poster Award.

Linguistic liminality: Tamazight, Arabic and French in Guigou, Morocco
Michael Miller
Student Research Symposium, Goshen College, 2013.
Audience Choice Award.

INVITED TALKS &
PANELS

Help, my research area is politicized! Perspectives from hate speech detection in NLP
NLP Seminar, University of Pittsburgh, Pittsburgh, PA, 2025.

Computational approaches to identities and language: case studies in online extremism
A.W. Mellon Digital Humanities Workshop, Pittsburgh, PA, 2024.

Using NLP to track white supremacist narratives over time
Carnegie Mellon University Data Science Club, Pittsburgh, PA, 2024.

Rethinking data: Subjectivity in the data science process and implications for bias in AI systems
School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, 2024.

Algorithms against hate: Mapping online identities and ideologies with computational techniques
Department of Mathematics and Computer Science, Duquesne University, Pittsburgh, PA, 2023.

Computational approaches to identities and language
A.W. Mellon Digital Humanities Workshop, Pittsburgh, PA, 2023.

Beyond a hate speech binary
Social Cybersecurity Working Group (virtual), 2022.

Computational models for social good: Beyond bias and representation
(panel)
Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation (SBP-BRIMS), Pittsburgh, PA, 2022.

Creating a resource directory for crisis *(panel)*
Eradicate Hate Global Summit, Pittsburgh, PA, 2022.

Detecting hate speech with natural language processing
IDEaS Summer Institute, Carnegie Mellon University, Pittsburgh, PA, 2022.

Computational models of identity presentation in language
Department of Linguistics, University of Pittsburgh, Pittsburgh, PA, 2020.

ADVISING

Thesis committees:

- Master's Thesis Committee, Aparna Srinivasan, "Adapting Large Language Models to Genomic Data: Tokenization Methods and Applications for Whole-Genome Sequence Analysis", Intelligent Systems Program, University of Pittsburgh, April 2025
- Master's Thesis Committee, Deyasini Mitra, "Transfer Learning in Molecular Data Analysis: Unleashing the Potential of Large-Language Models", Intelligent Systems Program, University of Pittsburgh, April 2025
- Master's Thesis Committee, Raja Krishnaswamy, "A Framework for Concept Extraction from Course Presentations", Department of Computer Science, University of Pittsburgh, April 2024

Capstone projects:

- Master's capstone project, John Bowen, "Enhancing Document-Level English Machine Translation for Korean Literature", Department of Computer Science, University of Pittsburgh, April 2025

HONORS &
AWARDS

Best Reviewer Award

EMNLP 2023 Computational Social Science and Cultural Analytics

Outstanding Reviewer Award

ACL 2021

Best Poster Award

2019

Student Research Symposium

Language Technologies Institute, Carnegie Mellon University

Best Student Paper Award

2017

International Joint Conference on Natural Language Processing (IJCNLP)

NSF Graduate Research Fellowship

2015–2018

Awarded national fellowship with a focus area in natural language processing.

Audience Choice Award

2013

Student Research Symposium, Goshen College

President's Leadership Award Scholarship, Goshen College

2009–2013

Awarded top merit scholarship for leadership in academic studies.

ACADEMIC
SERVICE

Organizing:

- Collaboratory Against Hate Workshop on Practitioner Research Needs, 2022
- Eradicate Hate Global Summit 2022 Poster Session

Committees:

- Advisor, Technical Track, Hacking 4 Humanity 2025
- Judge Panel, Hacking 4 Humanity 2023
- Co-Director Search Committee, Collaboratory Against Hate, 2023
- Diversity, Equity, and Inclusion Committee, CMU Language Technologies Institute, 2020-2021
- Anti-Racism Working Group, CMU School of Computer Science Dean's PhD Advisory Committee, 2020-2021
- Update.org Data Science for Social Good Fellows Mentor Committee, 2019

Reviewing (NLP and Computational Social Science):

- ACL Rolling Review (2022-)
- CHI 2025
- SPELLL 2023
- *ACM Journal on Computing & Sustainable Societies*, 2023
- *EPJ Data Science*, 2023
- *Computational and Mathematical Organization Theory*, 2022, 2024

- USEC 2022
- EMNLP Computational Social Science and Cultural Analytics track 2022, 2023 (**Best Reviewer Award**)
- ACL 2021 (**Outstanding Reviewer Award**), 2023
- CoNLL 2020, 2022
- Bridging Human-Computer Interaction and NLP Workshop (HCI+NLP) 2021
- NAACL, ACL and ACL Student Research Workshops 2020, 2021, 2022, 2024
- Human and Machine-in-the-Loop Evaluation and Learning Strategies Workshop (HAMLETS @ NeurIPS) 2020

Reviewing (Digital Humanities):

- *Between Invention and Audience: DocuScope and the Rhetorical Approach to Corpus Analysis* (Studies in Corpus Linguistics, John Benjamins Press), 2022
- *Humanities* journal, special issue on fanfiction, 2022
- Association for Computers and the Humanities (ACH) 2021
- Digital Humanities (DH) 2020

INTERNATIONAL
EXPERIENCE

Bicycle Tourist

2013–2014

Paraguay, South America to Indiana, USA

Bicycled with a group of three over 10,000 miles from South America to the United States on an 11-month self-supported journey. Learned conversational Spanish and gained familiarity with Latin American culture.

Researcher and Service Worker

2012

Meknes and Guigou, Morocco

Studied Moroccan language and culture for 12 weeks while living with host families in urban and rural settings. Researched code-switching among Tamazight, French, Colloquial and Modern Standard Arabic speakers in the Middle Atlas Mountains.

Youth Camp Volunteer

2011

West Bank, Palestine & Israel

Studied and practiced Levantine and Modern Standard Arabic for 9 weeks. Participated in history and learning tours of the Israeli-Palestinian conflict while living with a Palestinian friend's family. Facilitated crafts and games for Palestinian youth.