

# Michael Miller Yoder

yoder@cs.cmu.edu

cs.cmu.edu/~mamille2

EDUCATION	<b>Carnegie Mellon University</b> , Pittsburgh, PA <i>Ph.D.</i> , Language and Information Technologies 2021 Advisor: Carolyn P. Rosé Thesis: “Computational models of identity presentation in language” PDF <i>M.S.</i> , Language Technologies 2017  <b>Goshen College</b> , Goshen, IN <i>B.A.</i> , Computer Science 2013 Minors: English, International Studies
RESEARCH INTERESTS	Natural language processing, computational social science, hate speech, language and identity, sociolinguistics, discourse analysis, digital humanities
CONFERENCE & WORKSHOP PUBLICATIONS	<b>A Weakly Supervised Classifier and Dataset of White Supremacist Language</b> Michael Miller Yoder, Ahmad Diab, David West Brown, Kathleen M. Carley <i>Proceedings of the Association for Computational Linguistics (ACL)</i> . 2023.  <b>Identity Construction on a Misogynist Incels Forum</b> Michael Miller Yoder, Chloe Perry, David West Brown, Kathleen M. Carley, Meredith Pruden <i>Proceedings of the Workshop on Online Abuse and Harms (WOAH)</i> . 2023.  <b>Online to Offline Crossover of White Supremacist Propaganda</b> Ahmad Diab, Bolor-Erdene Jagdagdorj, Lynnette Hui Xian Ng, Yu-Ru Lin, Michael Miller Yoder <i>International Workshop on Cyber Social Threats (CySoc), Companion Proceedings of the ACM Web Conference (WWW)</i> , 2023.  <b>How Hate Speech Varies by Target Identity: A Computational Analysis</b> Michael Miller Yoder, Lynnette Hui Xian Ng, David West Brown, Kathleen M. Carley <i>Proceedings of the Conference on Computational Natural Language Learning (CoNLL)</i> , 2022. PDF.  <b>FanfictionNLP: A Text Processing Pipeline for Fanfiction</b> Michael Miller Yoder, Sopan Khosla, Qinlan Shen, Aakanksha Naik, Huiming Jin, Hariharan Muralidharan, Carolyn P. Rosé <i>Proceedings of the Workshop on Narrative Understanding</i> , 2021. PDF.  <b>Modeling Unstructured Data: Teachers as Learners and Designers of Technology- enhanced Artificial Intelligence Curriculum</b> Cansu Tatar, Michael Miller Yoder, Madeleine Coven, Kenia Wiedemann, Jie Chao, William Finzer, Shiyan Jiang, Carolyn P. Rosé <i>Proceedings of the International Society of the Learning Sciences</i> , 2021. PDF.  <b>Phans, Stans and Cishets: Self-presentation Effects on Content Propagation in Tumblr</b> Michael Miller Yoder, Qinlan Shen, Yansen Wang, Alex Coda, Yunseok Jang, Yale Song, Kapil Thadani, Carolyn P. Rosé <i>Proceedings of the ACM Conference on Web Science (WebSci)</i> , 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.

**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.

**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

**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.

WHITEPAPERS &  
DEMONSTRATIONS

**The Relationship Between Antisemitism and COVID-19 Conspiracy on Twitter**  
Grace Garner, Madeleine McGrann, Maja Lynn, Rachel Kranson, Michael Miller Yoder, Daniel Klug  
Whitepaper, Collaboratory Against Hate, 2022. 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

**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.

**The Relationship Between Antisemitism and COVID-19 Conspiracy on Twitter**

Grace Garner, Madeleine McGrann, Maja Lynn, Rachel Kranson, Daniel Klug,  
Michael Miller Yoder  
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

**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.

TEACHING	<b>Computational Ethics for Natural Language Processing</b> <i>Teaching Assistant</i> Carnegie Mellon University	<b>Spring 2020</b>
	<b>Applied Machine Learning</b> <i>Teaching Assistant</i> Carnegie Mellon University	<b>Spring 2019</b>
	<b>Introduction to Artificial Intelligence</b> <i>Tutor</i> Goshen College	<b>Spring 2013</b>
PROFESSIONAL EXPERIENCE	<b>Postdoctoral Fellow</b> 2022–present Researcher with the Collaboratory Against Hate, an interdisciplinary initiative of Carnegie Mellon University and the University of Pittsburgh. Developing new methods for tracking and detecting hate speech in online contexts. Advised by Prof. Kathleen M. Carley in the Software and Societal Systems Department and Prof. David West Brown in the Department of English.	Carnegie Mellon University Pittsburgh, PA
	<b>Graduate Research Assistant</b> 2015–2021 Apply 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.	Carnegie Mellon University Pittsburgh, PA
	<b>Software Engineering Intern</b> 2019 Detected communities of Tumblr users from social network structure. Explored community-specific language use. Advised by Dr. Kapil Thadani.	Yahoo Research/Tumblr New York, NY
	<b>Research Programmer/Analyst</b> 2015 Designed and implemented neuroimaging database to support machine learning research. Advised by Prof. Timothy Verstynen in the Department of Psychology.	Carnegie Mellon University Pittsburgh, PA
	<b>Software Developer &amp; IT Administrator</b> 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.	GSAREH Austin, TX area
HONORS & AWARDS	<b>NSF Graduate Research Fellowship</b> 2015–2018 Awarded national fellowship with a focus area in natural language processing.	
	<b>President’s Leadership Award Scholarship, Goshen College</b> 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:*

- 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

*Reviewing (NLP and Computational Social Science):*

- *ACM Journal on Computing & Sustainable Societies*, 2023
- *EPJ Data Science*, 2023
- *Computational and Mathematical Organization Theory*, 2022
- USEC 2022
- EMNLP 2022 Computational Social Science track
- ACL Rolling Review (2022- )
- ACL 2021 (**Outstanding Reviewer Award**), 2023
- CoNLL 2020, 2022
- Bridging Human-Computer Interaction and NLP Workshop (HCI+NLP) 2021
- NAACL, AACL and ACL Student Research Workshops 2020, 2021, 2022
- 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.