

## Michael Miller Yoder

yoder@cs.cmu.edu  
319-804-8987  
cs.cmu.edu/~mamille2

5003 Broad St  
Pittsburgh, PA 15224

---

EDUCATION	<b>Carnegie Mellon University</b> , Pittsburgh, PA <i>Ph.D.</i> , Language and Information Technologies Advisor: Carolyn P. Rosé Expected 2021
	<i>M.S.</i> , Language Technologies 2017
	<b>Goshen College</b> , Goshen, IN <i>B.A.</i> , Computer Science Minors: English, International Studies 2013
RESEARCH INTERESTS	Identity presentation in language, social factors in language, computational sociolinguistics, natural language processing, discourse analysis, digital humanities
CONFERENCE PUBLICATIONS	<b>Perceptions of Censorship and Moderation Bias in Political Debate Forums</b> Qinlan Shen, Michael Miller Yoder, Yohan Jo, Carolyn P. Rosé <i>Proceedings of ICWSM 2018</i> .
	<b>Roles and Success in Wikipedia Talk Pages: Identifying Latent Patterns of Behavior</b> Keith Maki, Michael Miller Yoder, Yohan Jo, Carolyn P. Rosé <i>Proceedings of IJCNLP 2017</i> . <b>Best Student Paper Award</b> .
	<b>Modeling Dialogue Acts with Content Word Filtering and Speaker Preferences</b> Michael Miller Yoder, Shruti Rijhwani, Carolyn P. Rosé, Lori Levin <i>Proceedings of the Second Workshop on Natural Language Processing and Computational Social Science, 2017</i> .
	<b>Code-Switching as a Social Act: The Case of Arabic Wikipedia Talk Pages</b> Keith Maki, Michael Miller Yoder, Yohan Jo, Carolyn P. Rosé <i>Proceedings of EMNLP 2017</i> .
	<b>Metaphor Detection with Topic Transition, Emotion and Cognition in Context</b> Hyeju Jang, Yohan Jo, Qinlan Shen, Michael Miller, Shane Moon, Carolyn P. Rosé <i>Proceedings of ACL 2016</i> .
	<b>Unpacking a Political Icon: “Bike Lanes” and Orders of Indexicality</b> Michael Miller Yoder, Barbara Johnstone <i>Discourse &amp; Communication 12(2), 192–208, 2018</i> .
PROJECTS	<b>Hate Speech Identification:</b> Developed hate speech identification system without relying on derogatory terms. Improved identification performance on two Twitter hate speech datasets and provided analysis of potential improvements for better generalization to other hate speech domains.
	<b>Identity Performance on Tumblr:</b> Designed task and assembled dataset of social media content related to self-identification among age, gender/sexuality, personality type, ethnicity/nationality, and other dimensions of identity. Implemented machine learning system to identify social media content related to each dimension, as well as predictors of influence based on self-identification.

**Gender/Sexuality Representations in Fanfiction:** Designed embedding approaches for gender/sexuality identity terms from text and metadata of fanfiction. Led team of master's students using NLP with fanfiction data to address questions from digital humanities.

## EXPERIENCE

### Graduate Research Assistant

Carnegie Mellon University

2015–present

Pittsburgh, PA

Apply computational models of discourse to study social meaning in language on social media. Advised by Prof. Carolyn P. Rosé in the Language Technologies Institute.

### Research Programmer/Analyst

Carnegie Mellon University

2015

Pittsburgh, PA

Designed and implemented neuroimaging database to support machine learning research. Advised by Prof. Timothy Verstynen in the Department of Psychology.

### Software Developer & IT Admin.

GSAREH

2014

Austin, TX area

Developed applications to recognize and map named entities in English and Arabic articles and social media. Managed software development and IT administration of small GIS startup, named Geospatial Applied Research Experts House.

## GLOBAL 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.