

Xinchen Yu

CONTACT INFORMATION	1040 E 4th St, Rm 829 Tucson, AZ 85719	xinchenyu@arizona.edu xinchenyu.github.io
RESEARCH INTERESTS	My research lies at the intersection of <b>natural language processing</b> (NLP) and <b>computational social science</b> (CSS). The primary goal is to <b>make computers understand and facilitate better communication</b> between people. My latest work centers on <b>explaining, predicting, and preventing</b> incivility in user-generated content. Specifically, I enhance the robustness of NLP systems in identifying <b>hate speech and counter speech</b> . Additionally, I develop reliable frameworks to <b>measure and forecast the incivility of conversations</b> , as well as bridge the gap between computer science and computational social science to facilitate the use of cutting-edge AI models in other disciplines, including communication and social science.	
ACADEMIC APPOINTMENTS	<b>Assistant Professor of Practice</b> , Department of Computer Science University of Arizona, Tucson, AZ	Aug. 2023 - present
EDUCATION	<b>Ph.D. in Information Science</b> (Data Science Concentration) University of North Texas, Denton, TX Committee: Lingzi Hong (Chair), Eduardo Blanco, Junhua Ding, Diyi Yang Thesis: Countering Hate Speech: Modeling User-generated Web Content Using Natural Language Processing (NLP)	Jul. 2023
	<b>M.S. in Information Science</b> University of Pittsburgh, Pittsburgh, PA	May 2019
	<b>B.S. in Library and Information Science</b> Zhejiang University, Hangzhou, China	May 2017
HONORS AND AWARDS	UNT Margaret Irby Nichols Endowed Scholarship (\$1,000)	2023
	UNT Dewey E. Carroll Endowed Graduate Fellowship (\$4,000, top 1)	2022
	UNT Graduate Student Research Awards (\$1,200)	2021, 2022
	UNT Department of Information Science Travel Grant	2020, 2022
	UNT COVID-19 Student Success Award (\$1,000)	2020, 2021
	ZJU Excellent Student Awards	2014-2017
	ZJU Hengyi Scholarship (top 1/326)	2016
PUBLICATIONS	ZJU Real Estate Fund Scholarship	2015
	<b>Xinchen Yu</b> , Hayden Arnold, Benjamin Su, Eduardo Blanco (2025). Measuring and Forecasting Conversation Incivility: the Role of Antisocial and Prosocial Behaviors. To Appear at <u>ICWSM'25</u> ( <b>Early acceptance rate: 9.2%</b> ).	
	Xiaoying Song, Sharon Lisseth Perez, <b>Xinchen Yu</b> , Eduardo Blanco , Lingzi Hong (2025). Echoes of Discord: Forecasting Hater Reactions to Counterspeech. To Appear at Findings of <u>NAACL'25</u> .	
	<b>Xinchen Yu</b> , Eduardo Blanco, Lingzi Hong (2024). Hate Cannot Drive out Hate: Forecasting Conversation Incivility following Replies to Hate Speech. In <i>Proceedings of the 18th International AAAI Conference on Web and Social Media (ICWSM'24)</i> . Vol. 18, pp. 1740-1752. Buffalo, NJ, USA.	

**Xinchen Yu**, Ashley Zhao, Eduardo Blanco, Lingzi Hong (2023). A Fine-Grained Taxonomy of Replies to Hate Speech. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP'23)*. pp. 7275–7289. Singapore.

**Xinchen Yu**, Zhuoli Xie, Afra Mashhadi, Lingzi Hong (2022). Multi-task Models for Multi-faceted Classification of Pandemic Information on Social Media. In *Proceedings of the 14th ACM Web Science Conference (WebSci'22)*. pp. 327-335. Barcelona, Spain.

**Xinchen Yu**, Eduardo Blanco, Lingzi Hong (2022). Hate Speech and Counter Speech Detection: Conversational Context Does Matter. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL'22)*. pp. 5918–5930. Seattle, WA, USA.

**Xinchen Yu**, Jeremy Boy, Rene Clausen Nielsen, Lingzi Hong (2022). Linguistic Characteristics of Social Media Messages Spreading across Geographic and Linguistic Boundaries. In *Proceedings of the 9th European Conference on Social Media (ECSM'22)*. pp. 211-218. Virtual.

**Xinchen Yu**, Afra Mashhadi, Jeremy Boy, Rene Clausen Nielsen, Lingzi Hong (2022). Causal Impact Model to Evaluate the Diffusion Effect of Social Media Campaigns. In *Proceedings of 20th European Conference on Computer-Supported Cooperative Work (EUSSET'22)*. Vol. 6, no. 1. Coimbra, Portugal.

Alexander Kahanek, **Xinchen Yu**, Lingzi Hong, Ana Cleveland, Jodi Philbrick (2021). Temporal Variations and Spatial Disparities in Public Sentiment Toward COVID-19 and Preventive Practices in the United States: Infodemiology Study of Tweets. In *JMIR Infodemiology*. 2021;1(1):e31671.

Lingzi Hong, William Moen, **Xinchen Yu**, Jiangping Chen (2021). The Disciplinary Research Landscape of Data Science Reflected in Data Science Journals. In *Journal of Information Discovery and Delivery*. Vol. 49, pp. 287-297.

**Xinchen Yu**, Daida Shashidhar Reddy, Lasya Bentula, Lingzi Hong (2020). Characteristics of Information Spreading across Nations. In *Proceedings of the Association for Information Science and Technology (ASIS&T'20)*. Vol 57:e309. Pittsburgh, PA, USA.

**Xinchen Yu**, Shashidhar Reddy Daida, Jeremy Boy, Lingzi Hong (2020). The Effect of Structural Affinity on the Diffusion of a Transnational Online Movement: The Case of #MeToo. In *Proceedings of the 12th International Conference on Social Informatics (SocInfo'20)*. pp. 447-460. Virtual.

## TEACHING

**Instructor of Record** at the University of Arizona.

- CSC 110: Introduction to Computer Programming I (Fall 2023 - Spring 2025)
- CSC 380: Principles of Data Science (Fall 2023, Spring 2024)

**Instructor of Record** at the University of North Texas.

- INFO 4670: Data Analysis and Knowledge Discovery (Fall 2022)

**Teaching Assistant** at the University of North Texas.

- INFO 3010: Introduction to Data Science (Fall 2020, Fall 2021)
- INFO 5709: Data Visualization and Communication (Spring 2021, Spring 2022)

## SERVICE

Mentoring and Advising

- Hayden Arnold, UofA, undergraduate student → KAIST

- Benjamin Su, UofA, undergraduate student → USC
- Selina Lu, UofA, undergraduate student → CMU
- Ashley Zhao, TAMS, high school student → Cornell

#### Conference and Workshop Organizing

- Area Chair of the 31st International Conference on Computational Linguistics (COLING 2025): Offensive Speech Detection and Analysis track
- Organizer of the 7th Women in Data Science Tucson Conference (WiDS-Tucson 2025), [wids-tucson.github.io](https://wids-tucson.github.io)
- Organizer and Moderator of the 2025 Workshop on Considering Cultural and Linguistic Diversity in AI Applications, [cald-ai.github.io](https://cald-ai.github.io)

#### Reviewing

- ACL Rolling Review (emergency reviewer)
- The Annual Meeting of the Association for Computational Linguistics (ACL 2024)
- The Conference on Empirical Methods in Natural Language Processing (EMNLP 2023, 2024)
- The Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)
- The ACM Conference on Human Factors in Computing Systems (CHI 2024)
- Computational Linguistics
- Computer Standards & Interfaces
- Information Discovery and Delivery
- The Electronic Library

#### To University of Arizona

- |  |                       |
|--|-----------------------|
| • Faculty Advisor, Women in Information and Computer Science | Fall 2023-Spring 2025 |
| • Member, Transfer Student Support Committee                 | Fall 2023-Spring 2024 |
| • Member, Intro Sequence Committee                           | Spring 2025           |
| • Member, Career-Track Evaluation Committee                  | Spring 2025           |

#### INVITED TALKS

##### **Countering Hate Speech: Detection, Strategies, and Forecasting.**

- Lucy Wang's Group Seminar, University of Washington, Oct. 2024

##### **Forecasting Conversation Incivility following Replies to Hate Speech.**

- ICWSM, Jun. 2024

##### **A Fine-Grained Taxonomy of Replies to Hate Speech**

- EMNLP, Dec. 2023

##### **Hate speech and counter speech detection.**

- AJCAI Workshop on Toxic Language Detection (TLD), Dec, 2022
- NAACL, *Oral Presentation*, Aug. 2022

##### **Multi-Faceted Classification of Pandemic Information.**

- WebSci, *Oral Presentation*, Jun. 2022

##### **Characteristics of Information Spreading across Nations.**

- ASIS&T, *Oral Presentation*, Oct. 2020