

# Caleb Ziems

## Curriculum Vitae

✉ [cziems@stanford.edu](mailto:cziems@stanford.edu)  
🌐 [www.calebziems.com](http://www.calebziems.com)  
🐦 [cziems](https://twitter.com/cziems)  
🌐 [cziems](https://github.com/cziems)

I work on **socially intelligent** and **equitable NLP** systems  
that are aware of **social norms** and **robust to language  
variation**

### Education

- 2023–Present **Stanford University**, *Ph.D. in Computer Science*.  
Advisor: Diyi Yang
- 2020–2023 **Georgia Institute of Technology**, *M.S. in Computer Science (Ph.D. track)*.  
Advisor: Diyi Yang
- 2016–2020 **Emory University**, *B.S. in Computer Science (highest honors), B.A. in Mathematics*.  
Advisor: Ymir Vigfusson

### Pre-prints

- 2024 **Ziems, C.**,\* Yang, D.,\* Held, W.,\* Shaikh, O.,\* Bernstein, M.S., Mitchell, J. (2024). *Social Skill Training with Large Language Models*. Under Review. [PDF](#)
- 2024 Shi, W., Li, R., Zhang, Y., **Ziems, C.**, Yu, C., Horesh, R., Abreu de Paula, R., Yang, D. (2024). *CultureBank: An Online Community-Driven Knowledge Base Towards Culturally Aware Language Technologies*. Under Review. [PDF](#)

### Publications

- 2024 **Ziems, C.**, Held, W., Dwivedi-Yu, J., & Yang, D. (2024). *Measuring and Addressing*  
**ACL** *Indexical Bias in Information Retrieval*. In Proceedings of the 61st Annual Meeting of  
*Findings* the Association for Computational Linguistics. Findings.  
[PDF](#)
- 2024 Li, M., Shi, W., **Ziems, C.**, & Yang, D. (2024). *Social Intelligence Data Infrastructure:*  
**ACL** *Structuring the Present and Navigating the Future*. In Proceedings of the 61st Annual  
*Findings* Meeting of the Association for Computational Linguistics. Findings.  
[PDF](#)
- 2024 Kruk, J., Marchini, M., Magu, R., **Ziems, C.**, Muchlinski, D., & Yang, D. (2024). *Silent*  
**ACL** *Signals, Loud Impact: LLMs for Word-Sense Disambiguation of Coded Dog Whistles*.  
In Proceedings of the 61st Annual Meeting of the Association for Computational  
Linguistics.  
[PDF](#)
- 2024 **Ziems, C.**, Held, W., Shaikh, O., Chen, J., Zhang, Z., & Yang, D. (2023). *Can Large*  
**CL** *Language Models Transform Computational Social Science?* In Computational  
Linguistics, 49(4).  
Included in the curriculum at Dartmouth, GaTech, Purdue, UTAustin, UToronto, UVA.  
[PDF](#)

- 2023 **EMNLP** Kruk, J., **Ziems, C.**, & Yang, D. (2023). *Impressions: Visual Semiotics and Aesthetic Impact Understanding*. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing.  
23% Accept. Rate  
[PDF](#)
- 2023 **EMNLP** Li, E., Shi, T., **Ziems, C.**, Kan, M., Chen, N., Liu, Z., & Yang, D. (2023). *CoAnnotating: Uncertainty-Guided Work Allocation between Human and Large Language Models for Data Annotation*. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing.  
23% Accept. Rate  
[PDF](#)
- 2023 **ACL** **Ziems, C.**,\* Held, W.,\* Yang, J., & Yang, D. (2023). *Multi-VALUE: A Framework for Cross-Dialectal English NLP*. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics.  
22% Accept. Rate  
[PDF](#)
- 2023 **ACL** **Ziems, C.**, Dwivedi-Yu, J., Wang, Y., Halevy, A. Y., & Yang, D. (2023). *NormBank: A Knowledge Bank of Situational Social Norms*. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics.  
22% Accept. Rate  
[PDF](#)
- 2023 **ACL Findings** **Ziems, C.**,\* Shaikh, O.,\* Held, W., Pariani, A. J., Morstatter, F., & Yang, D. (2023). *Modeling Cross-Cultural Pragmatic Inference with Codenames Duet*. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics: Findings.  
[PDF](#)
- 2023 **ACL Findings** Held, W., **Ziems, C.**, & Yang, D. (2023). *TADA: Task-Agnostic Dialect Adapters for English*. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics: Findings.  
[PDF](#)
- 2022 **ACL** **Ziems, C.**, Chen, J., Harris, C., Anderson, A., & Yang, D. (2022). *VALUE: Understanding Dialect Disparity in NLU*. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing.  
21% Accept. Rate  
[PDF](#)
- 2022 **ACL** **Ziems, C.**, Yu, J. A., Wang, Y., Halevy, A. Y., & Yang, D. (2022). *The Moral Integrity Corpus: A Benchmark for Ethical Dialogue Systems*. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing.  
21% Accept. Rate  
[PDF](#)
- 2022 **ACL** **Ziems, C.**,\* Li, M.,\* Zhang, A., & Yang, D. (2022). *Inducing Positive Perspectives with Text Reframing*. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing.  
21% Accept. Rate  
[Outstanding Paper Award \(Top .8%\).](#)  
[PDF](#)

- 2021 **EMNLP** **Ziems, C.,\*** ElSherief, M.,\* Muchlinski, D., Anupindi, V., Seybolt, J., De Choudhury, M., & Yang, D. (2021). *Latent Hatred: A Benchmark for Understanding Implicit Hate Speech*. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing.  
25% Accept. Rate  
[PDF](#)
- 2021 **EMNLP Findings** **Ziems, C.,** Yang, D. (2021). *To Protect and To Serve? Analyzing Entity-Centric Framing of Police Violence*. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: Findings.  
*Included in the curriculum at Stanford.*  
[PDF](#)
- 2021 **ASONAM** He, B., **Ziems, C.,** Soni, S., Ramakrishnan, N., Yang, D., & Kumar, S. (2021). *Racism is a Virus: Anti-Asian Hate and Counterhate in Social Media during the COVID-19 Crisis*. In 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE.  
22% Accept. Rate  
*Included in the curriculum at CMU, GaTech, Stanford, and Tübingen.*  
[PDF](#)
- 2020 **ICWSM** **Ziems, C.,** Vigfusson, Y., & Morstatter, F. (2020) *Aggressive, Repetitive, Intentional, Visible, and Imbalanced: Refining Representations for Cyberbullying Classification*. In Proceedings of the 14th International AAAI Conference on Web and Social Media.  
24% Accept. Rate  
*Honorable Mention for Best Paper Award (Top 11%).*  
[PDF](#)

## Experience

- Sept 2023 - **Stanford NLP Group**, *Graduate Research Assistant*.  
Present Advisor: Diyi Yang
- May 2022 - **Meta AI Integrity**, *Research Intern*.  
August 2022 Manager: Alon Halevy  
Peers: Jane Yu, Yi-Chia Wang
- May 2021 - **Meta AI Integrity**, *Research Intern*.  
Sep 2021 Manager: Alon Halevy  
Peers: Jane Yu, Yi-Chia Wang
- August 2020 - **Georgia Tech SALT Lab**, *Graduate Research Assistant*.  
August 2023 Advisor: Diyi Yang
- March 2020 - **Georgia Tech CLAWS Research Group**, *Undergraduate Research Assistant*.  
May 2020 Advisor: Srijan Kumar
- August 2019 - **Emory SimBioSys Lab**, *Undergraduate Research Assistant*.  
March 2020 Advisor: Ymir Vigfusson
- June 2019 - **USC/ISI MINDS Research Group**, *NSF REU Fellow*.  
August 2019 Advisor: Fred Morstatter
- June 2018 - **Stanford InfoLab**, *NSF REU Fellow*.  
August 2018 Advisor: Srijan Kumar  
PI: Jure Leskovec
- August 2017 - **Emory Concept Mining Lab**, *Undergraduate Research Assistant*.  
May 2018 Advisor: Phillip Wolff

## Teaching

- January 2024 **Stanford CS 224N – NLP with Deep Learning**, *Teaching Assistant*.  
- March 2024
  - Pretested homework assignments
  - Graded homework assignments and project reports (~520 students)
  - Held weekly office hours
- January 2022 **Georgia Tech CS 6471 – Computational Social Science**, *Teaching Assistant*.  
- May 2022
  - Co-designed the CS 6471 curriculum
    - Co-designed the course schedule of assigned readings
    - Co-designed project report and presentation rubrics
    - Designed quizzes and bi-weekly reading response questions
  - Graded reports and presentations
  - Held weekly office hours
- August 2017 - **Emory MATH 111, 112 – Calculus**, *EPASS Peer Tutor*.  
May 2018
  - Worked one-on-one with 10 to 20 different undergraduates each semester
  - Explained Calculus I and II concepts, designed problem sets, and offered individual feedback

## Mentoring

- Spring 2024 **Yicheng Fu**, *Stanford University* (M.S.)
- Spring 2024 **William Z. Li**, *Stanford University* (B.S., M.S.)  
*Topic*: Social Skill Training
- Fall 2023 - **Vyoma Raman**, *Stanford University* (M.S.)  
Present *Topic*: Power Dynamics in NLP
- Fall 2023 - **Seonghee Lee**, *Stanford University* (M.S.)  
Present *Topic*: LLM Agent Consistency
- Summer 2023 **Chijioke Mgbahurike**, *Stanford University* (B.S.)  
*Topic*: Dialectal Bias in ASR Systems  
*Result*: One presentation at the INSPIRE Research Poster Session.
- Fall 2022 - **Julia Kruk**, *Georgia Institute of Technology* (M.S.)  
Present *Topic*: Computational Aesthetics and Photojournalism  
*Results*: One EMNLP 2023 Paper; One ACL 2024 Paper.
- Spring 2021 - **Minzhi (Ella) Li**, *National University of Singapore* (B.S., Ph.D)  
Present *Topic*: Positive Reframing; Human-LLM Co-Annotation  
*Results*: One ACL 2022 **Outstanding Paper**; One EMNLP 2023 Paper; One ACL 2024 Paper.
- Spring 2021 - **Anthony Zhang**, *Georgia Institute of Technology* (B.S.)  
Fall 2021 *Topic*: Positive Reframing  
*Results*: One ACL 2022 **Outstanding Paper**
- Spring 2022 - **Aryan Pariani**, *Georgia Institute of Technology* (B.S.)  
Spring 2023 *Topic*: Pragmatic Reasoning in Language Games  
*Results*: One ACL 2023 Findings Paper.
- Fall 2021 - **Taiwei Shi**, *Georgia Institute of Technology* (B.S.)  
Spring 2022; *Topics*: Dialect-Aware NLP; Human-LLM Co-Annotation  
Summer 2023 *Results*: Undergraduate Thesis; One EMNLP 2023 Paper.

---

## Awards and Honors

- May 2023 **Magic Grant**, *The Brown Institute for Media Innovation*.
- May 2022 **Outstanding Paper**, *ACL 2022*.
- March 2021 **NSF Graduate Research Fellowship**, *National Science Foundation*.
- June 2020 **Honorable Mention for Best Paper**, *ICWSM 2020*.
- May 2020 **Summa Cum Laude**, *Emory University Computer Science Department*.
- May 2020 **Academic Excellence Award**, *Emory University Computer Science Department*.
- November 2018 **Phi Beta Kappa**, *Emory University*.  
Elected as a Junior
- August 2017 **Dean's Achievement Scholarship**, *Emory University*.  
\$30,000 scholarship recognizing academic achievement, leadership, and community involvement

---

## Service

- Reviewing ACL / ARR Review (2024, 2023, 2022, 2021);  
ACM CSCW (2021)
- Volunteer ACL (2023); EMNLP (2021)

---

## Talks and Miscellanea

- April 15, 2024 **Guest Lecturer** for CS 6501 Natural Language Processing at UVA.  
*Talk*: Can Large Language Models Transform Computational Social Science?
- March 19, 2024 **Guest Lecturer** for LIN 393 Social Applications and Impact of NLP at UT Austin.  
*Talk*: Can Large Language Models Transform Computational Social Science?
- October 9, 2023 **Speaker** at the AWS Responsible AI Group.  
*Talk*: Multi-VALUE: A Framework for Cross-Dialectal English NLP
- June 19, 2023 **Speaker** at the Summer Institutes in Computational Social Science (SICSS).  
*Talk*: Can Large Language Models Transform Computational Social Science?
- October 4, 2022 **Panel Member** for CS 7001 Introduction to Graduate Studies at Georgia Tech.  
*Topic*: National Fellowships
- February 24, 2022 **Student Host** for Yian Yin, Faculty Candidate in the School of Interactive Computing  
*Topic*: Science of Science and Network Science
- February 14, 2022 **Speaker** at IST-Unbabel Seminar  
*Talk*: Social NLP for Positive and Civil Interactions
- November 17, 2021 **Guest Lecturer** for CS 4650 Natural Language Processing at Georgia Tech  
*Topic*: Dialect-Aware NLP + Computational Social Science
- October 5, 2021 **Panel Member** for CS 7001 Introduction to Graduate Studies at Georgia Tech.  
*Topic*: National Fellowships