

Katy Felkner

felkner@isi.edu | <http://www.katyfelkner.com> | (405) 808-2703

EDUCATION

University of Southern California

Ph.D. in Computer Science

Los Angeles, CA

Aug. 2020 – Present

- Expected May 2027; GPA: 3.88/4.00; Supported by **NSF Graduate Research Fellowship**
- Coursework: Advanced Natural Language Processing, Fairness in AI, Probability and Statistics for Data Science, Representation Learning in NLP, Advanced Algorithms, Crafting a Research Agenda in Computer Science

University of Oklahoma

B.S. in Computer Science (summa cum laude), B.A. in Letters (summa cum laude)

Norman, OK

Aug. 2016 – May 2020

- Minor: Mathematics; GPA: 4.00/4.00

RESEARCH INTERESTS

Research agenda focuses on developing high-quality, **well-grounded benchmark datasets** for robust evaluation of language models, with particular emphases on participatory, community-engaged methods and on rigorous measurement of fairness and social bias.

PUBLICATIONS

Felkner, Virginia K., Allison Lim, and Jonathan May (Oct. 2025). *Textual Entailment and Token Probability as Bias Evaluation Metrics*. Under review. arXiv: 2510.07662 [cs.CL]. URL: <https://arxiv.org/abs/2510.07662>.

Felkner, Virginia K., Jennifer Thompson, and Jonathan May (Aug. 2024). “GPT is Not an Annotator: The Necessity of Human Annotation in Fairness Benchmark Construction”. In: *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*. DOI: 10.18653/v1/2024.acl-long.760.

Felkner, Virginia K., Ho-Chun Herbert Chang, et al. (July 2023). “WinoQueer: A Community-in-the-Loop Benchmark for Anti-LGBTQ+ Bias in Large Language Models”. In: *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. DOI: 10.18653/v1/2023.acl-long.507.

Felkner, Virginia K., Ho-Chun Herbert Chang, et al. (July 2022). “Towards WinoQueer: Developing a Benchmark for Anti-Queer Bias in Large Language Models”. In: *QueerInAI Affinity Group Workshop*. NAACL 2022. arXiv: 2206.11484 [cs.CL]. URL: <https://arxiv.org/abs/2206.11484>.

Weideman, Nicolaas, **Virginia K. Felkner**, et al. (Nov. 2021). “PERFUME: Programmatic Extraction and Refinement for Usability of Mathematical Expression”. In: *Proceedings of the 2021 Research on Offensive and Defensive Techniques in the Context of Man At The End (MATE) Attacks*. ACM SIGSAC CCS 2021. DOI: 10.1145/3465413.3488575.

Felkner, Virginia K. and Elisabeth (Lissa) Moore (Nov. 2020). “Investigating the Efficacy of Unstructured Text Analysis for Node Failure Detection in Syslog”. In: *Second Workshop on Machine Learning for Computing Systems*. Supercomputing 2020. URL: https://mlcsworkshop.weebly.com/uploads/1/1/6/9/116940258/ws_mlcs103s2-file1.pdf.

TECHNICAL SKILLS

Programming Languages and Skills: Python, C, C++, R, SQL, Bash Scripting, HPC, Slurm

ML/NLP/Data Science Tools: HuggingFace, PyTorch, TensorFlow, SpaCy, Pandas, Numpy, Scikit-Learn, NLTK

RESEARCH EXPERIENCE

LLM Fairness and Bias, USC Information Sciences Institute

Graduate Research Assistant

Los Angeles, CA

Jan. 2022 – Present

- Develop two community-sourced benchmark datasets for detecting anti-LGBTQ+ and antisemitic bias in LLMs.
- Design and distribute two community surveys and analyze 500 responses to inform LLM bias benchmarks.
- Lead team of 3 to collect 8M sentences of community data, then train 32 debiased models via continued pretraining.
- Assess LLM capability to assist in fairness benchmark construction and conduct detailed human analysis of model failures.
- Analyze viability of natural language inference (NLI) as an LLM bias metric, developing a novel NLI bias dataset and conducting detailed study of differences between token probability and NLI bias metrics.
- **Publications:** Lead interdisciplinary teams to publish four papers (two ACL, one workshop, one under review).
- **Mentoring:** Mentor undergraduate student on NLI bias metric project.
- **Advisor/PI:** Jonathan May

DARPA ReMATH, USC Information Sciences Institute	Los Angeles, CA
<i>Graduate Research Assistant</i>	Jan. 2021 – Jun. 2022
<ul style="list-style-type: none"> ○ Trained neural machine translation (NMT) models for extraction of symbolic math from assembly code and bit-vector expressions. ○ Developed novel metrics for evaluation of NMT-generated mathematical expressions. ○ Publication: Wrote relevant sections of workshop paper accepted to CheckMATE, ACM CCS 2021. ○ PI: Luis Garcia 	
Border Translation Project, USC Information Sciences Institute	Los Angeles, CA
<i>Graduate Research Assistant</i>	Aug. 2020 – Dec. 2021
<ul style="list-style-type: none"> ○ Built and evaluated NMT systems for extremely low-resource languages. ○ Developed baseline NMT system for K'iche', a Mayan indigenous language, with intended application of helping NGOs provide legal assistance to immigrants and asylum seekers. ○ Advisor/PI: Jonathan May 	
Ultrascale Systems Research Center, Los Alamos National Laboratory	Remote
<i>Post-Baccalaureate Research Intern</i>	May 2020 – Aug. 2020
<ul style="list-style-type: none"> ○ Collected 750,000 lines of non-secret syslog data using a 20-node virtual compute cluster. ○ Tested several word embedding and clustering algorithms to extract natural language features from supercomputer syslog. ○ Found these natural language features can be used to improve HPC node failure detection and prediction. ○ Publication: Authored workshop paper accepted to Second Workshop on Machine Learning for Computing Systems, Supercomputing 2020. ○ Mentor: Elisabeth (Lissa) Moore 	
Digital Latin Library, University of Oklahoma	Norman, OK
<i>Undergraduate Research Assistant</i>	Jan. 2017 – May 2020
<ul style="list-style-type: none"> ○ Wrote Python scripts to parse HTML documents containing Latin texts, generating 125000 lines of SQL output. ○ Designed scripts to automate XML encoding of Latin prose and poetry, reducing time to encode by a factor of 400. ○ Only undergraduate student accepted to National Endowment for the Humanities Institute on Advanced Topics in Digital Humanities for Summer 2018 at Tufts University. ○ Presentations: Conference talks at Society for Classical Studies 2019, Classical Association of the Middle West and South 2018, National Conference on Undergraduate Research 2018 ○ Mentor: Samuel Huskey 	

TEACHING EXPERIENCE

CSCI 662 Advanced Natural Language Processing, University of Southern California	Los Angeles, CA
<i>Teaching Assistant</i>	Aug. 2025 – Dec. 2025
<ul style="list-style-type: none"> ○ Substantially update all homework assignments to better cover finetuning, attention, transformers, and RAG. ○ Migrate automated grading for all programming assignments from Vocareum to Gradescope. ○ Configure all GradeScope assignments and rubrics. ○ Grade assignments, including programming homework, research reports, and in-class presentations. ○ Manage course logistics for about 30 students, including webpage, Slack channel, and Brightspace LMS. ○ Host weekly office hours and answer student questions on Slack. ○ Professor: Jonathan May 	
CSCI 350 Introduction to Operating Systems, University of Southern California	Los Angeles, CA
<i>Teaching Assistant</i>	May 2025 – Aug. 2025
<ul style="list-style-type: none"> ○ Graded C programming assignments in XV6 teaching operating system. ○ Hosted weekly office hours and answered student questions on Piazza. ○ Professor: Bill Cheng 	

WORK EXPERIENCE

Epic Systems Corporation	Madison, WI
<i>Software Developer Intern</i>	May 2019 – Aug. 2019
<ul style="list-style-type: none"> ○ Prototyped a menstrual cycle tracking feature within MyChart for iOS to allow patients to share cycle data with providers. ○ Designed and implemented server and client-side data structures to save cycle data to patient's medical record. ○ Built four complete and interactive user-interface screens and partial prototypes of several others. ○ Conducted user research, including about 100 internal survey responses and 5 user interviews. 	

PRESENTATIONS

Invited Talks

- Celebrate Track Invited Speaker, *Community-Sourced Bias Metrics for LLMs*, UCLA QWERHacks. Feb. 2025
- Guest Lecturer on *Ethics and Power in NLP*, USC CSCI 662 Advanced NLP Fall 2023, 2024, 2025
- Invited Panelist on *Language Matters: Addressing Bias in LLMs*, IAPP Privacy. Security. Risk. Sep. 2024
- Invited Speaker, John Brooks Slaughter Leadership in Engineering DEI Summit. Feb. 2024
- Invited Panelist on *Bias in Generative AI*, QueerInAI Workshop at NeurIPS. Dec. 2023
- Invited Speaker, *Anti-Queer Bias in LLMs*, USC ISI Natural Language Seminar. Jun. 2022
- Guest Lecturer, *Servius: An Object-Oriented Approach*, Sunoikisis Digital Classics May 2019

Conference Talks and Posters

- Lightning Talk, *Participatory Methods for LLM Bias Benchmarks*, Responsible Data Workshop. Sep. 2025
- Poster Presentation, *Textual Entailment as a Bias Evaluation Metric for LLMs*, USC ShowCAIS. Apr. 2025
- Poster Presentation, *GPT is Not an Annotator*, Southern California NLP Symposium. Nov. 2024
- Roundtable Discussion, *Defining and Disrupting Antisemitism*, Association for Jewish Studies. Dec. 2023
- Poster Presentation, *Anti-Queer Bias in LLMs*, CRA-WP Grad Cohort for Women. Apr. 2023
- Poster Presentation, *Anti-Queer Bias in LLMs*, Southern California NLP Symposium. Nov. 2022
- *Automatically Encoding Digital Editions of Latin Texts*, Society for Classical Studies. Jan. 2019
- *Coding And/As Classical Scholarship*, Classical Association of the Midwest and South. Apr. 2018
- *Coding And/As Classical Scholarship*, National Conference on Undergraduate Research. Apr. 2018

PROFESSIONAL SERVICE

USC ISI Natural Language Seminar

Los Angeles, CA

Coordinator Aug. 2024 – May 2025

- Recruited external speakers and coordinated publicity for 15 talks, average 20 attendees each.

Conference and Workshop Reviewing

- Conferences: EMNLP 2023, ACL 2023, EMNLP 2022 Theme Track
- Workshops: NAACL Student Research Workshop 2025
- Ethics Reviewing: EMNLP 2025

LEADERSHIP AND OUTREACH

USC Center for AI in Society

Los Angeles, CA

Student Leader Aug. 2025 – Present

- Assist with publicizing seminars, mentoring CAIS++ undergraduates, and planning ShowCAIS Symposium.

Queers in Engineering, Science, and Technology (QuEST)

Los Angeles, CA

Graduate Representative Sep. 2020 – Present

- Recruit graduate students to QuEST and advocate for graduate student needs in the organization.
- Conduct general meetings on graduate school/fellowship applications and mentor undergraduate applicants.
- 2022 oSTEM Region F Conference Organizer: Present graduate application workshop; speak on Queer in Academia panel; organize and hoste charity livestream, raising \$950 for LGBTQ+ community organizations in LA

Alpha Sigma Kappa -- Women in Technical Studies

Norman, OK

President Apr. 2019 – Apr. 2020

- Oversaw 6 executive board members and served as chief risk management officer.
- Led chapter's transition to virtual operations due to COVID-19, rapidly establishing new programs to maintain sisterhood and encourage academic excellence during the pandemic.

Vice President Dec. 2017 – Dec. 2018

- Planned and executed nine recruitment events per semester, recruiting 32 new members.
- Mentored 9 officers on 5 committees, resulting in effective chapter social and new member education programs.

Out in Science, Technology, Engineering, and Math

Norman, OK

Secretary Mar. 2017 – Dec. 2019

- Served on founding executive board of first oSTEM chapter in the state of Oklahoma.
- Planned and executed about two medium-scale events per month and managed event promotion and member communication.

Halliburton Women's Welcome

Norman, OK

Program Committee Chair Aug. 2018, Aug. 2019

- Planned and executed a week-long program for 80+ incoming female students in engineering.
- Organized campus resource fair and inclusive professional dress workshop which received 91% positive participant feedback.