Chan Young Park

park@utexas.edu | https://chan0park.github.io

Professional Experience

University of Texas at Austin
Assistant Professor, Information School, School of Computing — Aug 2026 - Present

Microsoft Research
Visiting Researcher, AI Interaction and Learning — Mar 2025 - July 2026

University of Washington
Postdoctoral Researcher, Computer Science & Engineering — Sep 2024 - Mar 2025
· Advisor: Yejin Choi

Education

Carnegie Mellon University
Ph.D., Language Technologies Institute, School of Computer Science — Aug 2018 - Aug 2024
· Advisor: Yulia Tsvetkov

Yonsei University
M.S., Computer Science, School of Integrated Technology — Mar 2016 - May 2018
B.S., Computer Science, School of Integrated Technology — Mar 2013 - Feb 2016
· Advisor: Prof. Songkuk Kim, M.S. GPA: 4.13/4.00, B.S. GPA: 4.09/4.00
· Magna Cum Laude

Awards

K&L Gates Presidential Fellowship	Oct 2023
Asian Deans' Forum 2023 - The Rising Stars Women in Engineering	Aug 2023
ACL Best Paper Award	July 2023
Wikimedia Foundation Research Award of the Year 2023	May 2023
UChicago Rising Stars in Data Science	Nov 2022
Overseas PhD Scholarship, Korea Foundation of Advanced Studies	Aug 2018 - May 2023
Undergraduate Scholarship, Korea Foundation of Advanced Studies	Mar 2015 - Feb 2016
Excellence Award, Ministry of Science of Korea	May 2015
Full-Ride Scholarship, Yonsei University	Mar 2013 - May 2018

Grants

Schmidt Sciences: Reasoning Up the Instruction Ladder: Inference-Time Deliberation for Safer AI
Co-PI (with PI Sachin Kumar (OSU))
$178,319. January 2026-June 2027

Publications

Preprints

Jillian Fisher, Jennifer Neville, Chan Young Park. Response-Aware User Memory Selection for LLM Personalization. In *arXiv*. **arXiv 2026**

Dustin Wright, Sarah Masud, Jared Moore, Srishti Yadav, Maria Antoniak, Peter Ebert Christensen, Chan Young Park, Isabelle Augenstein. Epistemic Diversity and Knowledge Collapse in Large Language Models. In *arXiv*. **arXiv 2026**

Conferences/Journals

Zishuo Zheng, Vidhisha Balachandran, Chan Young Park, Faeze Brahman, Sachin Kumar. Reasoning Up the Instruction Ladder for Controllable Language Models. In *arXiv*. **ACL 2026 (Findings)**

Taylor Sorensen, Benjamin Newman, Jared Moore, Chan Young Park, Jillian Fisher, Niloofar Mireshghallah, Liwei Jiang, Yejin Choi. Spectrum Tuning: Post-Training for Distributional Coverage and In-Context Steerability. In *arXiv*. **ICLR 2026**

Chan Young Park*, Jillian Fisher*, Marius Memmel, Dipika Khullar, Seoho Yun, Abhishek Gupta, Yejin Choi. Making VLMs More Robot-Friendly: Self-Critical Distillation of Low-Level Procedural Reasoning. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. **EMNLP 2025 (Oral)**

Yejin Son*, Minseo Kim*, Sungwoong Kim, Seungju Han, Jian Kim, Dongju Jang, Youngjae Yu*, Chan Young Park*. Subtle Risks, Critical Failures: A Framework for Diagnosing Physical Safety of LLMs for Embodied Decision Making. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. **EMNLP 2025**

Shuyue Stella Li, Melanie Sclar, Hunter Lang, Ansong Ni, Jacqueline He, Puxin Xu, Andrew Cohen, Chan Young Park, Yulia Tsvetkov, Asli Celikyilmaz. PrefPalette: Personalized Preference Modeling with Latent Attributes. In *Second Conference on Language Modeling*. **COLM 2025 (Spotlight)**

Jillian Fisher, Ruth E. Appel, Chan Young Park, Yujin Potter, Liwei Jiang, Taylor Sorensen, Shangbin Feng, Yulia Tsvetkov, Margaret E. Roberts, Jennifer Pan, Dawn Song, Yejin Choi. Political Neutrality in AI is Impossible- But Here is How to Approximate it. In *42nd International Conference on Machine Learning*. **ICML 2025 (Oral)**

Yu Ying Chiu, Liwei Jiang, Bill Yuchen Lin, Chan Young Park, Shuyue Stella Li, Sahithya Ravi, Mehar Bhatia, Maria Antoniak, Yulia Tsvetkov, Vered Shwartz, Yejin Choi. CulturalBench: a Robust, Diverse and Challenging Benchmark on Measuring the (Lack of) Cultural Knowledge of LLMs. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*. **ACL 2025**

Quan Ze Chen, K.J. Kevin Feng, Chan Young Park, Amy X. Zhang. SPICA: Retrieving Scenarios for Pluralistic In-Context Alignment. In *Findings of the Association for Computational Linguistics: ACL 2025*. **ACL 2025 (Findings)**

Chan Young Park*, Sachin Kumar*, Yulia Tsvetkov, Noah A Smith, Hannaneh Hajishirzi. ComPO: Community Preferences for Language Model Personalization. In *Proceedings of the 2025 Conference of the North American Chapter of the Association for Computational Linguistics*. **NAACL 2025**

Chan Young Park*, Shuyue Stella Li*, Hayoung Jung*, Svitlana Volkova, Tanushree Mitra, David Jurgens, Yulia Tsvetkov. ValueScope: Unveiling Implicit Norms and Values via Return Potential Model of Social Interactions. In *Findings of the Association for Computational Linguistics: EMNLP 2024*. **EMNLP 2024 (Findings)**

Farhan Samir, Chan Young Park, Anjalie Fields, Vered Shwartz, Yulia Tsvetkov. Locating Information Gaps and Narrative Inconsistencies Across Languages: A Case Study of LGBT People Portrayals on Wikipedia. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. **EMNLP 2024**

Shangbin Feng, Taylor Sorensen, Yuhan Liu, Jillian Fisher, Chan Young Park, Yejin Choi, Yulia Tsvetkov. Modular Pluralism: Pluralistic Alignment via Multi-LLM Collaboration. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. **EMNLP 2024**

Yuhan Liu, Shangbin Feng, Xiaochuang Han, Vidhisha Balachandran, Chan Young Park, Sachin Kumar, Yulia Tsvetkov. What Constitutes a Faithful Summary? Preserving Author Perspectives in News Summarization. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics.* **NAACL 2024**

Sachin Kumar, Chan Young Park, Yulia Tsvetkov. GEN-Z: Generative Zero-Shot Text Classification with Contextualized Label Descriptions. In *The Twelfth International Conference on Learning Representations.* **ICLR 2024**

Chan Young Park*, Lucille Njoo*, Octavia Stappart, Marvin Thielk, Yi Chu and Yulia Tsvetkov. TalkUp: Paving the Way for Understanding Empowering Language. In *Findings of the Association for Computational Linguistics: EMNLP 2023.* **EMNLP 2023 (Findings)**

Jihyung Moon, Dong-Ho Lee, Hyundong Cho, Woojeong Jin, Chan Young Park, Minwoo Kim, Jonathan May, Jay Pujara, and Sungjoon Park. Analyzing Norm Violations in Live-Stream Chat . In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing.* **EMNLP 2023**

Shangbin Feng, Chan Young Park, Yuhan Liu and Yulia Tsvetkov. From Pretraining Data to Language Models to Downstream Tasks: Tracking the Trails of Political Biases Leading to Unfair NLP Models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics.* **ACL 2023**

- Received ACL Best Paper Award.
- Media Coverage: "AI language models are rife with different political biases" (MIT Tech Review, 2023), "ChatGPT leans liberal, research shows" (The Washington Post, 2023)

Chan Young Park*, Julia Mendelsohn*, Anjalie Field*, Yulia Tsvetkov. Challenges and Opportunities in Information Manipulation Detection: An Examination of Wartime Russian Media. In *Findings of the Association for Computational Linguistics: EMNLP 2022.* **EMNLP 2022 (Findings)**

Chan Young Park*, Anjalie Field*, Antonio Theophilo*, Jamelle Watson-Daniels, and Yulia Tsvetkov. An Analysis of Emotions and the Prominence of Positivity in #BlackLivesMatter Tweets. In *Proceedings of the National Academy of Sciences*, 2022. **PNAS 2022**

- Featured: "Analysis of #BlackLivesMatter social media content points to the power of positivity in online activism and large-scale social movements" (UW Allen School News, 2022), "Researchers Study Tweets To Understand Roles of Optimism and Hope in the Black Lives Matter Movement" (CMU SCS News, 2022)

Anjalie Field, Chan Young Park, Kevin Z. Lin, and Yulia Tsvetkov. Controlled Analyses of Social Biases in Wikipedia Bios. In *Proceedings of the ACM Web Conference*, 2022. **WWW 2022**

- Received Wikimedia Foundation Research Award of the Year 2023. Two papers were chosen among 180+ research papers published on or about Wikimedia projects during 2022.

Chan Young Park, Julia Mendelsohn, Karthik Radhakrishnan, Kinjal Jain, Tushar Kanakagiri, David Jurgens, and Yulia Tsvetkov. Detecting Community Sensitive Norm Violations in Online Conversations. In *Findings of the Association for Computational Linguistics: EMNLP 2021.* **EMNLP 2021 (Findings)**

Chan Young Park*, Jimin Sun*, Hwijeen Ahn*, Yulia Tsvetkov, and David R. Mortensen. Ranking Transfer Languages with Pragmatically-Motivated Features for Multilingual Sentiment Analysis. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics*, 2021. **EACL 2021**

Chan Young Park*, Xinru Yan*, Anjalie Field*, and Yulia Tsvetkov. Multilingual Contextual Affective Analysis of LGBT People Portrayals in Wikipedia. In *Proceedings of International AAAI Conference on Web and Social Media*, 2021. **ICWSM 2021**

- Impact: collaboration with investigative journalists from the Washington Post to collect Chinese social media data and analyze sentiments expressed toward the black population during the pandemic, published as "Video evidence of anti-black discrimination in China over coronavirus fears" (The Washington Post, 2020)

Chan Young Park*, Hwijeen Ahn*, Jimin Sun*, and Jungyun Seo. NLPDove at SemEval-2020 Task 12: Improving Offensive Language Detection with Cross-lingual Transfer. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, 2020. **SemEval 2020**

Chan Young Park and Yulia Tsvetkov. Learning to generate word-and phrase-embeddings for efficient phrase-based neural machine translation. In *Proceedings of the 3rd Workshop on Neural Generation and Translation*, 2019. **WNGT 2019**

Industry Experience

Facebook AI

Research Intern May 2021 - Sep 2021

· Developed models for *personalized sentiment classification*, specifically explored various pre-training objectives and methodologies to train transformer-based personalized language models. Location: Seattle, WA, Team: AI Integrity, Mentor: Hao Ma, Additional Mentors: Madian Khabsa, Simon Wang, Yuxiao Dong

Adobe Research

Research Intern May 2019 - Aug 2019

· Developed *language generation models* for procedural texts such as instructions and recipes. Location: San Jose, CA, Team: NLP, Mentor: Doo Soon Kim

Research Intern Jun 2018 - Aug 2018

· Developed *text segmentation* models. Location: San Jose, CA, Team: NLP, Mentor: Seokhwan Kim

Naver Clova AI

Research Intern Jul 2017 - Nov 2017

· Conducted research on dialog state tracking and developed a tool to visualize dialog response templates. Location: Seoul, Korea, Team: Dialog Systems, Mentor: Kyungduk Kim

Teaching

Guest lecture for Deep Learning (CS 4644 / 7643) Spring 2026

· Georgia Institute of Technology, Gave a lecture on Post-Training of LLMs

Guest lecture for Natural Language Processing (CSE 447) Fall 2024

· University of Washington, Gave lectures on LLM+Reasoning, knowledge distllation, and text summarization

Guest lecture for Natural Language Processing (CSE 447) Spring 2022

· University of Washington, Gave a lecture on text summarization

Guest lecture for Algorithms for NLP Fall 2021

· Carnegie Mellon University, Gave a lecture on text summarization

Lecture for Low-Resource NLP Bootcamp Fall 2020

· Carnegie Mellon University, Led a hands-on learning session for Multilingual NLP

TA for Algorithms for NLP Fall 2019, Spring 2020

· Carnegie Mellon University, Developed assignments and quizzes, delivered lectures and recitations, and mentored student research projects.

TA for Theory of Algorithm Fall 2016, Fall 2017

· Yonsei University, Developed homework assignments

Invited Talks

Modeling Social Influence in Language: Application, Integration, and Measurement

Workshop on Social Influence in Conversations@EMNLP 2024, Nov 2024

Measuring Social Biases Using Language Technologies

University of Arizona, Ethics Colloquium, Oct 2023

Socioculturally Aware Language Technologies

KAIST School of AI, Aug 2023

Socioculturally Aware Language Technologies

Aptima, Aug 2023

Multilingual Contextual Affective Analysis of LGBT People Portrayals in Wikipedia

Wikimedia Research Showcase, Jun 2023

Socioculturally Aware Language Technologies

KAIST School of Computing, Jun 2023

Challenges and Opportunities in Opinion Manipulation Detection

University of Chicago, Nov 2022

Challenges and Opportunities in Opinion Manipulation Detection

Text as Data (TADA) Conference, Oct 2022

Recent Advances in NLP and AI Ethics

Hyundai Vision Conference, Aug 2022

Service

Mentoring

- Farhan Samir, PhD Student, UBC 2023
- Stella Li, PhD Student, UW 2023
- Hayoung Jung, MS Student, UW 2023
- Lucille Njoo, PhD Student, UW 2023
- CMU K-12 Summit on AI for Social Good Symposium 2022
- CMU Graduate Application Support program 2020, 2021
- Hwijeen Ahn, CMU, Visiting MS Student → CMU LTI 2020
- Jimin Sun, CMU, Visiting MS Student 2020
- Xinru Yan, MS student, CMU 2020

Area Charing

- Senior Area Chair: ACL 2026, EMNLP 2026, ARR 2026 (Jan, Mar)
- Area Chair: NeurIPS 2026, ARR 2025 (May, July, Oct), EMNLP 2023

Workshop Organizing

- EMNLP 2024: *Customizable NLP: Progress and Challenges in Customizing NLP for a Domain, Application, Group, or Individual*

Reviewing

- ARR 2024-Present
- ACL 2023-Present
- EMNLP 2020-Present
- CoNLL 2021, 2022
- NAACL SRW 2020, 2021
- NLP4PI Workshop 2021, 2022

Skills

- Computer-related: Python, C++, C, Java, React, Django, AWS, SQL, R, MATLAB
- Natural Language: English, Korean