Hyundong Justin Cho

Research Interests	I design human-centric foundation models. My work explores post-training with custom computational environments and reward structures [P1, R5, R2], personalized alignment methods [R3, R4, R11], and specialized user interfaces for complex real-world evaluations [R6, R8].

Education

University of Southern California, Information Sciences Institute (USC ISI)   09/2020 - Present
Ph.D. in Computer Science, Viterbi School of Engineering
Advisor: Jonathan May

Hong Kong Unversity of Science and Technology (HKUST)   09/2013 - 06/2019
B.Eng. in Computer Science, School of Engineering

Experience

Amazon AGI, Research Intern   05/2024 - 09/2024
Developed Trial-Error-Explain ICL, a tuning-free personalized text generation method [R3].

Amazon Alexa, Research Intern   06/2023 - 09/2023
Post-trained LLMs with RL from human feedback on audio responses for voice-suitable
response generation [R5].

Meta AI, Research Intern   06/2022 - 09/2022
Implemented example-guided question answering for continual learning for dialogue state
tracking (DST) for on-device models [R9].

Meta AI, Research Intern   08/2021 - 03/2022
Developed CheckDST for comprehensive DST evaluation and PrefineDST for pre-finetuning
DST models to improve their robustness to adversarial perturbations [R12].

Stitch Fix Inc., Data Science Intern   06/2020 - 08/2020
Developed a customer review analysis system. [Blog post]

USC ISI, Programmer Analyst   09/2019 - 06/2020
Post-trained dialogue models for grounded response generation using improvised dialogues [R17].

Publications

∗: equal contribution.   Featured work

[R1] **Can Vision Language Models Understand Mimed Actions?**
Hyundong Cho, S. Lin, T. Srinivasan, M. Saxon, D. Kwon, N. Chavez, J. May. 2025
ACL 2025 Findings

[R2] **NewsInterview: Evaluating LLMs' Grounding Gap via Informational Interviews**
M. Lu, S. Kalyan, **Hyundong Cho**, W. Shi, J. May, A. Spangher. 2025
ACL 2025

[R3] **Tuning-Free Personalized Alignment via Trial-Error-Explain In-Context Learning**
Hyundong Cho, K. Sharma, N. Jedema, L. Ribeiro, A. Moschitti, J. May.
NAACL 2025 Findings

[R4] **Aligning Language Models with Demonstrated Feedback**
O. Shaikh, M. Lam, J. Hejna, Y. Shao, **Hyundong Cho**, M. Bernstein, D. Yang. 2024
ICLR 2025

[R5] Speechworthy Instruction-tuned Language Models
Hyundong Cho, N. Jedema, L. Ribeiro, K. Sharma, P. Szekely, A. Moschitti, R. Janssen, J. May
EMNLP 2024

[R6] Can Language Model Moderators Improve the Health of Online Discourse?
Hyundong Cho, S. Liu, T. Shi, D. Jain, B. Rizk, Y. Huang, Z. Lu, N. Wen, J. Gratch, E. Ferrara, J. May
NAACL 2024

[R7] Improving Covert Toxicity Detection by Retrieving and Generating References
D. Lee, **Hyundong Cho**, W. Jin, J. Moon, S. Park, P. Röttger, J. Pujara, R. Lee
NAACL 2024 Workshop on Online Abuse and Harms

[R8] BotEval: Facilitating Interactive Human Evaluation
Hyundong Cho, Thamme Gowda, Y. Huang, Z. Lu, Tianli Tong, J. May
ACL 2024 System Demonstration

[R9] Continual Dialogue State Tracking via Example-Guided Question Answering
Hyundong Cho, A. Madotto, Z. Lin, K. Raghavi Chandu, S. Kottur, J. Xu, J. May, C. Sankar
EMNLP 2023

[R10] Analyzing Norm Violations in Live-Stream Chat
J. Moon, D. Lee, **Hyundong Cho**, W. Jin, C. Park, M. Kim, J. May, Jay Pujara, S. Park
EMNLP 2023

[R11] RECAP: Retrieval-Enhanced Context-Aware Prefix Encoder for Personalized Dialogue Response Generation
S. Liu, **Hyundong Cho**, M. Freedman, X. Ma, J. May
ACL 2023

[R12] Know Thy Strengths: Comprehensive Dialogue State Tracking Diagnostics
Hyundong Cho, C. Sankar, C. Lin, A. Celikyilmaz, S. Shayandeh, J. May, A. Beirami
EMNLP 2022 Findings

[R13] Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality
P. Zhou, **Hyundong Cho**, P. Jandaghi, D. Lee, Y. Lin, J. Pujara, X. Ren
EMNLP 2022

[R14] Probing Commonsense Explanation in Dialogue Response Generation
P. Zhou, P. Jandaghi, **Hyundong Cho**, Y. Lin, J. Pujara, X. Ren
EMNLP 2021 Findings

[R15] Viola: A Topic Agnostic Generate-and-Rank Dialogue System
Hyundong Cho, B. Shbita, K. Shenoy, S. Liu, N. Patel, H. Pindikanti, J. Lee, J. May 2021
Alexa Prize Socialbot Grand Challenge 4 Proceedings

[R16] Agenda Pushing in Email to Thwart Phishing
Hyundong Cho, G. Bartlett, M. Freedman
ACL 2021 Doc2Dial Workshop

[R17] Grounding Conversations with Improvised Dialogues
Hyundong Cho, J. May
ACL 2020

Preprints [P1] Which Questions Improve Learning the Most? Utility Estimation of Questions with LM-based Simulations
D. Lee*, **Hyundong Cho***, J. May, J. Pujara. 2025

[P2] Uncovering Intervention Opportunities for Suicide Prevention with Language Model Assistants
J. Ranjit, **Hyundong Cho**, C. Smerdon, Y. Nam, M. Phung, J. May, J. Blosnich, S. Swayamdipta. 2025

Awards	**Alexa Prize Socialbot Grand Challenge 4 Semi-finalist**, $250,000 grant, Amazon, 2021. **Graduate School Fellowship for Incoming Students**, USC, 2020. **First Class Honors**, HKUST, 2019. **Academic Excellence Recruitment Scholarship**, Full tuition, HKUST, 2013-2019. **Reaching Out Award**, HKSAR Government Scholarship Fund, 2019. **CodeIT Suisse Coding Challenge Winner**, Credit Suisse APAC Region, 2017/9. **Dean's List**, HKUST, 2013, 2015, 2017, 2018, 2019.
Invited Talks	**CSCI 499 Language Models in NLP**: Guest Lecture on RLHF **CSCI 662 Advanced NLP**: Guest Lecture on RLHF **USC ISI NLG Seminar**: Contextualization for Human-AI Interactions **Tonal Talk**: LLMs and RLHF
Teaching Experience	**TA**, CSCI 662 Advanced NLP **TA**, CSCI 699 Probabilistc and Generative Models
Professional Services	**Reviewer**: *ACL, EMNLP, ACL ARR (2021-2025), NeurIPS (2023-2025), ICLR (2025), TMLR (2023-2024) **Lead Workshop Organizer**: What's left to TEACH chatbots Workshop at ICML 2023