Shikib Mehri

✉ 🎓

EXPERIENCE

Amazon (Alexa AI & Amazon General Intelligence)   Sunnyvale, CA
Applied Scientist   *Aug 2022 - Present*

- Scientist working on LLM research and product applications, with a concentration on reward modeling, learning from human feedback, data collection and curation, and model-in-the-loop.
- Lead scientist on **reward modeling**. Overseeing all stages from preference data collection, training methodology, analysis and validation. Trained all RMs for Alexa and AGI. Novel scientific enhancements to data collection and reward model training methodology.
- Co-led data **collection** for Alexa LLM, created a *'Taxonomy of User Goals for LLMs'* and iterated on annotation guidelines. Improved quality and throughput through novel **model-in-the-loop** data collection strategies.
- Independently established a **centralized LLM evaluation framework**, automating response assessment, Elo ratings, and visualizations. This was outside of my job scope, but I developed a functional end-to-end solution (over a weekend) to unblock an organization of 50+ members. My solution was used for 2+ months, and since then, I have transferred ownership of the various components to different engineering teams.
- Core science contributor to **LHF**, including work on iterative **reward-based data augmentation**, RLHF, **critique-and-revision**, knowledge grounding, automatic evaluation.
- Hosted and mentored two PhD interns: Steven Feng (Stanford) worked on CoT and reasoning, and Kaustubh Dhole (Emory) worked on user simulators. Publications in the works.

EDUCATION

Carnegie Mellon University   Pittsburgh, PA
Ph.D. Language Technologies   *Aug 2018 - Jul 2022*

- Thesis: *"Towards Generalization in Dialog through Inductive Biases"* [Link]
- Committee: Maxine Eskenazi, Graham Neubig, Yonatan Bisk, Dilek Hakkani-Tur, Tiancheng Zhao
- TA for Neural Networks for NLP (Graham Neubig) and Multi-Modal Machine Learning (LP Morency).
- Awarded CMU Presidential Fellowship for LTI (2021); secured funding from RASA, Google, Microsoft; co-organized NSF workshop on dialog evaluation.

University of British Columbia   Vancouver, BC
B.Sc. Honours Computer Science   *Sep 2013 - May 2018*

- Served as a TA for various undergraduate CS courses and a graduate-level multimodal machine learning class, teaching tutorials, managing labs, grading assignments, and tutoring students.
- Research assistant on several projects: conversational analysis (IJCNLP paper), video captioning (NeurIPS paper), and bioinformatics.
- Internships at Facebook ('16, '17), Microsoft ('17), and Amazon ('18), focusing on applied ML.

PUBLICATIONS

Preprint 2023	Data-Efficient Alignment of Large Language Models with Human Feedback Through Natural Language *Di Jin, **Shikib Mehri**, Devamanyu Hazarika, Aishwarya Padmakumar, Sungjin Lee, Yang Liu, Mahdi Namazifar*
EMNLP 2023	CESAR: Automatic Induction of Compositional Instructions for Multi-turn Dialogs *Taha Aksu, Devamanyu Hazarika, **Shikib Mehri**, Seokhwan Kim, Dilek Hakkani-Tur, Yang Liu, Mahdi Namazifar*
IWSDS 2023	Understanding the Effectiveness of Very Large Language Models on Dialog Evaluation *Jessica Huynh, Cathy Jiao, Prakhar Gupta, **Shikib Mehri**, Payal Bajaj, Vishrav Chaudhary, Maxine Eskenazi*
EMNLP 2022	Improving Zero and Few-shot Generalization in Dialogue through Instruction Tuning *Prakhar Gupta, Cathy Jiao, Yi-Ting Yeh, **Shikib Mehri**, Maxine Eskenazi, Jeffrey Bigham*

SIGdial 2022	LAD: Language Models as Data for Zero-Shot Dialog **Shikib Mehri**, *Yasemin Altun, Maxine Eskenazi*
SIGdial 2022	The DialPort Tools *Jessica Huynh,* **Shikib Mehri**, *Cathy Jiao, Maxine Eskenazi*
LREC 2022	Interactive Evaluation of Dialog Track at DSTC9 **Shikib Mehri**, *Yulan Feng, Carla Gordon, Seyed Hossein Alavi, David Traum, Maxine Eskenazi*
NSF Report	Report from the NSF Future Directions Workshop on Automatic Evaluation of Dialog: Research Directions and Challenges **Shikib Mehri**, *Jinho Choi, Luis Fernando D'Haro, Jan Deriu, Maxine Eskenazi, et al.*
SIGdial 2021	Schema-Guided Paradigm for Zero-Shot Dialog **Shikib Mehri**, *Maxine Eskenazi*
SIGdial 2021	GenSF: Simultaneous Adaptation of Generative Pre-trained Models and Slot Filling **Shikib Mehri**, *Maxine Eskenazi*
Preprint 2021	A Comprehensive Assessment of Dialog Evaluation Metrics *Yi-Ting Yeh, Maxine Eskenazi,* **Shikib Mehri**
NAACL 2021	Example-Driven Intent Prediction with Observers **Shikib Mehri**, *Mihail Eric, Dilek Hakkani-Tur*
AAAI 2022	Overview of the Ninth Dialog System Technology Challenge: DSTC9 *Many authors*
Preprint 2020	DialoGLUE: A Natural Language Understanding Benchmark for Task-Oriented Dialogue **Shikib Mehri**, *Mihail Eric, Dilek Hakkani-Tur*
Preprint 2020	STAR: A Schema-Guided Dialog Dataset for Transfer Learning *Johannes EM Mosig,* **Shikib Mehri**, *Thomas Kober*
SIGdial 2020	Unsupervised Evaluation of Interactive Dialog with DialoGPT **Shikib Mehri**, *Maxine Eskenazi*
ACL 2020	USR: An Unsupervised and Reference Free Evaluation Metric for Dialog Generation **Shikib Mehri**, *Maxine Eskenazi*
ACL 2020	None of the Above: Measure Uncertainty in Dialog Response Retrieval *Yulan Feng,* **Shikib Mehri**, *Maxine Eskenazi, Tiancheng Zhao*
EMNLP 2019	Multi-Granularity Representations of Dialog **Shikib Mehri**, *Maxine Eskenazi*
SIGdial 2019	Structured Fusion Networks for Dialog [**Best Paper Award**] **Shikib Mehri**, *Tejas Srinivasan, Maxine Eskenazi*
SIGdial 2019	Investigating Evaluation of Open-Domain Dialogue Systems With Human Generated Multiple References *Prakhar Gupta,* **Shikib Mehri**, *Tiancheng Zhao, Amy Pavel, Maxine Eskenazi, Jeffrey P Bigham*
DiGO 2019	CMU GetGoing: An Understandable and Memorable Dialog System for Seniors **Shikib Mehri**, *Alan W Black, Maxine Eskenazi*
ACL 2019	Pretraining Methods for Dialog Context Representation Learning **Shikib Mehri**, *Evgeniia Razumovskaia, Tiancheng Zhao, Maxine Eskenazi*

NeurIPS 2018   Middle-Out Decoding
Shikib Mehri, *Leonid Sigal*

IJCNLP 2017   Chat Disentanglement: Identifying Semantic Reply Relationships with Random Forests and Re-
current Neural Networks
Shikib Mehri, *Giuseppe Carenini*