
Alexa Prize Socialbot Grand Challenge Year IV

Dilek Hakkani-Tür

Amazon Alexa AI
hakkanit@amazon.com

We are completing the 4th year of the Alexa Prize Socialbot Grand Challenge that aims to advance human-computer interaction through conversations. 9 university teams participated in this year's challenge to create socialbots that can converse coherently and engagingly with humans on a range of current events and popular topics such as entertainment, sports, politics, technology, and fashion. The supported teams include returning ones: Stanford University with Chirpy Cardinal, University of California, Santa Cruz with Athena, Czech Technical University with Alquist, Emory University with Emora, and Moscow Institute of Physics and Technology with Dream; as well as new participants: University of Texas at Dallas with Caspr, The University at Buffalo with Proto, Universidad Politécnica de Madrid with Genuine², and University of Southern California with Viola.

The selected teams were announced in November 2020 and received a research grant, Alexa-enabled devices, free Amazon Web Services (AWS) to support their development efforts, and access to other tools (such as the Conversational Bot, i.e., Cobot, tools as described later in the proceedings), data sources, and support from Alexa engineering and science teams. In addition to the support for accessing Amazon's computing infrastructure and tools, Alexa Prize Challenges have been providing a unique opportunity for university teams to connect their systems with millions of real users for spoken interactions. While conversational systems that enable natural language interactions with machines has been an attractive research area since the early days of computing, automatically evaluating open domain conversational systems that is critical for advancing the quality of these systems still remains an open question [Yeh et al., 2021]. Real user ratings coupled with these conversations provide university teams a large-scale experimentation framework, accelerating the advances in open domain conversational response generation systems and socialbots. Millions of Alexa customers engaged in interactions with one of the participating socialbots throughout the challenge, contributing to the advancement of artificial intelligence in the challenging area of open-domain social conversations.

This year, the semifinals period took place between May 4 and June 25, 2021. During the semifinals, users interacting with Alexa were connected with the teams' socialbots by saying "*Alexa, let's chat*" to any Alexa-enabled device. Competing socialbots were randomly invoked in response to this utterance. At each turn, user's spoken utterances were transcribed using automatic speech recognition (ASR) and these utterances along with high-level metadata (such as, confidence scores) were provided to the socialbots, which in return produced textual responses (possibly with Amazon speech synthesis markup language (SSML) markings to format the prosody of these responses). These responses were then converted to spoken utterances by text-to-speech synthesis. At the end of the

interaction, users were prompted to provide a rating from 1 to 5 stars (5 being the highest) on how they felt about speaking with that socialbot. After the rating, the users were also asked if they would like to leave a verbal feedback to the university team that built the socialbot they just interacted with. The ASR output of user feedback is also provided to the teams along with the rating for the conversation.

The top 3 teams with the highest average user ratings during the semifinals automatically advanced to finals (Emory, Stanford and University of California, Santa Cruz), and Amazon also selected 2 teams (Czech Technical University and The University at Buffalo) as wildcard finalist teams based on the ratings from Alexa customers, depth and breadth of topics covered, and appropriateness and accuracy of their responses.

Throughout the challenge, the Amazon Alexa Prize team has provided a suite of tools and models to the participants including entity recognition trained to be robust to speech recognition errors [Shang et al., 2021], improved dialogue act and topic detection, sensitive content detection, conversation evaluator models, and finally neural response generators trained on the Topical Chat dataset [Gopalakrishnan et al., 2018] and other similar datasets, as well as response generators based on novel approaches to generate more appropriate and engaging responses through policy-driven response generation [Hedayatnia, 2020] as back-off response generators. Alexa science teams also invested in several new areas to better understand human-socialbot conversations and help university teams to build improved socialbots. These include entity resolution in open domain conversations [Shang et al., 2021], zero-shot controls for neural response generation [Hazarika et al., 2021], commonsense reasoning in social conversations [Chang et al., 2021 and Zhou et al., 2021], use of speech sentiment to estimate customer satisfaction [Kim et al., 2020], how personality effects user ratings [Papangelis et al., 2021], pre-training of warped language models to correct speech recognition errors [Namazifar et al., 2021], and knowledge selection for open-domain conversations [Eric et al., 2021].

The first article in the proceedings from the Alexa Prize team presents these advances, in order to support the university teams. These are in addition to the scientific contributions from the previous challenges that have been summarized in [Ram et al., 2018], [Khatri et al., 2020] and [Gabriel et al., 2020]. In the following chapters, the Alexa Prize Challenge proceedings are organized to include overviews and contributions of the participating university systems. Many thanks to the Amazon Alexa Prize team, participating university teams, and Alexa customers for making Amazon Alexa Prize Challenge established and influential, and congratulations to the finalists!

References

[Chang et al., 2021] Ting-Yun Chang, Yang Liu, Karthik Gopalakrishnan, Behnam Hedayatnia, Pei Zhou, Dilek Hakkani-Tur, “Go Beyond Plain Fine-tuning: Improving Pretrained Models for Social Commonsense”, to appear, SLT 2021.

[Eric et al., 2021] Mihail Eric, Nicole Chartier, Behnam Hedayatnia, Karthik Gopalakrishnan, Pankaj Rajan, Yang Liu, Dilek Hakkani-Tur. A New Paradigm for Knowledge Selection in Open-Domain Dialogue. In submission, 2021.

[Gabriel et al., 2020] Raefer Gabriel, Yang Liu, Anna Gottardi, Mihail Eric, Anju Khatri, Anjali Chadha, Qinlang Chen, Behnam Hedayatnia, Pankaj Rajan, Ali Binici, Shui Hu,

Karthik Gopalakrishnan, Seokhwan Kim, Lauren Stubel, Kate Bland, Arindam Mandal, Dilek Hakkani-Tür. Further Advances in Open Domain Dialog Systems in the Third Alexa Prize Socialbot Grand Challenge. Alexa Prize Socialbot Grand Challenge 3 Proceedings, 2020.

[Gopalakrishnan et al., 2018] Karthik Gopalakrishnan, Behnam Hedayatnia, Qinlang Chen, Anna Gottardi, Sanjeev Kwatra, Anu Venkatesh, Raefer Gabriel, Dilek Hakkani-Tür, AI AA. Topical-Chat: Towards Knowledge-Grounded Open-Domain Conversations. In INTERSPEECH 2019 (pp. 1891-1895).

[Hazarika et al., 2021] Devamanyu Hazarika, Mahdi Namazifar, Dilek Hakkani-Tur. Zero-Shot Controlled Generation with Encoder-Decoder Transformers. In submission. arXiv preprint arXiv:2106.06411. 11 June 2021.

[Hedayatnia et al., 2020] Behnam Hedayatnia, Seokhwan Kim, Yang Liu, Karthik Gopalakrishnan, Mihail Eric, Dilek Hakkani-Tur D. Policy-Driven Neural Response Generation for Knowledge-Grounded Dialogue Systems. arXiv preprint arXiv:2005.12529. 26 May 2020 and INLG, 2020.

[Kim et al., 2020] Yelin Kim, Joshua Levy, Yang Liu. Speech Sentiment and Customer Satisfaction Estimation in Socialbot Conversations. Interspeech 2020.

[Khatri et al., 2020] Chandra Khatri, Behnam Hedayatnia, Anu Venkatesh, Jeff Nunn, Yi Pan, Qing Liu, Han Song, Anna Gottardi, Sanjeev Kwatra, Sanju Pancholi, M. Cheng, Qinglang Chen, Lauren Stubel, Karthik Gopalakrishnan, Kate Bland, Raefer Gabriel, Arindam Mandal, Dilek Hakkani-Tur, G. Hwang G, Nate Michel, Eric King, Rohit Prasad. Advancing the state of the art in open domain dialog systems through the Alexa prize. arXiv preprint arXiv:1812.10757. Alexa Prize Proceedings, 2020.

[Namazifar et al., 2021] Mahdi Namazifar, John Malik, Li Erran Li, Gokhan Tur, Dilek Hakkani-Tur. Correcting Automated and Manual Speech Transcription Errors using Warped Language Models. Interspeech 2021.

[Papangelis et al., 2021] Alexandros Papangelis, Nicole Chartier, Pankaj Rajan, Julia Hirschberg, Dilek Hakkani-Tur, Understanding How People Rate Their Conversations. In submission, 2021.

[Ram et al., 2018] Ashwin Ram, Rohit Prasad, Chandra Khatri, Anu Venkatesh, Raefer Gabriel, Qing Liu, Jeff Nunn, Behnam Hedayatnia, M. Cheng, A. Nagar, Eric King, Kate Bland, A. Wartick, Yi Pan, H. Song, S. Jayadevan, G. Hwang, A. Pettigrew. Conversational AI: The science behind the Alexa prize. arXiv preprint arXiv:1801.03604. 2018.

[Shang et al., 2021] Mingyue Shang, Tong Wang, Mihail Eric, Jiangning Chen, Jiyang Wang, Matthew Welch, Tiantong Deng, Akshay Grewal, Han Wang, Yue Liu, Yang Liu and Dilek Hakkani-Tur. Entity Resolution in Open-domain Conversations. NAACL 2021. Industry track.

[Yeh et al., 2021] Yi-Ting Yeh, Maxine Eskenazi and Shikib Mehri. A Comprehensive Assessment of Dialog Evaluation Metrics. arXiv preprint, arXiv: 2106.03706, 7 Jun 2021.

[Zhou et al., 2021] Pei Zhou, Karthik Gopalakrishnan, Behnam Hedayatnia, Seokhwan Kim, Jay Pujara, Xiang Ren, Yang Liu and Dilek Hakkani-Tur. Commonsense-Focused Dialogues for Response Generation: An Empirical Study. SIGDial, 2021.