

ISCA INTERSPEECH-2022

January 26, 2022

Title:

APOLLO Fearless Steps: A Community Resource for Massive Naturalistic Communications

Introduction:

We are pleased to announce the **APOLLO Fearless Steps Special Session** to be held at INTERSPEECH-2022. The focus of this Special Session is to provide a forum for researchers working on the massive naturalistic audio collection stemming from the NASA Apollo Missions. UTDallas-CRSS under NSF support has led the Fearless Steps Initiative, a continued effort spanning eight years has resulted in the digitization, and recovery of over 50,000 hours of original analog audio data, as well as the development of algorithms to extract meaningful information from this naturalistic data resource, including an initial release of pipeline diarization meta-data for all 30 channels of APOLLO-11 and APOLLO-13 Missions. More than 500 sites world wide have accessed the initial data. A current NSF Community Resource project is continuing this effort to recover the remaining Apollo missions (A7-A17; estimated to be 150,000hrs of data) in addition to motivating collaborative speech and language technology research through the Fearless Steps Challenge series.



We invite all researchers to submit papers to this special session which use Fearless Steps speech data as a data set for their research. In addition, this special session is also connected with the next phase of the Fearless Steps Challenge (FS-4) which starts Feb. 1, 2022, and will remain open until Sept. 15, 2022. Any team worldwide can participate and submit their output scoring summaries with a leaderboard being updated regularly (see Fearless Steps FS-4 Challenge logistics below). If your organization, university, team wishes to submit a specific solution on FS-4 for Interspeech-2022, we welcome you to do this as well and submit your paper to this Special Session. A major goal of this Special Session is to allow for researchers to discuss and provide feedback on how they have used this public data resource, and how ongoing efforts can make this data more useful for the speech technology/psychology-team research/history-STEM education communities.

THE FEARLESS STEPS CHALLENGE SERIES: "The FEARLESS STEPS Challenge: (FSC-P1)" held during [ISCA INTERSPEECH-2019](#) encouraged the development of core unsupervised/semi-

supervised speech and language systems for single-channel data with low resource availability. This was followed with [ISCA INTERSPEECH-2020](#) which held the Special Session for [FEARLESS STEPS Challenge \(FSC-P2\)](#), which focused on developing supervised learning strategies for the 100 hour Challenge Corpus. 15 hours of additional data from a previously unseen Apollo-11 channel and audio from five Apollo-13 channels were added to the core challenge corpus as a part of the [Fearless Steps Challenge Phase-03](#), which tested the generalizability criterion for the SLT community using the 5 Challenge Tasks.

The [Fearless Steps Challenge Phase-04](#) held by UTDallas-CRSS in collaboration with National Institute of Standards and Technology (**NIST**) and Linguistic Data Consortium (**LDC**), extends the corpus and system advancements made in the previous challenges by including 10 additional hours of manually transcribed and previously unseen **Apollo-8** mission data. **Additional tasks of Speaker Verification (SV)** and **Topic Detection** are also being available to Fearless Steps Challenge participants to promote high level knowledge extraction from such massively naturalistic conversational audio.

TIMELINE: Challenge Start Date (Train/Dev Data Release): February 1st, 2022
Evaluation Data Release: March 7th, 2022
Evaluation Portal Open (Data Release): March 10th, 2022
INTERSPEECH-2022 Paper submission deadline: March 28th, 2022
Open Challenge Duration: from February 1st, 2022, to September 15th, 2022

Challenge Tasks in Phase-4 (FS#4):

1. Speech Activity Detection (SAD)
2. Speaker Recognition:
 - 2a. Track 1: Speaker Identification (SID)
 - 2b. Track 2: Speaker Verification (SV)
3. Speaker Diarization (SD):
 - 3a. Track 1: Diarization using reference SAD
 - 3b. Track 2: Diarization using system SAD
4. Automatic Speech Recognition (ASR):
 - 4a. Track 1: ASR using reference Diarization
 - 4b. Track 2: Continuous stream ASR
5. Topic Identification:
 - 5a. Track 1: Topic Detection
 - 5b. Track 2: Topic Segmentation

For Challenge Registration, please contact:
FearlessSteps@utdallas.edu

Website Link: <https://fearless-steps.github.io/ChallengePhase4/>

FEARLESS STEPS CHALLENGE

NSF

FEARLESS engineering

UTDALLAS CRSS RSTL NIST LDC NASA

PRESENTED BY: CENTER FOR ROBUST SPEECH SYSTEMS (CRSS) THE UNIVERSITY OF TEXAS AT DALLAS (UTDALLAS)

IN COLLABORATION WITH: LINGUISTIC DATA CONSORTIUM (LDC) AND NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST)

An Audio Corpus of Apollo Missions 8, 11, & 13

TO BE RELEASED FOR: INTERSPEECH 2022

25,000 HOUR NATURALISTIC MULTI-CHANNEL CORPUS

NASA's Apollo program stands as one of mankind's greatest achievements in the 20th century. The CRSS Lab successfully digitized the mission audio tapes, and are now making the data publicly available, intended to advance the Speech and Language Research Community

About the Corpus:
Full Corpus (25k hours): Mission Control, Air-To-Ground, all Back-Room communications in 30 synch. channels for the NASA Apollo-8, 11, & 13 Missions

Challenge Corpus: 125 Hours
Challenge Tasks: 6 (6 sub-tasks)

Challenge Corpus and supporting technologies will be Open Source, Freely Distributed

The Challenge Tasks:

1. Speech Activity Detection
2. Speaker Diarization
3. Speaker Identification
4. Speaker Verification
5. Automatic Speech Recognition
6. Topic Identification

For inquiries regarding the Corpus Delivery and Challenge Tasks, Please Contact Us at: FearlessSteps@utdallas.edu

Organizers

John H.L. Hansen (*john.hansen@utdallas.edu*) is Associate Dean for Research, Professor of Electrical and Computer Engineering, Director of Center for Robust Speech Systems (CRSS) at the Univ. of Texas at Dallas (USA), and serves as PI for the NSF funded projects for analysis, diarization, and linking of audio content for the Apollo program. His research interests span the areas of machine learning for analysis and modeling of speech and speaker traits, diarization for naturalistic data, robust speech recognition, speaker/language/dialect ID in diverse environmental conditions, and general problems in man-machine interaction. He has published 750 papers in the field; is an IEEE Fellow, ISCA Fellow, and currently serves as ISCA President.

Christopher Ceiri (*ccieri@ldc.upenn.edu*) is has been Executive Director of the Linguistic Data Consortium since January 1998. He holds BA, MA and PhD in Linguistics from the University of Pennsylvania where he focused on sociolinguistics, phonology, and educational linguistics. He oversees all aspects of LDC work including research, data collection, annotation, archiving distribution, and outreach as well as the technical infrastructure that enables that work.

James Horan (*james.horan@nist.gov*) is a Technical Manager at NIST, working as Technical Lead on the NIST Speaker and Language Recognition Evaluations. Heads the Multimodal Information Group at NIST, over, and has been responsible for overseeing OpenSAT challenges and DARPA evaluations held annually by NIST.

Aditya Joglekar (*aditya.joglekar@utdallas.edu*) is currently pursuing a Ph.D. at the Univ. of Texas at Dallas (USA). He led the efforts to host the Inaugural Fearless Steps Challenge, held in 2019. His research interests include speaker diarization, speaker identification, speech synthesis, conversational analysis, and multichannel speech processing.

Midia Yousefi (*midia.yousefi@utdallas.edu*) is a post-doctoral researcher and staff member at the Center for Robust Speech Systems (CRSS) at Univ. of Texas at Dallas (USA). She completed her Ph.D. under the guidance of Dr. John H.L. Hansen at CRSS, University of Texas at Dallas. Her research interests include overlap speech detection, speech separation, automatic speech recognition, and sequence to sequence modeling.

Meena Chandra Shekar (*meena.chandrashekar@utdallas.edu*) is currently pursuing a Ph.D. at the Univ. of Texas at Dallas (USA). Her research interests include Speaker tracking and analysis in a naturalistic environment, speaker identification, speaker verification and speaker diarization.