

FEARLESS STEPS: ADVANCEMENTS IN SPEECH TECHNOLOGY AND CORPUS DEVELOPMENT FOR NATURALISTIC AUDIO

A. Joglekar, C. Belitz, M.C. Shekhar, S.-J. Chen, J.H.L. Hansen

{Aditya.Joglekar, Chelzy.Belitz, Meena.ChandraShekhar, SzuJui.Chen, John.Hansen}@utdallas.edu
Center for Robust Speech Systems (CRSS), The University of Texas at Dallas (UTD), Richardson, Texas 75080.

INTRODUCTION: CRSS-UTDallas initiated and oversaw the efforts to recover APOLLO mission communications by re-engineering the NASA SoundScriber playback system, and digitizing 30-channel analog audio tapes – with the entire Apollo-11, Apollo-13, and Gemini-8 missions during 2011-17 [1,6]. This vast data resource was made publicly available along with supplemental speech & language technologies meta-data based on CRSS pipeline diarization transcripts and conversational speaker time-stamps for Apollo team at NASA Mission Control Center, [2,4]. In 2021, renewed efforts over the past year have resulted in the digitization of an additional +50,000hrs of audio from Apollo 7,8,9,10,12 missions, and remaining A-13 tapes. Cumulative digitization efforts have enabled the development of the largest publicly available speech data resource with unprompted, real conversations recorded in naturalistic environments. Deployment of this massive corpus has inspired multiple collaborative initiatives such as Web resources ExploreApollo (<https://app.exploreapollo.org>) LanguageARC (<https://languagearc.com/projects/21>) [3]. ExploreApollo.org serves as the visualization and play-back tool, and LanguageARC the crowd source subject content tagging resource developed by UG/Grad. Students, intended as an educational resource for k-12 students, and STEM/Apollo enthusiasts. Significant algorithmic advancements have included advanced deep learning models that are now able to improve automatic transcript generation quality, and even extract high level knowledge such as ID labels of topics being spoken across different mission stages. Efficient transcript generation and topic extraction tools for this naturalistic audio have wide applications including content archival and retrieval, speaker indexing, education, group dynamics and team cohesion analysis. Some of these applications have been deployed in our online portals to provide a more immersive experience for students and researchers. Continued worldwide outreach in the form of the Fearless Steps Challenges has proven successful with the most recent Phase-4 of the Challenge series. This challenge has motivated research in low level tasks such as speaker diarization and high level tasks like topic identification.

IMPACT: Distribution and visualization of the Apollo audio corpus through the above mentioned online portals and Fearless Steps Challenges have produced significant impact as a STEM education resource for K-12 students as well as a SLT development resource with real-world applications for research organizations globally. The speech technologies developed by CRSS-UTDallas using the Fearless Steps Apollo corpus have improved previous benchmarks on multiple tasks [1, 5]. The continued initiative will extend the current digitization efforts to include over 150,000 hours of audio recorded during all Apollo missions.

ILLUSTRATION: We will demonstrate [WebExploreApollo](#) and [LanguageARC](#) online portals with newly digitized audio playback in addition to improved SLT baseline systems, the results from ASR and Topic Identification systems which will include research performed on the corpus conversational. Performance analysis visualizations will also be illustrated. We will also display results from the past challenges and their state-of-the-art system improvements.

REFERENCES: *[1]* A. Joglekar, S.O. Sadjadi, M. Chandra-Shekar, C. Cieri, J.H.L. Hansen, “Fearless Steps Challenge Phase-3 (FSC P3): Advancing SLT for Unseen Channel and Mission Data across NASA Apollo Audio,” ISCA INTERSPEECH-2021, pp. 986-990, Sept. 2021; *[2]* A. Joglekar, J.H.L. Hansen, M. Chandra-Shekar, A. Sangwan, “FEARLESS STEPS Challenge (FS-2): Supervised Learning with Massive Naturalistic Apollo Data,” ISCA INTERSPEECH-2020, pp. 2617-2621, Oct. 2020; *[3]* J.H.L. Hansen, A. Joglekar, M. Chandra Shekhar, V. Kothapally, C. Yu, L. Kaushik, A. Sangwan “The 2019 Inaugural Fearless Steps Challenge: A Giant Leap for Naturalistic Audio,” ISCA INTERSPEECH-2019, pp. 1851-1855, Sept. 2019; *[4]* J.H.L. Hansen, A. Sangwan, A. Joglekar, A.E. Bulut, C. Yu, L. Kaushik, “Fearless Steps: Apollo-11 Corpus Advancements for speech technologies from Earth to the Moon,” ISCA INTERSPEECH-2018, pp. 2758-2762, Sept. 2018; *[5]* S.-J. Chen, W. Xia, J.H.L. Hansen, “Scenario Aware Speech Recognition: Advancements for Apollo Fearless Steps & Chime-4 Corpora,” IEEE ASRU-2021, Dec. 2021; *[6]* C. Yu, J.H.L. Hansen, D. Oard, “‘Houston, We have a solution’: A Case Study of the Analysis of Astronaut Speech during NASA Apollo 11 for Long-Term Speaker Modeling,” ISCA Interspeech-2014, Sept. 2014.