Building a Standalone AI Assistant for Deep Space Exploration

Lima Agnel Tony

Scientist at Robotics and Artificial Intelligence Group (RAIG), ISRO Inertial Systems Unit (IISU), India

Harikumar Ganesan

Scientist at Robotics and Artificial Intelligence Group (RAIG), ISRO Inertial Systems Unit (IISU), India

Sangeetha G. R.

Scientist at Robotics and Artificial Intelligence Group (RAIG), ISRO Inertial Systems Unit (IISU), India

Amritha A.

Scientist at Robotics and Artificial Intelligence Group (RAIG), ISRO Inertial Systems Unit (IISU), India

Durairaj R.

Scientist and Group Head, Robotics and Artificial Intelligence Group (RAIG), Inertial Sensors Entity (ISE), ISRO Inertial Systems Unit (IISU), India

Joji J. Chaman

Outstanding Scientist and Deputy Director, Inertial Sensors Entity (ISE), ISRO Inertial Systems Unit (IISU), India

Padmakumar E. S.

Distinguished Scientist and Director, ISRO Inertial Systems Unit (IISU), India

Abstract

This paper presents the initial experiences from the design and implementation of an AI visual and voice bot for astronaut assistance in the upcoming Indian Human Space missions. The speech pipeline of the voice bot has speech recognition, text to speech conversion, and an on-device Large Language Model (LLM) based chatbot module. This bot is equipped with a display to render an animated

doi: 10.18278/sesa.5.2.7

face and play media contents. The voice assistant is specifically designed for interaction in Indic-English and Hindi languages. The voice bot along with its display is packaged to occupy the volume of a football. The high-performance processor used is up-screened for use in space, and is size weight power and cost (swap-c) efficient compared to conventional solutions. The capability to function offline without cloud connectivity and ability to interact verbally and visually with the user makes this AI assistant unique in comparison to its contemporaries.

Keywords: AI voice assistant, LLM, ASR, TTS, Deep space missions

Construcción de un asistente de IA independiente para la exploración del espacio profundo

Resumen

Este artículo presenta las experiencias iniciales del diseño e implementación de un bot visual y de voz con IA para la asistencia a astronautas en las próximas misiones espaciales tripuladas de la India. El flujo de voz del bot incluye reconocimiento de voz, conversión de texto a voz y un módulo de chatbot integrado basado en el Modelo de Lenguaje Largo (LLM). Este bot está equipado con una pantalla para mostrar una cara animada y reproducir contenido multimedia. El asistente de voz está diseñado específicamente para la interacción en los idiomas índico-inglés e hindi. El bot, junto con su pantalla, tiene un tamaño similar al de un balón de fútbol. El procesador de alto rendimiento utilizado está optimizado para su uso en el espacio y es eficiente en tamaño, peso, potencia y coste (swap-c) en comparación con las soluciones convencionales. La capacidad de funcionar sin conexión a la nube y de interactuar verbal y visualmente con el usuario hacen que este asistente de IA sea único en comparación con sus predecesores.

Palabras clave: Asistente de voz con IA, LLM, ASR, TTS, Misiones en el espacio profundo

为深空探索构建独立的人工智能助手

摘要

本文介绍了在即将到来的印度载人航天任务中为宇航员提供协助的人工智能(AI)视觉语音机器人的设计和实施的初步经验。语音机器人的语音管道具有语音识别、文本到语音转

换、以及一个基于设备的大型语言模型(LLM)的聊天机器人模块。该机器人配备了一个显示器，能呈现动画脸部并播放媒体内容。语音助手专为印度语-英语和印地语交互而设计。语音机器人及其显示器的包装体积与一个足球相当。所使用的高性能处理器经过升级，可用于太空，与传统解决方案相比，其尺寸、重量、功率和成本(swap-c)效率更高。无需云连接即可离线运行的能力以及与用户进行口头和视觉交互的能力使这款AI助手与同时代产品相比独一无二。

关键词：AI语音助手，大型语言模型，ASR，TTS，深空任务

A. Introduction

During long duration space missions, astronauts experience mental stress and psychological imbalances primarily due to monotony of routine and social isolation (Shashkova, 2022). Time is at a premium in space, and it is important for humans in space to be productive and minimally affected by the environment. Astronaut assistants, including humanoids and flying robots (Ambrose, 2000; Fredrickson, 2004; Schroder, 2018; Salah, 2023; Mitani, 2019) have been deployed onboard the International Space Station (ISS) to provide technical support like video recording of experiments and carrying out efficient routine on-board operations. Recently, increasing efforts are being made to cater to leisure and astronaut engagement too through application of AI and cloud technologies (Bualat, 2018; Yuste, 2022). There is also a massive drive towards bringing intelligence into edge devices by both the industry and research community (Bualat, 2018; Yuste, 2022). In this context, the development of power optimal, locally operating, state-of-the-art voice assistant for astronauts is of great importance.

The idea of conversational assistants has been around for few years now (Klopfenstein, 2017; Landay, 2019; Kulkarni, 2019). In the terrestrial arena, voice assistants like Alexa, Siri, Cortana, and Google voice assistant have already pervaded the daily lives of millions of people in assisting with domain specific tasks (Tulshan, 2019). In terms of system architecture, they incorporate speech recognition, chatbot and speech synthesis technologies that largely rely on internet connectivity to cloud servers. The latest trend is towards enhancing the capability of these systems by leveraging the power of LLMs (Kumar, 2023; Dong, 2023; Ross, 2023).

Literature reports quite a few studies on voice assistants for space applications. Conversational assistants capable of supporting astronauts in configuring extra-terrestrial habitat are discussed in (Goncalves et al., 2021). The assistant can answer queries on experiments, general questions, setting configurations of habitat, etc.

Managing space stations using voice is another direct use case of conversational assistants, where an imaginary personality is developed to promote better task management experience (Vega 2022). SCARLET (Mihaylov, 2022) and CLARK (Henriquez, 2023) are interesting conversational assistants developed by European Space Agency (ESA), that learn mission related lessons from experts and provide this information with context to other working personnel. Conversational assistants have also been developed for satellite operations (Gonccalves, 2018), where the objective of the assistant is to answer queries related to satellite building and functioning. Most of these papers focus more on the application domain than on the system realization aspects.

This paper describes the software and hardware architecture of an AI voice bot which is developed as an astronaut assistant for Indian Space Research Organization's upcoming Gaganyaan (Sundararajan, 2021) and BAS (Ilavazhagi, 2023) missions targeting capability building in sustained habitation of humans in space. In space, astronauts are involved in various experiments and operation of different devices and machines, among many others. Earlier astronauts used to carry operation manuals, etc. in printed format but now are replaced by laptops and tablets (Hoppenbrouwers, 2017). The AI bot extends this digital convenience by adding the capability to retrieve information stored in its local database and connected on-board systems through voice requests. Using its generative AI capabilities, the voice assistant can also act as a warm companion helping astronauts alleviate boredom and loneliness in space. In addition, the graphical display of the voice bot showcases an animated face while in interaction to establish an emotional connection with the user. The system caters to these diverse functionalities by suitably adapting its internal operating modes according to user requirements. An AI space voice bot deployed on a ruggedized edge device and packaged to be fully portable with the state-of-the-art language processing, operating locally and in two languages, to the best of our knowledge, is not reported elsewhere in literature.

B. AI Voice Bot Software

The software system is developed as ROS (Robot Operating System) nodes over Ubuntu 20.04 LTS operating system. The sub-components are described in the following section.

a. **Speech Module**: The software layout of AI voice bot speech module is shown in Fig. 1.

    i. Speech Input Processing: This module converts the audio input into its text equivalent. The input signal goes through Voice Activity Detection (VAD) and keyword spotter followed by Automatic Speech Recognition (ASR). The ASR used in this work is a lightweight

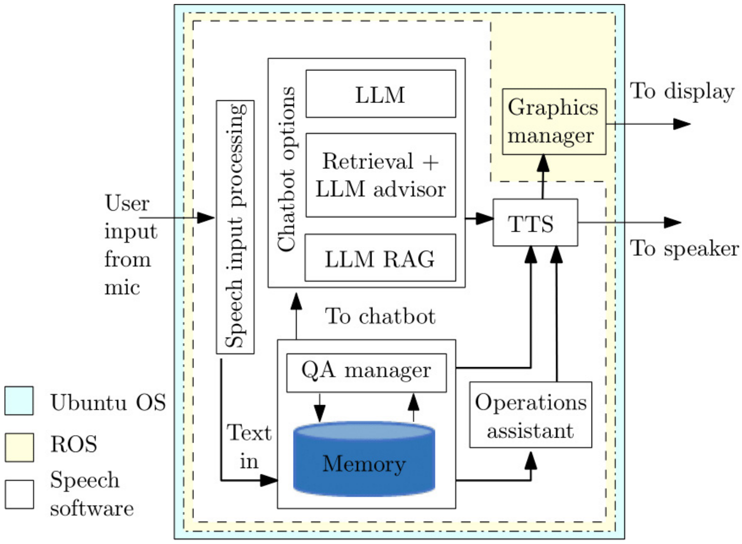

Fig. 1 Speech module block diagram

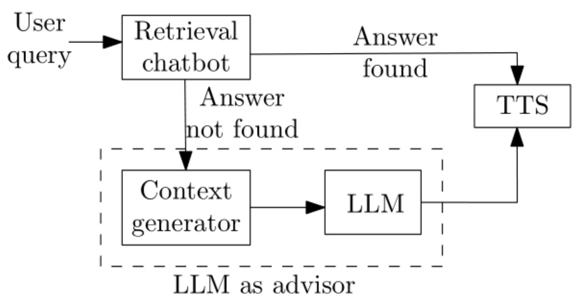

Fig. 2 Retrieval chatbot with LLM advisor

Kaldi (Povey, 2011) based Time Delay Neural Network (TDNN) model. ASR can process speech input in Hindi and English languages and has been trained on 50 hours of data with 40,000 unique words including common words relevant to the space domain.

ii. QA manager: The Question Answer manager discerns if the user input should be sent to the chatbot or operations assistant or be dealt directly. In case the user input is to display figures or to play media content, it is considered as a command and is handed over to operations assistant. Generic queries are directed to the chatbot. Requests like taking reminders are handled by the module itself and responses are sent to TTS.

iii. Chatbot: In this work, Large Language Model (LLM) used is Vicuna v1.5 (Touvron, 2023), fine-tuned LLaMA2 with 7 billion parameters. This is deployed locally on GPU device. Three different variants of LLM were developed, the details of which are given below.

(a) LLM only: In this variant, the vicuna model is used for interactions. A brief description on the AI voice bot is fed as context to enable it to respond in first person. Since LLM is trained on a huge database, it can generate responses on diverse topics. The limitation is that it is unaware of information specific to ISRO and those outside the training data.

(b) Retrieval chatbot with LLM advisor: The block diagram in Fig. 2 shows the modules involved in this version. It consists of two types of chatbot working together. The retrieval chatbot is a simple one with three neural network layers. Those queries that the retrieval chatbot cannot answer are consulted with the LLM. Otherwise, retrieval chatbot response is sent to the user. The advantages of this mode are (a) Average time of interaction with the AI bot reduces in comparison with the LLM only mode as the retrieval chatbot generates response within milliseconds. (b) ISRO specific and recent information are now available with the AI bot and can be updated easily by training the retrieval chatbot.

(c) RAG: RAG stands for Retrieval Augmented Generation (Lewis, 2020). This is a state-of-the-art method in which documents, directories, etc. are vectorized and stored in a Vector Database (VD). When queried, this information is pulled by a retriever model and sent to context generator. The context generator prepares the information as a prompt and feeds it to the LLM. LLM generates response using prompt. This is implemented using LangChain. By this method, large documents typically useful for astronauts like flight manuals, experiment details, etc. could be easily fed into the voice bot, in addition to new data updates including custom ISRO data.

iv. Text to Speech conversion (TTS): The response from the LLM is fed into the Text to Speech (TTS) module. This is an open source DNN based model by AI4Bharath (Kumar, 2023). The voice model generates voice in Indic-English and Hindi. The female version of the model is chosen currently for speech generation. The TTS also provides a trigger input to the graphics manager to sync between the audio-video content of AI voice bot.

v. Operations Assistant: The Operations Assistant (OA) handles the commanding mode of the AI voice bot. In case of command, the OA initiates a search in the database available locally to figure out the desired response. Block schematics, audio and video files relevant to the application are available in the voice bot memory. For example, the user could ask for block diagram of deboost procedure. In addition to these, OA provides answers to user queries on self-health of the voice bot, like present temperature, battery voltage, etc.

b. **Graphics Manager**: The AI voice bot is incorporated with a face to establish an emotional connection (Goncalves, 2021) with the user. The objective of graphics manager is to ensure that the animated face is in sync with the TTS output while being displayed to the user. Using the phonemes, the sequence of frames required to animate the text are generated. The delay between frames is computed using audio length and number of phonemes. The animator software used is PyGame (PyGame).

C. Bot Hardware Architecture

The voice bot hardware is centered around a high-performance computing platform, the NVIDIA Jetson AGX Xavier Industrial System On Module. A compact industrial carrier board of same form factor was chosen and ruggedized for use in intra vehicular space environment through card level component arresting, mechanical packaging and thermal solutions. The package has been qualified to random vibration level of 11.8 grms and thermo vacuum level of 60 deg and 10^{-5} mbar, typical of crew module avionics systems. Other major components include a microphone, speaker, battery and power management circuitry, which are all packaged into a self-contained flying saucer shaped main body as shown in Fig. 3. A TFT LCD display add on unit is currently used to render the graphical aspects of the voice bot. This is planned to be integrated to the voice bot body in the next version of voice bot as shown in Fig. 3 (inset).

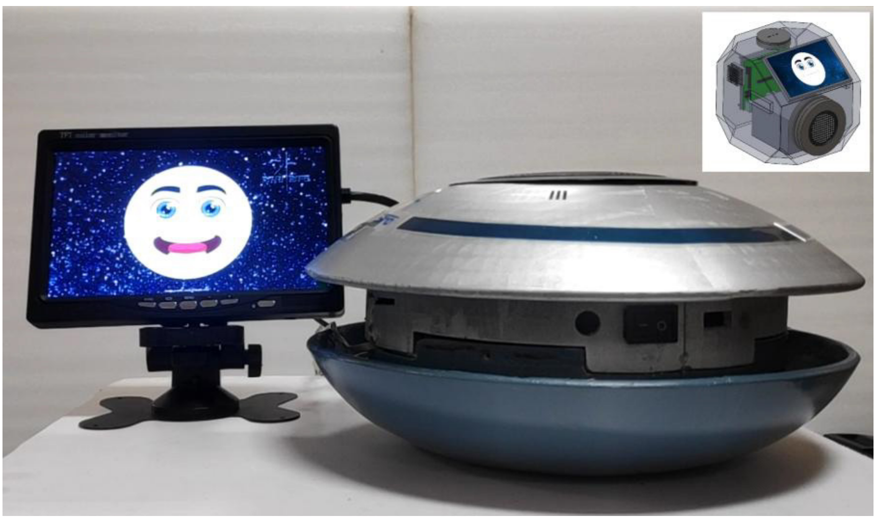

Fig. 3 Voice bot package. Inset is the CAD model of its next version.

D. Results and Discussion

The developed AI voice bot is evaluated using two methods (a) module level performance evaluation using suitable metrics (b) integrated functionality verification. The results obtained are given below.

a. **Module level Evaluation:** The individual modules are evaluated for speed and accuracy and results are tabulated in this section.

(a) **ASR Evaluation**: The metrics used for evaluation are (i) Word Error Rate (WER) and (ii) Real Time Factor (RTF). Nine candidates were asked to read out a sample text to evaluate the ASR performance. The candidates chosen had variability in accent, speed, and loudness. The mean WER is 10.67 % and mean RTF is 0.31 as shown in Table 1(a). Popular ASR engines are found to perform with a WER less than 20 % and RTF less than 1 (Azure, 2024). The ASR model is found to have acceptable performance.

Table 1. ASR and TTS performance evaluation summary

(a) ASR Results			(b) TTS Results			
Sl. No.	**WER (%)**	**RTF**	**Sl. No.**	**C**	**N**	**LE**
A	9.3	0.25	A	10	9	10
B	12.6	0.5	B	9	8.5	9
C	8.4	0.4	C	9	8	9
D	9.4	0.26	D	10	10	10
E	9.44	0.25	E	10	7	10
F	15.6	0.26	F	9	9	9
G	9.7	0.32	G	9	9	9
H	10.6	0.3	H	10	10	10
I	11.03	0.26	I	8	8	9

(b) **Chatbot Evaluation**: Here we evaluate the performance of vicuna v1.5 - 7b LLM deployed on Jetson Xavier 16 GB edge device, using the following metrics. (a) GPU utilization: For every query passed to the LLM, the number of forward passes equals the number of tokens generated. GPU utilization is assessed by the number of tokens generated. It is found that on an average it takes 0.727 seconds to generate one token and the GPU load at that point is 100%. The 16 GB RAM is fully occupied by the LLM model. (b) MT-Bench: Multi-turn benchmark (Bai 2024) is an effective way to test the ability of the LLM to involve in continuous and engaging conversations. Five sets of MT-Bench queries were asked in which each set had three to five correlated follow up questions, both ISRO specific and generic (snippet of test shown in Fig. 4). The LLM is found to generate relevant answers to the queries without losing context in each set. (c) LLM-as-judge: This method involves utilizing the MT-Bench responses for review by another LLM (Zheng 2024). Here, OpenAI's GPT 4 is prompted to evaluate the quality of the response considering factors such as the helpfulness, relevance, accuracy, depth, creativity, and level of detail of the response and scored out of 10. The average score obtained for ten questions is 8.4.

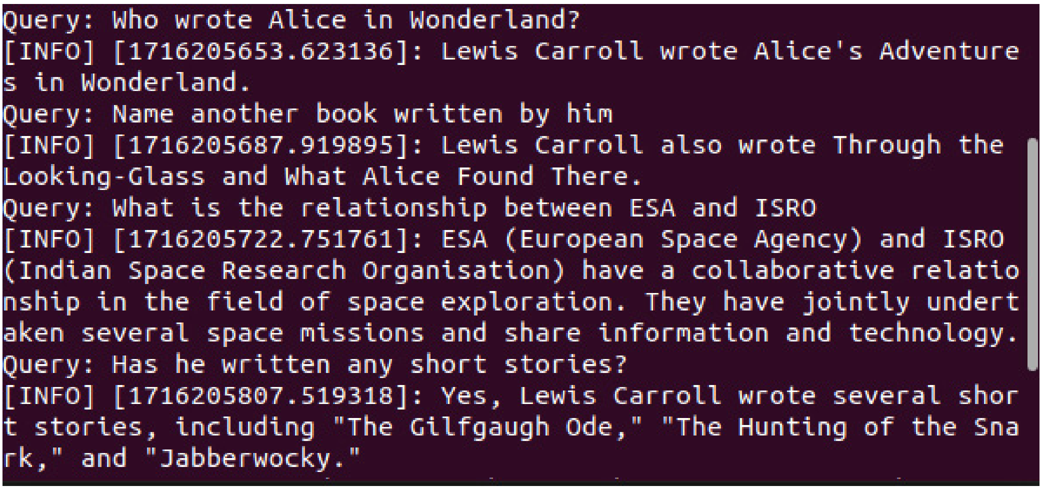

Fig. 4 Snapshot of MT-bench evaluation.

(a) **TTS Evaluation**: The TTS module was tested using Mean Opinion Score (MOS) index among nine candidates. The metrics considered are comprehensibility (C), naturalness (N) and listening effort (LE). The scores are listed in Table 1(b). The MOS for TTS is found to be 9.16 and the performance is found to be satisfactory.

b. **Integrated Functionality Tests:** A total of ten users were selected (six female, four male) for this initial pilot study. The group comprises members from different age groups, voice/speech characteristics, emotional-intellectual abilities and technical expertise. They could ask queries related to the experiments, request block schematic display, play media content, add reminder and recall them at a point later in time. At the end of the session, candidates were asked to rate the following aspects related to the voice assistant. The categories in which user queries came up and their relative percentage are shown in Fig. 5. It can be seen that the voice bot's diverse functionalities were exercised sufficiently by the considered user set. The number of queries per user varied between five to twenty-five, on both space domain and non-space domain topics. The feedback from users is given in Fig. 6 where the average response is noted against each aspect. The overall impression score by the users is found to be good. In the early versions of the chatbot, when the LLM was hallucinating more often, users had suggested incorporating a "stop" keyword. This was implemented by carefully interrupting the token generation using a separate thread and ensuring proper release of system resources by the LLM. The nominal power requirement is 10 - 12 watts and GPU load on LLM usage results in a peak power demand of about 30 watts.

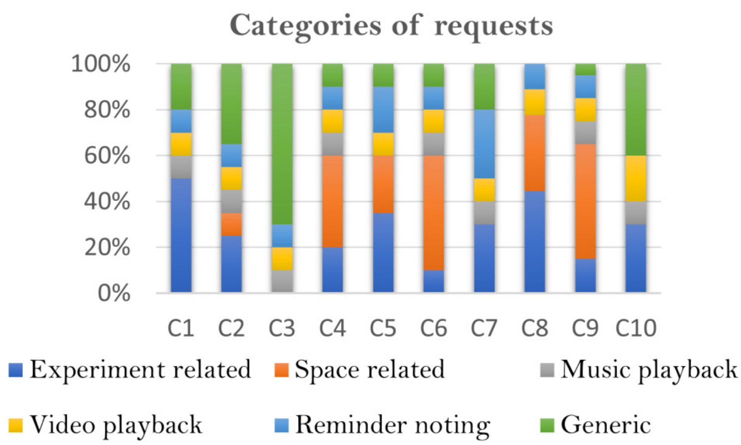

Fig. 5 Categories in which voice bot was tested.

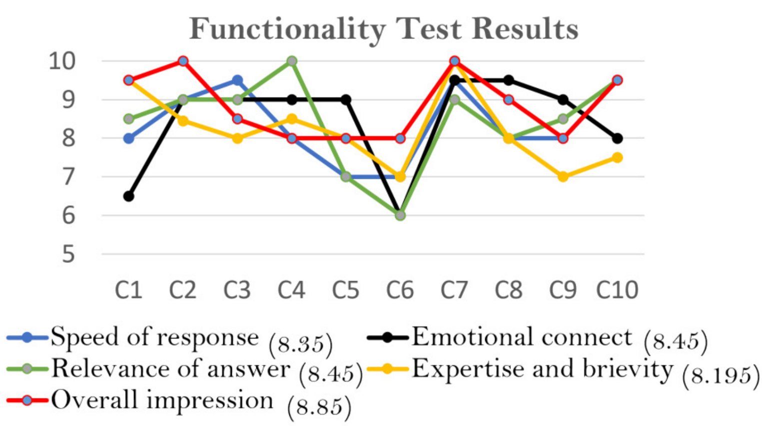

Fig. 6 Functionality test scores out of ten.

E. Discussion and Conclusions

The current version of AI voice bot is found to be promising. However, many more aspects need to be addressed as the system evolves. The speech input processing needs to be fine-tuned to be robust against noise that are likely to be present onboard. User queries in Hindi are currently handled by QA manager thereby limiting the scope. Future versions must incorporate LLMs that support text generation in Hindi for more seamless interactions. The TTS is at times noted to be monotonous and may lead to boredom. Incorporation of intonations in speech to make it more natural is being looked into. The time taken for the voice bot to generate response for a query is about 3.5 s (min) and 25 s (max) on the edge device considered. This could be optimized further by both software modification and hardware upgradation. Transformers with video encoders would enable the voice bot to perform image-based searching. A query interface to mission database could be designed so that the bot can access mission and other parameters and extend its scope of operations. Though the current version is designed to primarily assist the astronauts in carrying out experiments, the scope could be extended by deeper integration with mission databases. The voice bot could then serve as a hands-free ready reference for the astronaut on mission related parameters for quick decision making in critical scenarios. For enabling such applications, the design and realization processes must address safety and reliability aspects at all levels—hardware, software, and system. At hardware and system levels, standard practices in the space industry such as usage of high reliability components, screening, multiple redundancy etc. can be employed, while adhering to power and size constraints. With respect to software, in addition to institutional quality practices related to software life cycle, special attention needs to be paid to LLM due to its generative nature. More sophisticated methods integrating diverse data sources and improved model explainability need to be developed in order to enable application of such bots in supporting mission critical decision-making scenarios. Integration of camera to this bot, which is currently in progress, would enable the bot to make more personalized interactions using techniques like face detection, recognition, emotion identification, etc. Since the voice bot interacts with astronauts continuously, a feature to evaluate their emotional health would be another value addition to this product

F. Acknowledgement

The authors would like to thank Shaheen M. P., and Dhanesh S. for the design and prototyping of the voice assistant chassis. The authors express their gratitude to Aswathy A., Abhilash B. R., Jyothish M., Muhammed Rashid K., Abhay Dileep, Rinu Chacko, Ishan Meshram, Premkrishnan M. K., Abhilash M., Ahmed Rabeeh I. C., Devu Suresh, Bharath D., Shiniamol Jacob, Maya S. Pillai, Ardra Prasanth, Usha V., Aravind R., and Devika P. for participating in the test and evaluation of the AI voice bot.

References

Ambrose, R. O., Aldridge, H., Askew, R. S., Burridge, R. R., Bluethmann, W., Diftler, M., ... & Rehnmark, F. (2000). Robonaut: NASA's space humanoid. *IEEE Intelligent Systems and Their Applications*, *15*(4), 57-63.

Azure. *Test accuracy of a custom speech model.* (n.d.). Retrieved March 20, 2024, from https://learn.microsoft.com/en-us/azure/ai-services/speech-service/index-speech-to-text

Bai, G., Liu, J., Bu, X., He, Y., Liu, J., Zhou, Z., ... & Ouyang, W. (2024). Mt-bench-101: A fine-grained benchmark for evaluating large language models in multi-turn dialogues. arXiv preprint arXiv:2402.14762.

Bualat, M. G., Smith, T., Smith, E. E., Fong, T., & Wheeler, D. W. (2018). Astrobee: A new tool for ISS operations. In *2018 SpaceOps Conference*, 2517.

Dong, X. L., Moon, S., Xu, Y. E., Malik, K., & Yu, Z. (2023, August). Towards next-generation intelligent assistants leveraging llm techniques. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 5792-5793.

Fredrickson, S. E., Duran, S., Howard, N., & Wagenknecht, J. D. (2004, August). Application of the mini AERCam free flyer for orbital inspection. In *Spacecraft Platforms and Infrastructure*, *5419*, 26-35. SPIE.

Goncalves Freitas, A. R., Schülke, A., Glaser, S., Michelmann, P., Chi, T. N., Schröder, L. M., ... & Schöning, J. (2021, May). Conversational user interfaces to support astronauts in extraterrestrial habitats. In *Proceedings of the 20th International Conference on Mobile and Ubiquitous Multimedia*, 169-178.

Henriquez, M., Herd, A., Danènas, P., Dilijonas, D., & Ontiveros, J. (2023, September). CLARK: Building conversational intelligence for knowledge management in the space domain. In *European Conference on Knowledge Management*, *24*(1), 541-549.

Hoppenbrouwers, T., Ferra, L., Markus, M., & Wolff, M. (2017). Operations data files, driving force behind international space station operations. *Acta Astronautica*, *138*, 255-261.

Ilavazhagi, G. (2023). A comprehensive overview of ISRO's ambitious space station project and collaborative endeavors. *Acceleron Aerospace Journal*, *1*(5), 106-107.

Klopfenstein, L. C., Delpriori, S., Malatini, S., & Bogliolo, A. (2017, June). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. In *Proceedings of the 2017 conference on designing interactive systems*, 555-565.

Kulkarni, P., Mahabaleshwarkar, A., Kulkarni, M., Sirsikar, N., & Gadgil, K. (2019, September). Conversational AI: An overview of methodologies, applications &

future scope. In *2019 5th International conference on computing, communication, control and automation (ICCUBEA)*, 1-7. IEEE.

Kumar, V., Srivastava, P., Dwivedi, A., Budhiraja, I., Ghosh, D., Goyal, V., & Arora, R. (2023, December). Large-Language-Models (LLM)-Based AI Chatbots: Architecture, in-depth analysis and their performance evaluation. In *International Conference on Recent Trends in Image Processing and Pattern Recognition*, 237-249. Cham: Springer Nature Switzerland.

Kumar, G. K., Praveen, S. V., Kumar, P., Khapra, M. M., & Nandakumar, K. (2023, June). Towards building text-to-speech systems for the next billion users. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1-5. IEEE.

Landay, J. A., Oliver, N., & Song, J. (2019). Conversational user interfaces and interactions. *IEEE Pervasive Computing*, *18*(02), 8-9.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.

Mitani, S., Goto, M., Konomura, R., Shoji, Y., Hagiwara, K., Shigeto, S., & Tanishima, N. (2019, March). Int-ball: Crew-supportive autonomous mobile camera robot on ISS/JEM. In *2019 IEEE Aerospace Conference* (pp. 1-15). IEEE.

Mihaylov, D., Vasile, M., Herd, A., Broughton-Stuart, G., & Moshfeghi, Y. (2022, September). A Space Conversational Agent for Retrieving Lessons-learned and Expert Training. In *73rd International Astronautical Congress, Paris, France* (pp. 18-22).

PyGame. PyGame. Retrieved March 20, 2024, from https://www.pygame.org/

Ross, S. I., Martinez, F., Houde, S., Muller, M., & Weisz, J. D. (2023, March). The programmer's assistant: Conversational interaction with a large language model for software development. In *Proceedings of the 28th International Conference on Intelligent User Interfaces* (pp. 491-514).

Salah, M., & Elmowafy, O. (2023). The Analysis of Future Robots in Japan. *Multidisciplinary Journal of Engineering Sciences*, *1*, 12-35.

Schröder, V., Regele, R., Sommer, J., Eisenberg, T., & Karrasch, C. (2018, October). GNC system design for the crew interactive mobile companion (CIMON). In *Proceedings of IAF Human Spaceflight Symposium at 69th International Astronautical Congress.*

Shashkova, E., Navarro, G., Roy, R. N., Paillet, A., & Truntzler, L. (2022). Study and development of an AI assistant for future Moon and Mars stations in *73rd International Astronautical Congress IAC), Paris, France.*

Sundararajan, V. (2021). Gaganyaan Mission: Socio-technical Systems (STS) Design and Strategic Perspectives of India's Human Spaceflight Program. In *AIAA Scitech 2021 Forum* (p. 0306).

Yuste, M. D. (2022). EVA (ESA Virtual Assistant): The conversational AI platform for Space. *SciOps 2022: Artificial Intelligence for Science and Operations in Astronomy (SCIOPS)*, 9.

Tulshan, A. S., & Dhage, S. N. (2019). Survey on virtual assistant: Google assistant, siri, cortana, alexa. In *Advances in Signal Processing and Intelligent Recognition Systems: 4th International Symposium SIRS 2018, Bangalore, India, September 19–22, 2018, Revised Selected Papers 4* (pp. 190-201). Springer Singapore.

Vega, T., Bueno, A., Cerutti, C., Chang, H., Garvin, M., Garza, C., ... & Underwood, L. (2022, March). Enabling a Voice Management System for Space Applications. In *IEEE Aerospace Conference*.

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Zheng, L., Chiang, W. L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., ... & Stoica, I. (2023). Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in Neural Information Processing Systems*, *36*, 46595-46623.

Conflict of Interests

There are no conflicts of interest.