Intelligent Response and Interaction System (IRIS)

FY21 Closeout Report

Project Information

- **Team Members**
  - Aly Shehata – PI
  - Lui Wang – Advisor
  - William Baker – Lead Software Developer

- **Taxonomy Alignment**
  - TX02.2.4 - Low Power Embedded Computer Systems
  - TX02.2.8 - Use of Advanced Commercial-off-the-Shelf (COTS) Technologies
  - TX04.4.1 - Multi-Modal and Proximate Interaction
  - TX06.2.3 - Informatics and Decision Support Systems
  - TX06.6.4 - Operations Effectiveness
  - TX07.3.2 - Integrated Flight Operations Systems (Autonomous crew operations)

- **TRL Maturation**
  - TRL 3 (Start) to 4 and 5 (End)

Executive Summary

In the second year, the IRIS team developed the components necessary for successful offline deployment of the IRIS services. This includes custom automated speech recognition training on NASA audio data, and in-house development and integration of online and offline conversational services. Lastly, the team worked on integrating the IRIS technology with stakeholders and projects that have a strong need for voice interaction.

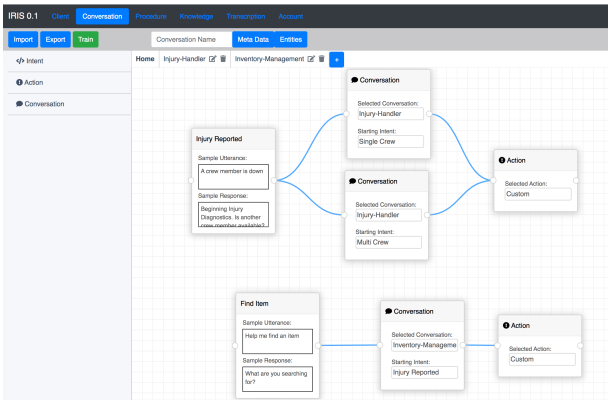


Figure 1: IRIS Web Client provides the capabilities needed for a novice user to author multi-modal conversations (Commands, Conversations, Actions)

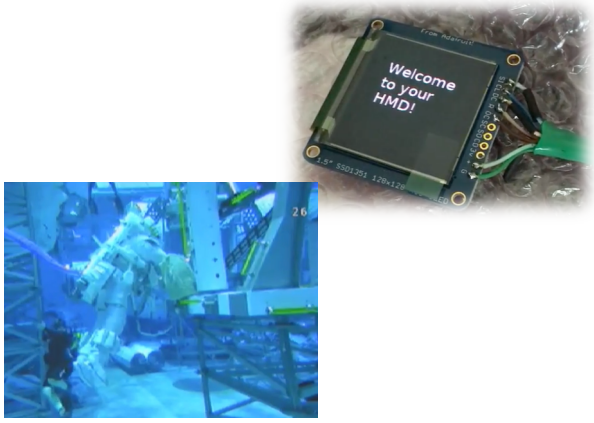


Figure 2: IRIS HMD Deployment at the NBL

Background

Crew Assistance technologies will be an essential element to the upcoming Lunar exploration missions. Operations and training burdens must be minimized, to meet Lunar & Mars challenges and compressed schedule. To achieve crew autonomy and independence from the ground support, IRIS provides scalable conversational technologies to assist crew during deep space missions, especially during communication outages.

The IRIS project continued maturing the intelligent response and interaction system developed in FY20, enhancing autonomy, and reducing training and operational burdens for crew during complex spaceflight operations. As a recap, IRIS provided the following (online) capabilities at the end of year 1 –

- Hands-free Conversational **Procedure Assistant**
- Auxiliary Spacecraft Systems **Command & Control**
- **Querying** of Mission Logistics and Timeline Management

While recent conversational voice systems have drastically improved in terms of accuracy and scalability, these systems do not work out of the box for unique harsh spaceflight environments. Several gaps must be addressed for scalable, robust, and intuitive voice interaction.

Plans and Objectives

Today, industry utilizes heavy cloud-based machine learning technologies to achieve state-of-the-art conversational systems (such as Siri and Alexa). These machine-learning based systems tend to produce a much more generalized and robust solution that can generalize to different variants of dialogue and noise environments. Typically, these technologies would not be available in an offline space environment with minimal resources.

Our approach is to use the on-prem machine learning computing paradigms to establish the needed voice infrastructure to support crew deep space missions. Building upon recent breakthroughs in voice and conversational processing technology, on-the-edge machine learning solutions are also quickly maturing, allowing NASA to begin deploying these conversational agents in an offline environment, with a minimal hardware footprint. Our objective is to build a highly dialogue-based voice assistant can be built to support a wide variety of space-flight use-cases and deployment environments.

For the second year, the plan was to focus on additional software development, technology maturation, infusion, integration, and testing. Our primary area of focus included developing on-premises solutions to the Automated Speech Recognition (ASR) and Natural Language Understanding (NLU) components. These on-premises solutions open the door for continuous improvement, customization per use-case, and end-to-end deployment strategies. The team also supported several existing use-cases and projects that utilized IRIS, such as the H3PO, JARVIS, and One Portal teams.

Through collaboration with commercial partners, a tailored automated speech recognition (ASR) solution for custom NASA use was trained and delivered to NASA (enabling further customizations in the future and supporting an offline environment). The team also developed offline Natural Language Understanding (NLU) components, ensuring that is interchangeable with the online alternatives, and similarly performant.

Partnerships

Internal Partnerships Include –

- JARVIS – EV3
  - Develop and integrate a light-weight voice control system for xEMU
- H3PO – SA
  - Support Head Mounted Display HITL Tests using IRIS Voice Control
- One Portal – ER
  - Integrate IRIS Voice Control with the One Portal Exercise Software

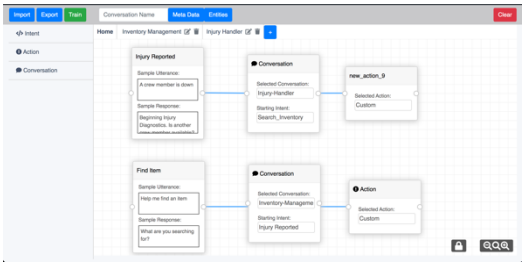
External Partnerships Include –

- Google Cloud Platform/Applied AI
  - Design and Development Consultation
  - Cloud Compute and AI Capabilities
- Deepgram
  - Custom NASA Automated Speech Recognition Models and Training Pipeline
  - Deployment of ASR model on NASA premises
- Picovoice Technologies
  - Lightweight voice utilities and infrastructure for wake word and intent detection.

Results and Knowledge Gained

❖ **End-to-End Conversational System**

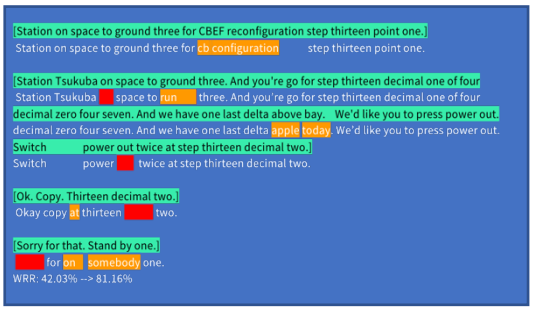
The IRIS team has developed the components needed to rapidly train highly conversational voice agents using a mix of off-the-shelf and custom in house technologies within the IRIS framework. IRIS provides the necessary architecture to support the end-to-end use-case for procedures, command-and-control, and knowledge-based questions. This allows users of the IRIS framework to rapidly prototype and wireframe conversations around already existing procedures or auxiliary systems.



IRIS Web Portal for Conversation Training

❖ **Automated Speech Recognition – Limitations/Constraints for NASA Audio**

The IRIS team collaborated with Deepgram (a commercial partner with tailored speech recognition solutions0 to train a speech recognition model robust enough for space-to-ground audio. The model can inherently understand acronyms and human spaceflight terminology, per the ~50 hour of EVA public data it was trained on. The word recognition rate improved from ~40% to over 80% in accuracy.



Custom NASA Model Improvements
Automated Speech Recognition (ASR)

❖ **Natural Language Understanding (NLU) – Scalability and Deployment Limitations**

The IRIS team developed the components for an offline NLU Manager. The offline component is compatible with existing IRIS conversations, allowing for seamless backwards compatibility between the online NLU (Dialogflow) and the offline NLU component (Snips AI)

❖ **Integration and Infusion**

The IRIS Voice Assistant has successfully been deployed and demonstrated in several ground and flight environments including the NBL, JARVIS (xEMU), and the crew exercise lab (One Portal)

- IRIS is seeking additional ground training data to improve robustness of current customized automated speech recognition solution.

Technology Maturation Opportunities

As machine learning applications become more widely adopted throughout the NASA software and hardware community, certification, and integration of these new computing platforms, within flight systems, will be necessary.

- The IRIS Voice Assistant has successfully been deployed and demonstrated in several ground and flight environments including the NBL, JARVIS (xEMU), and the crew exercise lab (One Portal)
- IRIS is seeking additional ground training data to improve robustness of current customized automated speech recognition solution.