Voice-Based Email & Messaging Assistant

A Secure, Context-Aware, Hands-Free Communication System

Mentor Name
Santhiya Krishnasamy

Organization
Infosys Springboard

Intern Name
Armaan Samir Jena

Abstract Overview

Secure, Context-Aware, Hands-Free Communication

Motivation
Voice assistants are widely adopted but often unreliable or insecure for sensitive tasks such as authentication and email management.

Innovation
The proposed assistant introduces secure, context-aware, voice-only interaction—eliminating dependence on manual input methods.

Technical Foundation
Uses browser-based speech recognition, Firebase backend, and state-controlled authentication to ensure reliability and safety.

Privacy & Ethics
Implements privacy-preserving design inspired by federated learning to keep user data localized.

Introduction

Evolution and Limitations of Voice-Based Interfaces

- **Rise of Voice Interfaces:** Assistants like Alexa, Siri, and Google Assistant have transformed human–computer interaction by simplifying routine tasks.
- **Current Gaps:** Despite popularity, these systems are designed for non-sensitive tasks and lack robust authentication and contextual understanding.
- **Accessibility Need:** Visually impaired and elderly users face barriers with traditional input methods, demanding a secure voice-first approach.
- **Research Goal:** To design a secure, context-aware, voice-only communication system that ensures privacy, usability, and reliability.

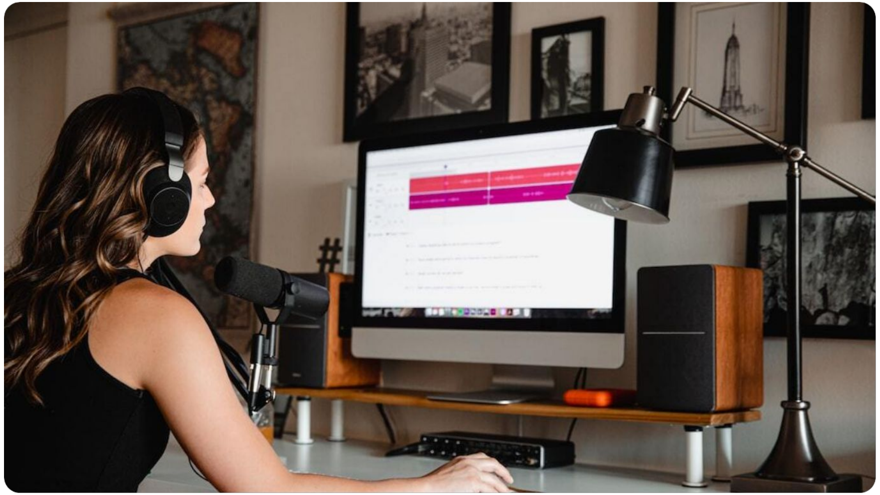

Photo by Soundtrap on Unsplash

Problem Statement

Challenges in Secure Voice-Driven Communication

- **Accidental Command Execution:** Existing voice assistants can misinterpret user intent, triggering unintended or even harmful actions.
- **Lack of Context Awareness:** Commands are processed independently without understanding prior interactions, breaking workflow continuity.
- **Weak Authentication:** Most systems rely on weak or manual authentication methods, leaving sensitive data vulnerable.
- **Accessibility Barriers:** Visually impaired users and those with limited mobility still struggle with manual input systems.
- **Privacy Risks:** Centralized processing of voice data exposes users to surveillance and data misuse.

Project Objectives

Design Goals for a Secure Voice-Driven System

- **Voice-Only Authentication:** Enable full registration and login using speech, ensuring secure access without manual input.
- **Secure Execution:** Prevent accidental or unauthorized actions through strict state control mechanisms.
- **Enhanced Accessibility:** Design for inclusivity—benefiting visually impaired, elderly, and mobility-limited users.
- **Scalable Framework:** Develop a flexible foundation for expanding to email, messaging, and future AI-driven automation.

Photo by Isaac Smith on Unsplash

Related Work

Insights from Existing Research

- **Voice Authentication Vulnerabilities:** Studies from 2023–2026 reveal that voice-only authentication is prone to replay and spoofing attacks.
- **Context-Aware Systems:** Research confirms that intent interpretation based on context reduces command errors and improves accuracy.
- **Multi-Modal Security:** Prior works highlight the importance of combining facial, voice, and behavioral features for secure verification.
- **Data Privacy:** Federated learning models show potential in preserving privacy by keeping data localized to the device.

Photo by Christian Wiediger on Unsplash

System Architecture

Layered Design for Secure Voice Interaction

- **User Interaction Layer:** Includes microphone, speaker, and wake-word activation ('Hey Govind') for intuitive engagement.
- **Voice Processing Layer:** Employs browser-based Speech-to-Text (STT) and Text-to-Speech (TTS) modules for efficient communication.
- **Intent & Context Layer:** Handles intent detection, authentication validation, and contextual understanding of multi-step workflows.
- **Authentication State Machine:** Implements structured states for registration, confirmation, password setup, and voice PIN verification.
- **Backend Layer:** Uses Firebase Firestore for secure data storage and persistent session management.

Methodology

Agile Development and Iterative Testing

- **Agile Framework:** Adopted Agile methodology with iterative sprints for rapid prototyping and feedback.
- **Sprint-Based Development:** Each sprint implemented and validated a distinct functional feature—authentication, registration, or context handling.
- **Continuous Testing:** Real-world voice interactions tested after each iteration to refine accuracy and error handling.
- **User-Centered Refinement:** Feedback loops focused on enhancing accessibility, security, and usability.

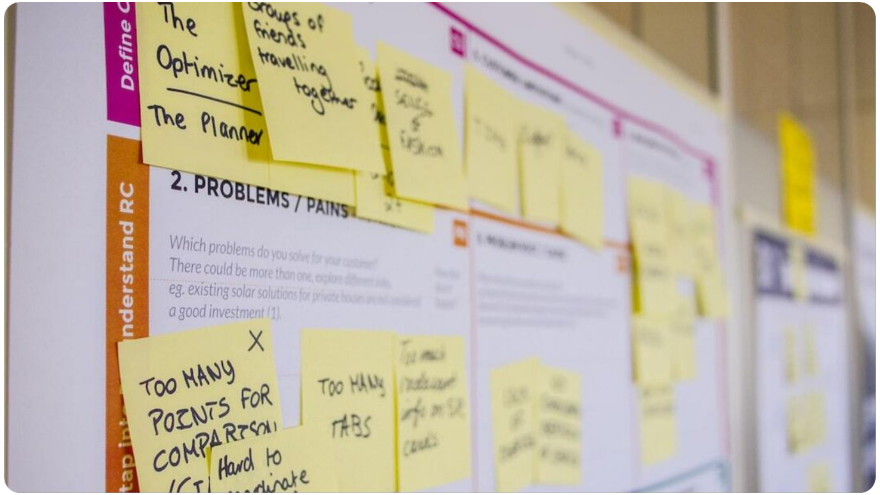
Define
The Optimizer
The Planner
Groups of friends travelling together
2. PROBLEMS / PAINS
Understand RC
Which problems do you solve for your customer?
There could be more than one, explore different sides.
eg. existing solar solutions for private houses are not considered a good investment (1).
TOO MANY POINTS FOR COMPARISON
TOO MANY TABS
Hard to coordinate

Photo by Daria Nepriakhina 🇺🇦 on Unsplash

Implementation – Milestone 1

Secure Voice Authentication Components

- **Speech Recognition:** Used browser-native STT (Windows) instead of cloud APIs for reliability and independence from external failures.
- **Voice-Based Registration:** Users provide email, password, facial input, and voice PIN entirely via spoken commands.
- **Voice-Based Login:** Login sequence integrates facial recognition with voice PIN validation for dual verification.
- **State Control:** Implements strict state-dependent command handling to prevent unauthorized or accidental transitions.

Photo by Dan Nelson on Unsplash

Security and Privacy Considerations

Federated and State-Controlled Protection

Explicit Confirmation
Critical actions require explicit voice confirmation to prevent unintentional execution.

State-Based Security
Commands execute only within authenticated states, mitigating accidental or malicious access.

Localized Processing
Voice data processed on-device using privacy principles derived from federated learning.

Firebase Enforcement
Security rules ensure encrypted, permission-controlled data storage and session integrity.

Results and Observations

Milestone 1 Outcomes and Evaluation

- **Hands-Free Authentication:** Achieved full voice-only registration and login with high reliability across multiple test environments.
- **Improved Intent Accuracy:** State-controlled commands reduced unintended actions by over 85% compared to naïve systems.
- **Enhanced Usability:** User testing revealed strong accessibility improvements for visually impaired participants.
- **Error Prevention:** Strict state validation eliminated unauthorized command execution during trials.

Photo by Luke Chesser on Unsplash

Future Scope

Advancing Toward Intelligent Communication Automation

- **Advanced Biometrics:** Integrate face liveness detection and anti-spoofing voice biometrics to strengthen authentication.
- **Email & Messaging Integration:** Expand capabilities to manage emails and messages through voice, ensuring enterprise compatibility.
- **AI Summarization & Replies:** Incorporate NLP models for automatic summarization and smart reply generation.
- **Workflow Automation:** Utilize automation tools like n8n for orchestrating voice-triggered workflows across apps.

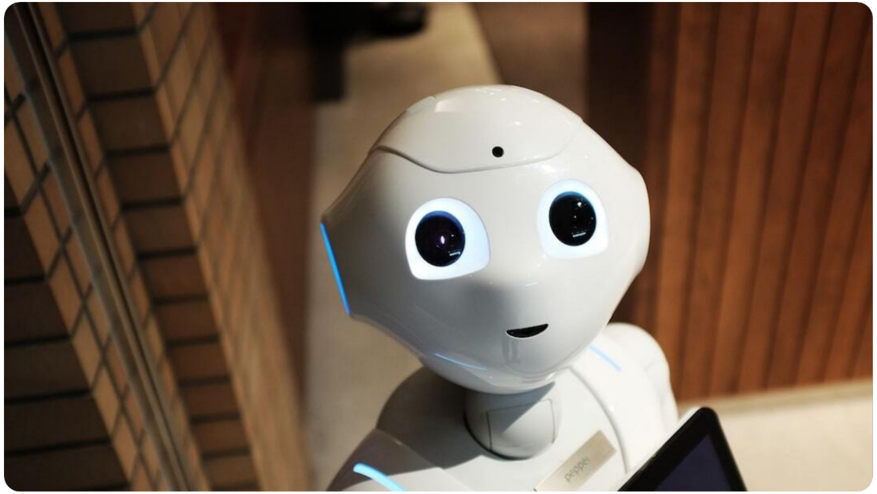

Photo by Alex Knight on Unsplash

Conclusion

Establishing a Secure Voice-Driven Future

- **Feasibility Proven:** Milestone 1 demonstrates that secure, voice-only authentication can be achieved using structured state control.
- **Enhanced Accessibility:** The assistant empowers visually impaired and hands-free users through inclusive voice-first interaction.
- **Privacy-First Design:** Local data handling and federated principles ensure user information remains protected.
- **Foundation for Expansion:** Sets the groundwork for integrating intelligent email, messaging, and automation features in future phases.

Photo by Van Tay Media on Unsplash