Voice-Based Email & Messaging Assistant

Hands-Free, Secure, Context-Aware Communication

Milestone -2 Presentation

Intern Name
Armaan Samir Jena

Mentor Name
Santhiya Krishnasamy

Organization
Infosys Springboard

Project Overview

Voice-Based Email & Messaging Assistant

- **Project Title:** Voice-Based Email & Messaging Assistant
- **Core Goal:** Enable secure, voice-driven email and messaging operations with contextual awareness.
- **Target Users:** Individuals seeking hands-free, accessible communication (especially mobility-impaired users).
- **Key Challenge:** Balancing automation, privacy, and platform restrictions.
- **Final Architecture:** Hybrid Model integrating Gmail API, Telegram Bot, and assistive WhatsApp Web.

Problem Statement

Challenges in Voice-Based Communication

Typing Limitations
Typing-based communication is time-consuming and not suitable for hands-free use cases.

Accessibility Issues
Users with mobility or vision impairments face barriers in using traditional email/messaging apps.

Voice Assistant Gaps
Existing assistants like Alexa or Google Assistant cannot perform secure, context-aware email or chat automation.

Accidental Actions
Voice misinterpretation can lead to incorrect or unintended message sending, posing security and privacy risks.

Project Objective

Security-First Voice Communication Design

- **Hands-Free Authentication**: Enable complete voice-only authentication without reliance on keyboards or touch input, ensuring usability for accessibility-focused and hands-busy environments.
- **Controlled Command Execution**: Prevent unintended or unsafe actions by enforcing explicit authentication and confirmation before any communication-related command is executed.
- **Context-Aware State Control**: Introduce a state-driven interaction model where the system understands whether the user is unauthenticated, authenticated, or executing privileged actions.
- **Secure Communication Access**: Provide authenticated access to emails and messages while preserving confidentiality, integrity, and intentional user control.
- **Scalable Automation Foundation**: Design the system architecture to support future integrations, workflows, and AI-driven automation without compromising security.

System Overview

Voice-Based Email & Messaging Assistant

- **Voice-First Interaction Model**: The system is designed around a voice-only interaction paradigm, eliminating dependence on screens or keyboards and enabling true hands-free communication.
- **Authentication-Gated Operations**: All sensitive operations such as reading, sending, or managing emails are strictly blocked until the user successfully completes multi-factor authentication.
- **Cloud-Backed Architecture**: Backend services are powered by cloud infrastructure to manage user identities, session states, and secure integrations with third-party communication platforms.
- **Accessibility-Centric Design**: The assistant is optimized for visually impaired users by providing auditory feedback, confirmation prompts, and error handling through voice.
- **Core Security Philosophy**: The guiding principle of the system is simple and enforceable: No authentication, no action.

System Architecture

High-Level Layered Design

- **User Interaction Layer**: Handles audio input and output through microphones and speakers, capturing voice commands and delivering spoken feedback to the user.
- **Voice Processing Layer**: Performs speech-to-text and text-to-speech conversion, transforming raw audio into structured commands and responses.
- **Intent and Context Layer**: Interprets user intent and maintains conversational context to determine whether commands are permissible based on the current system state.
- **Authentication State Machine**: Controls user state transitions such as unauthenticated, partially authenticated, and fully authenticated, enforcing security boundaries.
- **Backend and Cloud Services**: Manages user data, authentication records, session persistence, and secure integration with external services such as email providers.

Milestone-1

Focus and Scope

- **Primary Objective**: Design and implement a fully voice-driven registration and login pipeline without relying on traditional input devices such as keyboards or touchscreens.
- **Wake-Word Activation**: Introduce controlled wake-word detection to ensure the assistant listens only when explicitly invoked by the user.
- **Voice-Based Registration**: Enable new users to create accounts using voice commands, capturing identity attributes without manual input.
- **Multi-Factor Authentication**: Combine face verification and a voice PIN to strengthen identity assurance and reduce impersonation risk.
- **State-Controlled Workflow**: Enforce a strict step-by-step authentication sequence using a state machine to prevent bypass or accidental progression.

Milestone-1

Key Outcomes

- **Fully Hands-Free Authentication**: Achieved end-to-end user registration and login using only voice interaction, eliminating dependence on keyboards, touchscreens, or manual inputs.
- **Multi-Factor Security Enforcement**: Implemented layered authentication combining facial verification and a voice-based PIN, significantly strengthening identity assurance.
- **Accidental Command Prevention**: State-controlled workflows ensured that no sensitive operation could be triggered without completing required authentication steps.
- **Accessibility Improvement**: Enabled secure system access for visually impaired users through voice prompts, confirmations, and audio-based feedback.
- **Extensible Authentication Pipeline**: Delivered a robust and reusable authentication framework ready to be integrated with real-world communication services.

Why Milestone-2 Was Needed

From Authentication to Real Communication

- **Authentication Without Utilization**: Milestone-1 successfully secured user identity, but the system lacked any real-world communication functionality such as email or messaging access.
- **No Email or Messaging Integration**: The assistant could authenticate users but could not perform meaningful post-login actions, limiting practical applicability.
- **Missing Session Continuity**: Authenticated state was not persisted across actions, preventing smooth transitions from login to task execution.
- **Need for End-to-End Workflow**: A complete system required extending authentication into controlled execution of real communication workflows.
- **Milestone-2 Objective**: Bridge the gap between secure authentication and real-world email interaction while preserving strict security guarantees.

Milestone-2

Core Focus

- **Secure Gmail Integration**: Extend the authenticated voice assistant to interact with Gmail, enabling users to read and manage emails only after successful verification.
- **Authenticated Session Persistence**: Maintain user authentication state across multiple voice commands to support continuous, interruption-free workflows.
- **Post-Verification Command Control**: Ensure that voice commands are executed only when the system confirms the user is in a valid authenticated state.
- **User Data Persistence**: Store user profiles, connected services, and session metadata securely to enable reliable and repeatable interactions.
- **UI and State Synchronization**: Synchronize interface views with backend authentication states to prevent unauthorized access or state mismatches.

Milestone-2

Technical Implementation

- **Gmail Integration via OAuth 2.0**: Implemented secure Gmail access using OAuth 2.0, ensuring delegated authorization without exposing user credentials to the client application.
- **Firebase Firestore**: Used Firestore to store user profiles, linked email accounts, authentication states, and session metadata in a scalable NoSQL structure.
- **Firebase Storage**: Stored facial images securely for verification purposes while enforcing access controls and storage security rules.
- **Session State Management**: Maintained authenticated session continuity across voice commands, enabling seamless transition from login to email interaction.
- **State-Synchronized UI**: User interface components dynamically adapt based on authentication and session states to prevent unauthorized navigation.

Milestone-2

Security Enhancements

- **OAuth Token Isolation**: Access tokens are securely scoped and isolated per user session, preventing token reuse, leakage, or cross-user privilege escalation.
- **No Client-Side Credential Storage**: Sensitive credentials and authentication artifacts are never stored on the client, reducing exposure to device-level compromise.
- **Explicit Voice Confirmation**: Before executing critical email actions, the system requires explicit spoken confirmation to ensure user intent and prevent accidental execution.
- **Firebase Security Rules**: Strict backend security rules enforce access control policies at the database and storage layers, independent of client behavior.
- **Session Termination on Logout**: User sessions are explicitly invalidated on logout, immediately revoking access tokens and clearing authentication state.

Results and Observations

Milestone-2 Evaluation

- **Successful Gmail Integration**: Authenticated users were able to reliably access and interact with Gmail using voice commands without security violations or session failures.
- **Zero Accidental Email Actions**: State validation and explicit confirmation mechanisms prevented unintended email reads or sends during testing.
- **Stable Session Persistence**: Authentication state was maintained consistently across multiple commands, enabling smooth login-to-action transitions.
- **Improved User Trust**: Users demonstrated higher confidence in issuing voice commands due to visible authentication checks and confirmations.
- **Enhanced Hands-Free Usability**: The system proved effective in scenarios requiring minimal physical interaction, validating its accessibility goals.

Future Scope

Next Milestones and Enhancements

- **Advanced Face Liveness Detection**: Introduce liveness checks to prevent spoofing attacks using photos or recorded videos during facial verification.
- **Voice Biometric Anti-Spoofing**: Enhance voice authentication using biometric models capable of detecting replay and synthesis attacks.
- **Multi-Platform Messaging Integration**: Extend secure voice interaction to platforms such as WhatsApp and Slack for unified communication.
- **AI-Based Email Intelligence**: Enable automated email summarization and context-aware reply generation using NLP models.
- **Workflow Automation**: Integrate automation tools such as n8n to enable secure, voice-triggered workflows and task orchestration.

Conclusion

Key Takeaways

- **Feasibility of Secure Voice Authentication**: The project demonstrates that robust, multi-factor authentication can be achieved using voice-first interaction models.
- **Importance of Context Awareness**: State-driven control is essential to prevent accidental actions and enforce security in voice-based systems.
- **Bridging Authentication and Communication**: Milestone-2 successfully connected secure authentication with real-world email usage.
- **Scalable and Secure Foundation**: The system architecture supports future expansion into automation and multi-platform messaging without compromising security.
- **Redefining Voice Assistants**: This work shifts voice assistants from convenience tools to secure communication platforms.

Demo and Next Steps

Milestone-3 Roadmap

- **Live System Demonstration**: Demonstrate the complete workflow including wake-word activation, voice-based login, authentication verification, and secure Gmail interaction.
- **End-to-End Flow Validation**: Show real-time transition from unauthenticated to authenticated state and controlled execution of email commands.
- **Codebase Extension**: Extend the existing architecture to incorporate advanced biometric security and additional communication platforms.
- **Milestone-3 Focus**: Shift focus toward intelligence, automation, and anti-spoofing while maintaining security-first design principles.
- **Research and Deployment Path**: Prepare the system for extended testing, performance evaluation, and potential real-world deployment scenarios.

Thank You

Questions & Discussion

Project
Voice-Based Email & Messaging Assistant – A Secure, Context-Aware, Hybrid Voice Automation System

Internship Platform
Infosys Springboard

Next Steps
Open for Q&A, feedback, and mentor discussion.