Archives available at journals.mriindia.com

International Journal on Advanced Electrical and Computer Engineering

ISSN: 2349-9338
Volume 14 Issue 01, 2025

Jarvis 2.0: AI-Powered Conversational Assistant with Voice and Chatbot Interaction and Dashboard Management

[1]Prof. P.U. Chavan, [2] Nimbhore Rushika, [3] Shaikh Nilophar, [4] Sutar Gayatri
[1]Assistant Professor, Dept. of Computer Engineering, SBPCOE, Indapur
[2,3,4]Student of Final Year, Dept. of Computer Engineering, SBPCOE, Indapur
Email: Poojabhoite17@gmail.com , rushikanimbhore2003@gmail.com, shaikhnilofar0027@gmail.com, gayatrisutar838@gmail.com

Peer Review Information

Submission: 11 Sept 2025

Revision: 10 Oct 2025

Acceptance: 22 Oct 2025

Keywords

AI Voice Assistant, MERN Stack, MongoDB, Express.js, React.js, Node.js, Speech Recog- nition (ASR), Text-to-Speech (TTS), NLP, NLU, Dialogflow, Rasa, Browser-Based Voice Assistant, Real-Time Interaction, Multi-Turn Conversations, Task Automation, RESTful APIs, Dashboard Management, GUI Design, Voice Command Processing, IoT Integra- tion, HCI, Conversational AI, Emotion Sentiment Analysis, UX, Performance, Security

Abstract

Artificial Intelligence (AI) virtual assistants have rapidly evolved, becoming integral to modern user interaction across various domains, including customer service, personal productivity, and smart environments. This literature survey investigates the develop- ment and implementation of AI-powered voice and chatbot assistants, with a specific focus on virtual voice assistants. The study reviews key technologies such as automatic speech recognition (ASR), natural language processing (NLP), and text-to-speech (TTS), which enable seamless human-AI interaction. It also explores the use of the MERN stack (MongoDB, Express.js, React.js, Node.js) for building scalable, real-time, and interactive assistant platforms. Particular emphasis is placed on dashboard management systems that enhance transparency, allow customization of assistant behavior, and provide in- sights into user interactions. Through a comprehensive analysis of existing systems and methodologies, this paper aims to provide a foundation for the design and development of Jarvis 2.0, an AI virtual assistant that supports both voice and text-based communication with intuitive dashboard control.

Introduction

In recent years, AI-based voice assistants have revolutionized human-computer interaction by enabling hands-free, voice-controlled operations. Widely integrated into smartphones, smart speakers, and web applications, they perform tasks such as controlling home ap- pliances, playing music, answering queries, and more. Early virtual assistants referred to remote human support, but modern voice assistants leverage AI to understand user in- tent and perform actions autonomously. A typical voice assistant functions in three main stages: speech-to-text,

intent recognition, and text-to-speech. These systems are increas- ingly applied in fields like IT support, home automation, HR services, and voice-based search—paving the way for a voice-first digital future.

This paper surveys existing technologies and methods used in building AI-based voice assistants, focusing on implementations using the MERN stack (MongoDB, Express.js, React.js, Node.js) and external services such as Google Text-to-Speech (GTTS). This survey explores the development of such assistants, focusing on systems that support both voice and chatbot interfaces. It also highlights the role of the MERN stack in building scalable solutions and the importance of dashboard management for user control and transparency. The goal is to provide insights for developing Jarvis 2.0, an intelligent assistant with real-time, customizable interaction features.

Literature Survey

1. Artificial Intelligence-based Voice Assistant (Subhas S, Ullas A, Prajwal Srivatsa, 2020): Developed using Python, integrating gTTS for speech synthesis and PlaySound for audio playback. Performs web search, media playback, weather updates, screenshots, and emergency alerts via voice commands. Uses ASR and basic NLP. Desktop-based, not scalable, lacks web integration.

2. Adheetee: A Comprehensive Bangla Virtual Assistant (Islam et al., 2019): Bangla-language assistant for smartphones and PCs supporting voice and text com- mands, alarms, browser control, and email. Uses Google STT, NER, Cosine Similarity, and Google Translate API with 94.07% accuracy. Hybrid knowledge base + external API model. Limitations: dependency on external APIs, limited commands, scalability issues.

3. Unified Test Framework for Voice Enabled Devices (Chittaranjan Pradhan, Sunil A Kinange, Jayavel Kanniappan, 2020): Developed by Samsung R&D for end-to-end automation testing of AI-powered voice-enabled devices. Handles noise levels, accents, distances, and demographics. Compatible with Android, iOS, and Tizen.

4. The Rise of Conversational AI Platforms (Anna Maj, 2020): Conversational AI revolutionizes banking and payments through voice- and chatbot-based services. Enhances customer experience with instant, personalized support. Examples include DNB's Aino.

5. Next-Generation of Virtual Personal Assistants (Veton Kepuska, Gamal Bohouta, 2018): Advanced model integrating speech, gesture, image, video recog- nition, and body movement tracking for natural human-computer interaction. Tested on major AI services.

6. Virtual Assistants and Self-Driving Cars (Giuseppe Lugano, 2017): AI in autonomous vehicles helps with navigation, safety, and communication. Current systems limited; future AI should enhance safety and user interaction.

7. INTELLIBOT – Intelligent Voice Assisted Chatbot (2022): Integrates voice interaction, sentiment analysis, COVID-19 data visualization, and offensive text detection.

8. "JARVIS" - AI Voice Assistant (2023): Uses Speech Recognition, NLP, and TTS technologies for voice-driven tasks.

Methodology

The methodology for developing and evaluating Jarvis 2.0 follows a structured, layered approach inspired by existing AI-based voice assistant research.

Requirement Analysis

Objectives:

- Enable real-time, browser-native voice interaction with an intelligent assistant.
- Support voice-driven task execution including opening websites, retrieving live data, and controlling third-party services using APIs.
- Provide a dashboard that tracks assistant activity, stores user sessions, and allows configuration of assistant behavior.

Functional Requirements:

- Voice input/output using Web Speech API and TTS/STT engines.
- Natural Language Understanding (NLU) for intent recognition, entity extraction, and dialogue flow control.
- RESTful API support to fetch and manipulate external data.
- Context-aware conversation memory for multi-turn interactions.
- User management system for authentication, session handling, and privacy compliance.

System Design

Technology Stack (MERN):

- Frontend: React.js for dynamic interface and dashboard visualizations.
- Backend: Node.js + Express.js for API handling and microservices

communication.
- Database: MongoDB for user profiles, chat history, and system preferences.

System Modules:
- Voice Interface: Web Speech API for STT and TTS functionalities.
- Natural Language Understanding Engine: Dialogflow, spaCy, HuggingFace Transform- ers.
- Response Generator: Maps intent to responses and converts to audio.
- Task Manager: Executes system-level/external tasks via APIs and scripts.
- Dashboard Module: Visualizes user session history, logs, analytics, and settings.

Implementation Details

Frontend (React.js): Real-time interface with chat history and voice buttons, commu- nicates via REST APIs and WebSocket.

Backend (Node.js + Express.js): Handles user requests, command routing, AI module interaction, and interfaces with NLP microservices.

Database (MongoDB): Stores user credentials, preferences, chat history, contextual data, and admin logs.

AI Tools & APIs:
- STT: Google Cloud Speech-to-Text, Web Speech API
- TTS: Google Cloud TTS, Web Speech Synthesis API
- NLP/NLU: Dialogflow, spaCy, Rasa, HuggingFace Transformers
- Task Automation: JS triggers, shell commands, REST API calls

Evaluation

Usability Testing

Evaluated voice recognition accuracy, response relevance, and dashboard usability.

Performance Metrics
- Latency: Speech input to response delivery
- Intent Accuracy: Correct intent matches from NLP engine
- System Stability: Uptime and consistency under load

User Feedback

Collected through surveys/interviews for iterative improvements in dialogue flow, UI design, NLP accuracy, and dashboard functionality.

Future Enhancements
- Offline NLP/STT via on-device models (Whisper Tiny, Vosk, ONNX Runtime)
- Multilingual support using Indic NLP Library, AI4Bharat, HuggingFace models
- IoT integration via MQTT or REST APIs
- Contextual memory for multi-turn dialogues
- Cross-platform compatibility with React Native

Limitations
- Latency and real-time constraints with MERN stack
- Heavy dependency on third-party APIs
- No offline mode
- Security and privacy concerns
- Scalability issues without optimization

Research Gaps
- Lack of full-stack open-source voice assistants
- Underutilization of browser-based assistants
- Weak multi-turn contextual handling
- Limited language support for regional languages

Problem Statement

Most existing voice assistants are closed-source, platform-dependent, and rely heavily on external APIs. Jarvis 2.0 addresses this gap with a modular, MERN-based assistant supporting real-time voice communication, chatbot functionality, task execution, and dashboard management.

Conclusion

Jarvis 2.0 integrates AI-driven voice interaction, NLP, and dashboard management into a browser-based MERN stack system. It is open-source, platform-independent, and cus- tomizable. While relying on third-party APIs and lacking offline support, it provides a foundation for future enhancements like multilingual support and IoT integration.

References

S. S. Chowdury, A. Talukdar, A. Mahmud, and T. Rahman, (2018), ”Domain specific

intelligent personal assistant with bilingual voice command processing," IEEE.

E. V. Polyakov, M. S. Mazhanov, A. Y. Voskov, L. S. Kachalova, and S. V. Polyakov (2018), "Investigation and development of the intelligent voice assistant for the IoT using machine learning," in Moscow Workshop on Electronic Technologies.

K. Mahmood, T. Rana, and A. R. Raza (2018), "Singular adaptive multi-role intelli- gent personal assistant (SAM-IPA) for human-computer interaction," in International Conference on Open Source Systems and Technologies.

P. Vashishta, J. P. Singh, P. Jain, and J. Kumar (2019), "Raspberry Pi based voice- operated personal assistant," in International Conference on Electronics, Communica- tion and Aerospace Technology (ICECA).

L. Burbach, P. Halbach, N. Plettenberg, J. Nakyama, M. Ziefle, and A. C. Valdez (2019), "Ok Google, Hey Siri, Alexa: Acceptance relevance of virtual voice assistants," in IEEE International Communication Conference.

V. Kepuska and G. Bohota (2018),"Next generation of virtual assistant: Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home," in IEEE Conference.

Khairunisa Sharif & Bastian Tenbergen (2020), "Smart Home Voice Assistants: A Literature Survey of User Privacy and Security Vulnerabilities", Complex Systems Informatics and Modeling Quarterly.

Saavedra, Á., Chocarro, R., Cortiñas, M., & Rubio, N., "Impact of process and out- come quality on intention for continued use of voice assistants", Spanish Journal of Marketing ESIC, 2024.

Deepak Muralidharan, Joel Ruben Antony Moniz, Sida Gao, et al., "Noise Robust Named Entity Understanding for Voice Assistants", arXiv, 2020.

Jakob, D., "Voice Controlled Devices and Older Adults – A Systematic Literature Review", in Human Aspects of IT for the Aged Population – HCII 2022.

Runting Zhong, Mengyao Ma, Yutong Zhou, Qingxia Lin, Leiling Li & Nengjing Zhang, "User acceptance of smart home voice assistant: a comparison among younger, middle aged, and older adults", Universal Access in the Information Society, accepted in 2022 (published 2024).

Gadiparthy, H. S., Nair, M. M., Vani, V. & Shivani. (2022). INTELLIBOT – In-telligent Voice Assisted Chatbot with Sentiment Analysis, COVID Dashboard and Offensive Text Detection. Springer.

Priya Dalal, Tripti Sharma, Parth Gambhir.(2023). "JARVIS" - AI Voice Assistant. IJISRT.

Yida Chen, Aoyu Wu, Trevor DePodesta, Catherine Yeh, Kenneth Li, Nicholas Castillo Marin, Oam Patel, Jan Riecke, Shivam Raval, Olivia Seow, Martin Wattenberg, Fer- nanda Viégas.(2024). Designing a Dashboard for Transparency and Control of Con- versational AI. arXiv preprint arXiv:2406.07882.