iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VII **Month of publication:** July 2025

DOI: https://doi.org/10.22214/ijraset.2025.72820

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

Real-Time Voice-Based Web Form Interaction using Natural Language Processing

Carolina Pasham Frazer[1], G. Narasimham[2]
[1]*Post Graduate Student, M. Tech (Data Science),* [2]*Associate Professor, Department of Information Technology, Jawaharlal Nehru Technological University Hyderabad, UCESTH*

Abstract: ***This project implements an intelligent voice-driven form filling assistant using Streamlit, speech recognition, and natural language processing. The system enables users to interact with online forms through voice commands, dynamically extracting form fields from any URL and guiding users through a conversational interface to complete them. The Voice Based Form Filler system features multi-language translation support (English, Hindi & Telugu), speech synthesis for instructions, efficient form field extraction using Selenium, NLP-based context verification, and automatic form submission capability. It employs caching strategies, input verification algorithms, and field context recognition to maintain responsive performance while handling complex interactions. This solution makes web forms more accessible, especially for users with limited typing ability or who prefer voice interaction over traditional input methods.***

Keywords: Context Verification, Natural Language processing (NLP), Speech Recognition, Voice-driven form filling, Web Automation, Multilingual Support.

I. INTRODUCTION

Digital form completion represents a significant barrier to accessing online services for many populations, including the visually impaired, those with limited motor skills, elderly users, and individuals with limited digital literacy. While web forms remain the primary method for data collection across digital services, traditional keyboard and mouse interactions present challenges for large segments of the population.

Voice-driven interfaces offer a promising solution but have historically been limited to specific domains or simple commands. Our research project aims to bridge this gap by developing a comprehensive voice-driven form completion assistant that can navigate complex web forms in real-time, providing a seamless alternative input method.

A. Objective

To improve the accessibility and usability of web forms for individuals with disabilities and develop a real-time voice-based form-filling system that enhances the overall UX (User Experience) and efficiency of form-filling processes through intelligent NLP-based context verification and multilingual support.

II. LITERATURE SURVEY

The development of voice-based form filling systems has emerged as a critical area of research, particularly in addressing accessibility challenges faced by differently-abled individuals. This literature survey examines recent advancements in voice-assisted technologies designed to enhance digital inclusion and independence for users with various disabilities.

The work by Usharani et al. (2020) provides valuable insights into the practical challenges of implementing voice-based systems for visually challenged users in real-world environments such as banks and hospitals. Their research focuses on basic voice-to-text conversion coupled with mobile printing solutions, revealing both the potential and limitations of simpler approaches. The study identifies the need for additional hardware dependencies, such as OTG adaptors and printers, which can create barriers to widespread adoption. This research underscores the importance of developing comprehensive digital solutions that eliminate physical printing requirements while maintaining accessibility standards [1].

Mani et al. (2021) contribute significantly to understanding the technical foundations of speech-enabled form filling systems. Their research utilizes WaveRNN for speech synthesis and Hidden Markov Models (HMM) for speech recognition, targeting both visually impaired users and senior citizens. The study demonstrates the effectiveness of interactive voice prompts in guiding users through form completion processes. However, the research also highlights the limitations of traditional speech recognition technologies,

particularly in terms of accuracy and language support, suggesting the need for more advanced approaches using modern machine learning techniques and cloud-based APIs [2].

The research by Ramasubramanian et al. (2022) establishes the fundamental problem that existing form filling processes, particularly for UDID (Unique Disability ID) registration, create significant barriers for differently-abled individuals. Their work highlights how traditional systems force visually impaired users to rely on manual assistance, compromising privacy, dignity, and independence. The authors identify critical gaps in current systems, including the absence of accessibility features, voice-based interaction capabilities, and automated identity verification. Their proposed solution emphasizes the importance of intelligent, autonomous systems that can provide complete independence through voice-guided prompts and automated data processing with built-in security measures [3].

The research by Gaud et al. (2022) explores the broader context of conversational AI and natural language processing in creating personal assistant systems. While their work primarily focuses on chatbot development for general conversational services, it provides important insights into the theoretical foundations of NLP applications. The study emphasizes the potential for extending conversational AI beyond entertainment purposes to address practical accessibility challenges, particularly for users who cannot easily type or read digital forms due to various disabilities [4].

Building upon foundational work, Syed (2025) presents a comprehensive analysis of AI-powered multi-modal form filling systems that combine voice and image recognition technologies. This research demonstrates the proven effectiveness of AI-driven accessibility solutions in achieving enhanced processing efficiency and reduced error rates across different sectors. The study validates the use of advanced speech recognition with intelligent field validation and context-aware input correction, particularly beneficial for differently-abled individuals who currently depend on human assistance. However, the research also reveals the complexity inherent in multi-modal systems, suggesting that specialized voice-only approaches might offer superior reliability and simplicity for specific use cases [5].

III.METHODOLOGY OF THE PROPOSED SYSTEM

A. Proposed System

The proposed Voice-Based Form Filler system significantly outperforms existing solutions through its intelligent, comprehensive approach. Unlike basic browser extensions that only provide speech-to-text conversion, this system offers context-aware field validation and semantic understanding that automatically corrects inputs based on field types. While virtual assistants require pre-configured templates and raise privacy concerns through cloud processing, the proposed system dynamically parses any web form with local processing for enhanced security. This web-based solution provides superior accessibility features and multilingual support that platform-specific desktop tools lack. The system's real-time error correction, comprehensive performance analytics with WER tracking, and end-to-end automation from form parsing to submission verification represents a significant advancement over existing partial solutions that offer limited functionality without accuracy monitoring or complete workflow management.

B. System Architecture

The Voice-Based Form Filler system architecture can be explained as follows:

1) User Input Layer: Handles voice input and manual text entry through Streamlit interface
2) NLP and Context Verification Layer: Processes speech using spaCy for intelligent field validation and context-aware correction
3) Data Management Layer: Manages form data, session state, and Excel export functionality
4) Web Automation Layer: Selenium-based form parsing and submission to target websites
5) Response/Feedback Layer: Provides audio feedback via pyttsx3 and visual confirmation to users

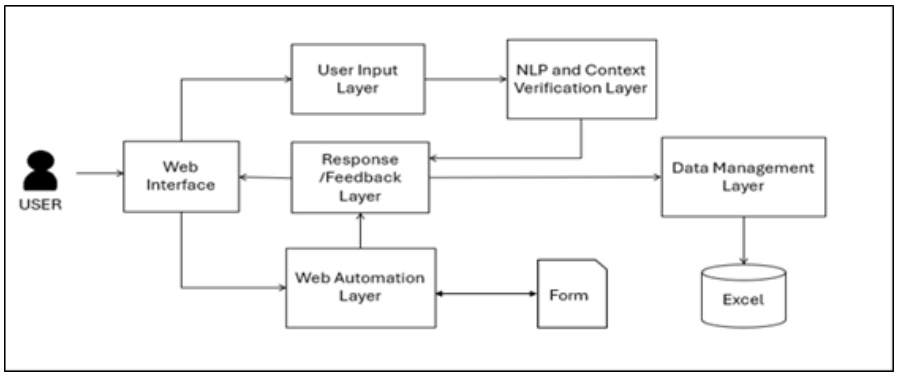


Fig 1: System Architecture of the Voice Based Form Filler System

C. Methodology

The methodology section outlines the plan and method that shows how the project is organized and implemented:

Form Structure Identification and Extraction

This module involves analysing the form to determine its layout, including the types and labels of fields. The system employs Selenium WebDriver to dynamically extract various form elements. By classifying fields and understanding their contexts, the system can effectively guide users in filling out the form.

Voice Interaction and Multilingual Support

In this module, a set of voice commands is developed to facilitate user interaction with the form. The system provides language selection (English, Telugu, Hindi) with translation capabilities for field labels. Using the Google Speech Recognition API, and text-to-speech synthesis with pyttsx3.

NLP-Based Context Verification

This module employs the spaCy NLP library to analyse field context and verify user input. The system categorizes form fields into semantic contexts (personal, contact, demographic, location, etc.) and applies context-specific verification rules. For example, it reformats spoken email addresses by converting "at" to "@", processes phone numbers by extracting only digits, and capitalizes names appropriately.

Intelligent Form Field Population

The system maps the transcribed and verified text to the corresponding form fields, automatically populating them based on user commands. For fields with options (radio buttons, checkboxes, select dropdowns), the system employs semantic similarity matching to find the closest match to the user's spoken input

Error Handling and Evaluation

This module focuses on detecting errors in user input and providing prompts for correction. By implementing word error rate metrics, the system verifies the accuracy of filled fields based on speech recognition, allowing users to make adjustments easily and ensuring data integrity.

Data Storage

The captured Form Field Data is stored in an excel sheet for record keeping.

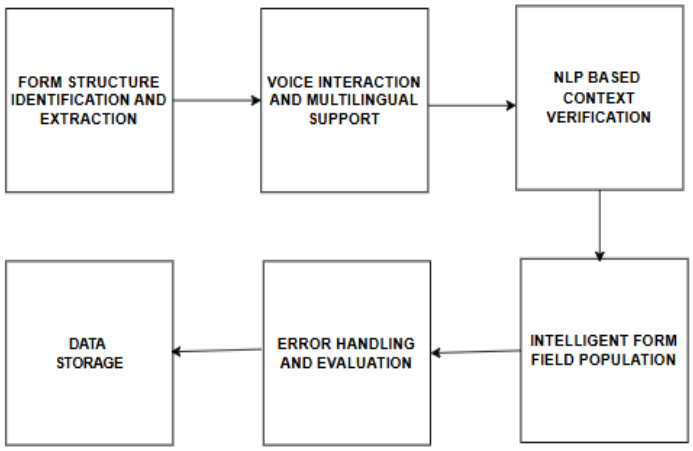


Fig 2: Block Diagram of the Voice Based Form Filler System

IV. IMPLEMENTATION AND RESULTS

A. *Key Features*

Features implemented are as follows:

1) Web Automation with Selenium
2) Speech Recognition and Translation
3) Natural Language processing for context verification
4) Streamlit Interface with State Management
5) Error Handling and Metrics Tracking
6) Data Storage

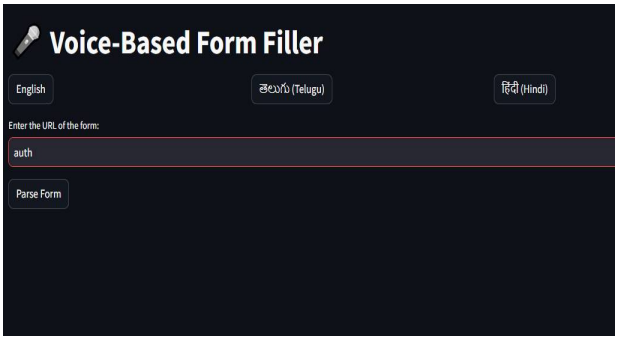


Fig 3: Main Display

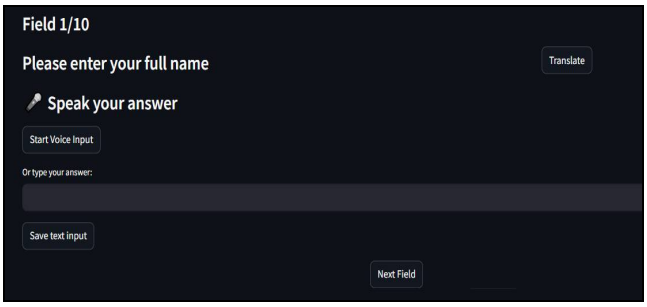


Fig 4: User Input Methods

B. *Results*

The voice-based form filling system demonstrates significant improvements in accessibility and user experience:

Context-Aware Input Verification: The system achieves high accuracy in field input through NLP-based verification, reducing errors commonly encountered in speech recognition.

Complete Voice-Driven Experience: The system not only recognizes speech input but also verbally communicates with users by speaking field labels and instructions aloud, making it truly accessible for visually impaired users.

Language Accessibility: Multilingual support (English, Telugu, Hindi) enables users to interact with web forms in their preferred language, with both visual and auditory translation.

Speech Recognition Metrics: Detailed tracking of Word Error Rate (WER) provides insights into system performance and user interaction patterns. The system performed quite well with a WER of 0% for 10 input attempts at every turn.

$$\mathbf{WER = (S+D+I)/N}$$

S= Substitutions, D= Deletions, I = Insertions

Here, we calculate it as manual overrides which replace unwanted substitutions, deletions or insertions:

$$WER = \frac{\text{Number of Manual Overrides}}{\text{Total Attempts}}$$

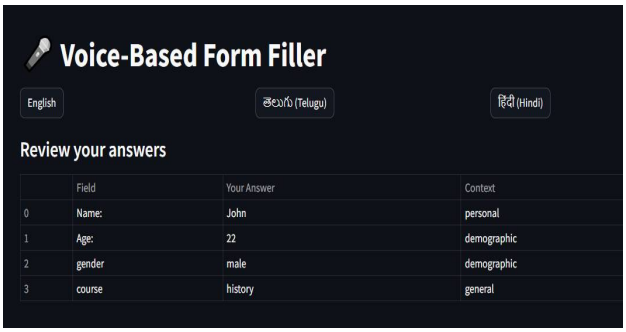


Fig 4: Review and Submit Page

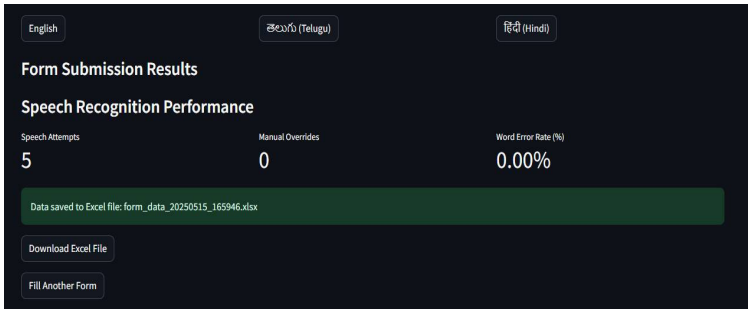


Fig 5: Form Submission Results

Field	Value	Context
Name:	John	personal
Age:	22	demographic
gender	male	demographic
course	history	general

Fig 6: Form Data saved in Excel

V. LIMITATIONS AND FUTURE SCOPE

While our system demonstrates promising results, several limitations suggest directions for future research. The current implementation struggles with dynamically generated content, custom controls, and JavaScript-heavy forms, limiting its applicability across modern web applications. The absence of persistent user data storage prevents automatic filling of repetitive fields like personal information, which would significantly enhance user efficiency. Enhanced context recognition through more sophisticated semantic extraction algorithms could improve field categorization accuracy and better understand complex form structures. Additionally, extended multilingual support beyond English, Telugu, and Hindi would broaden accessibility to diverse linguistic communities, requiring advanced NLP models that handle regional dialects and culturally specific form conventions.

VI. CONCLUSIONS

This research demonstrates the feasibility and value of voice-driven form completion as an accessibility solution for web interactions. Our implementation shows significant benefits for users with various accessibility needs, reducing completion time and error rates while improving overall satisfaction. The solution presented in this paper can help bridge the digital divide and ensure equitable access to web-based resources.

VII. ACKNOWLEDGMENT

I sincerely thank Sri G. Narasimham sir for his continual support in helping me carry out this project. Also extend my heartfelt gratitude to the authors of the research papers reviewed in this study.

REFERENCES

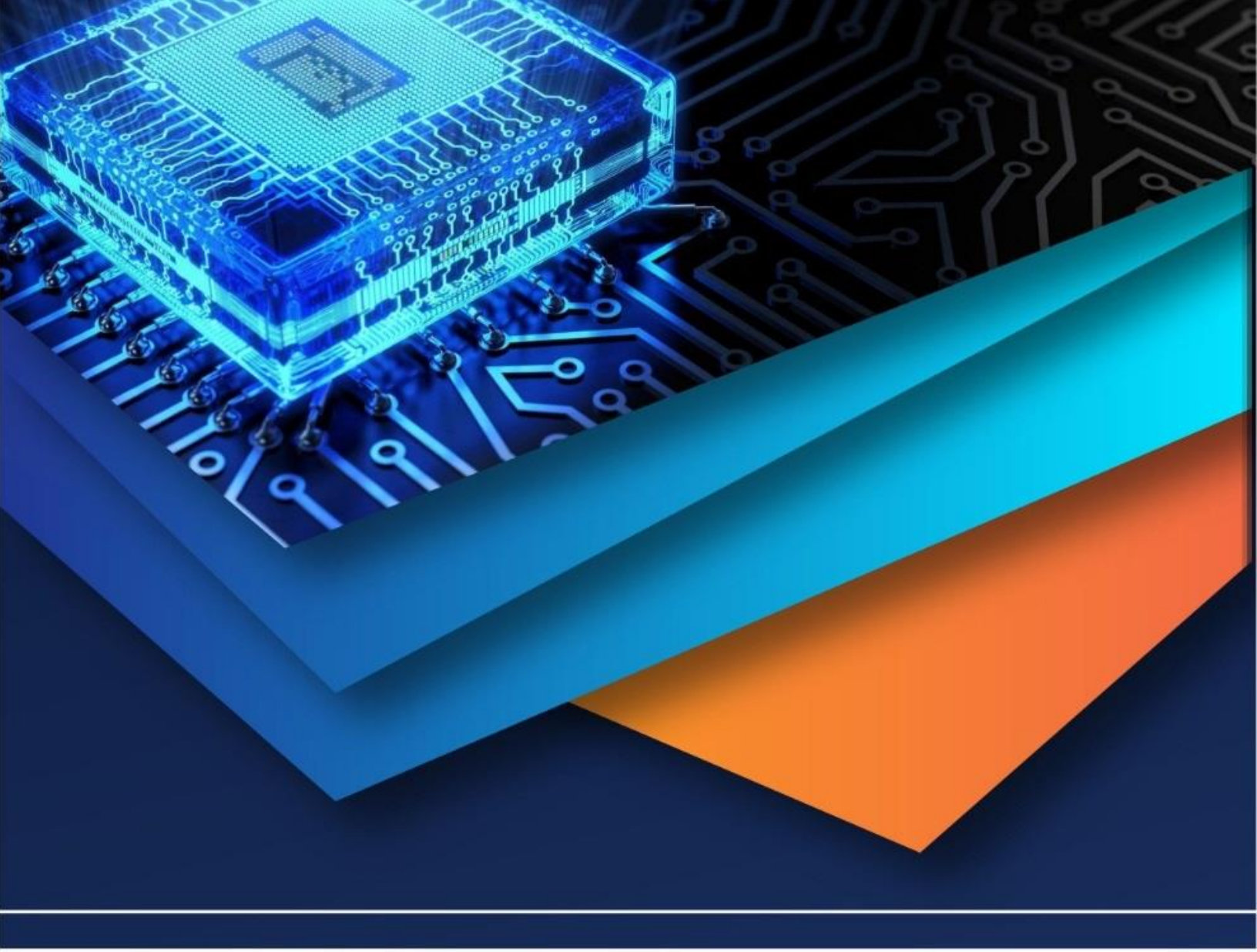
[1] S. Usharani, P. M. Bala and R. Balamurugan, "Voice Based Form Filling System For Visually Challenged People", 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2020.

[2] Mani, V.A., Dhanalakshmi, A., Dharani, S., Bharathi, B. (2021). Speech Enabled Automatic Form Filling System. In: Reddy, V.S., Prasad, V.K., Wang, J., Reddy, K.T.V. (eds) Soft Computing and Signal Processing. Advances in Intelligent Systems and Computing, vol 1325. Springer, Singapore.

[3] S. S. Ramasubramanian, S. Koodli, P. S. Nair, M. Sadique and H. Mamatha, "Voice Assisted Form Filling for the Differently Abled," 2022 International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Shivamogga, India, 2022.

[4] Asha Gaud, Bhumi Mota, Dhananjay Kumbhar, Veer Kumar, Prof. Shashank Gothankar, “Chatbot Personal Assistant Using Natural Language Processing (NLP)”, 2022, International Journal Of Innovative Research In Technology, MCT Rajiv Gandhi Institute of Technology, Mumbai, India. April 2022.

[5] Waseem Syed "AI-Powered Multi-Modal Form Filling: Advancing Accessibility through Voice and Image Recognition", 2025 International Journal of Scientific Research in Computer Science, Engineering and Information Technology, JNTU, Hyderabad, India.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY