International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 9, pp 19-25, September-2026

https://doi.org/10.47001/IRJIET/2026.109002

AI-driven Smart Form

[1]Matti Kiran, [2]P Kiran, [3]Ch. Shivani, [4]G. Sreenidhi

[1,2]Assistant Professor, Department of Computer Science and Engineering, G.Narayanamma Institute of Technology and Science for Women, Hyderabad, India – 500104

[3,4]Student, Dept of Computer Science and Engineering, G.Narayanamma Institute of Technology and Science for Women, Hyderabad, India – 500104

E-mail: [1]mkiran@gnits.ac.in, [2]pkiran@gnits.ac.in

***Abstract* - The digital world today has made online forms a crucial part of everyday life, from the registration of bank accounts to the application for insurance packages, accessing healthcare services, learning applications, and much more. But filling data in these forms is tedious, repetitive, and can be difficult, particularly for users who are not proficient in reading, typing, or filling out complex form structures. Smart Form – An AI-Powered Form Filling Assistant is the proposed project that would utilize state-of-the-art AI technology and voice commands to streamline and automate the process of filling out forms. The system can identify and extract form fields from uploaded PDF or image form and automatically fill the details known and stored in a pre-stored database, with intelligence. When there are unassigned or unknown fields in the Smart Form, Smart Form includes a voice assistant which reads every question or label for the fields and writes the user's answer in the corresponding field. Once the user completes the form, he/she is provided with a review interface to review, edit and confirm the answers he/she entered in the form and submit it. Smart Form's automation, AI and voice technology integration is an efficient, easy-to-use and accurate solution that increases accessibility and efficiency. This is particularly beneficial for individuals who struggle to see the screen or for older people or anyone who wishes to complete electronic forms without the need for hand typing.**

Keywords: Artificial Intelligence, Smart Form Filling, AI-Powered Form Assistant, Automated Form Filling, Voice Assistant, Voice-Based Form Filling, Intelligent Form Processing, Form Field Extraction, PDF Form Processing, Image Form Processing, Optical Character Recognition (OCR), Natural Language Processing (NLP), Speech Recognition.

I. INTRODUCTION

The online form has become a significant component in the digital environment in today's rapidly changing digital world. The online form has various applications in various sectors, including the banking sector, healthcare sector, education sector, and government sector. The online form is being used in different areas, but the form filling process is tedious, repetitive and time consuming. The online form has to be filled in the same way more than once, making it tedious and prone to mistakes.

In recent times, various advancements have been made in the field of artificial intelligence. This Includes computer vision, natural language processing, and speech recognition. These developments have paved the way for intelligent systems to gain insights into documents and communicate with users in a more natural and intuitive manner. In addition, voice-controlled technologies have been added. This has enabled users to input information using speech. This is a significant progress which has made digital technologies easier to use.

To overcome these challenges, this paper proposes a system named Smart Form - An AI-Powered Form Filling Assistant, a system that is capable of automatically filling digital forms given a PDF or image input. It can automatically identify the form fields and extract the necessary information, and then use voice assistance to input information that is missing. Smart Form aims to save time, improve accuracy and enhance the user experience by incorporating automation and interactive voice input. The way AI tools are changing the user experience of digital systems into an efficient process is an example of how such technology can be used.

II. LITERATURE SURVEY

In the last few years, tremendous progress has been made in the field of automated form processing and intelligent user interactions.

Researchers Carolina Pasham and Frazer *et al.* (2025) [1] have put forward a voice-based web form interaction system that uses speech recognition and text-to-speech technology to aid the user in filling out the form by voice commands. This system is highly effective for illiterate people since it reads out the form fields and converts the voice commands into text format. However, the system fails to provide the functionality

ISSN (online): 2581-3048
Volume 10, Issue 9, pp 19-25, September-2026
https://doi.org/10.47001/IRJIET/2026.109002

of storing the user information, which would otherwise increase the efficiency of the system in repetitive tasks.

Bobo Li *et al.* (2025) [2] proposed the intelligent web-based Form Factory platform, which facilitates the automation of form filling. The platform uses Optical Character Recognition (OCR), Natural Language Processing (NLP), and analysis of the form's structure to identify the input fields. Form Factory helps in the creation of digital forms by scanning the filled-out forms. Although the platform is efficient in automating the process, it lacks the functionality of voice assistance and the ability to handle multiple languages.

The multi-modal form filling system using voice recognition and image processing techniques has been proposed by Waseem Syed (2025) [3]. In this system, OCR is used for extracting text information from documents. Speech-to-text is used to input information. This is a more flexible and accessible system for filling forms. Although this system is effective, it does not allow real-time feedback and error correction while using speech input. This might cause errors in the input information.

It was found that in a study conducted by Deshmukh & Chalmeta in 2024 [4], Recognized the usability and effectiveness of Voice User Interfaces.

The study shows the significance of speech clarity, context, and user-centered design in improving the quality of interaction. However, the current Voice User Interfaces are still facing challenges in understanding different accents, emotions, and nuances, which might have a negative effect on the user experience.

Similarly, Bercaru and Popescu (2024) [5] have reviewed different techniques of accessibility that are implemented in online environments. In this context, the importance of multimodal interaction techniques like voice, visual aids, and accessible interfaces has been highlighted. Inconsistencies in different accessibility guidelines and a lack of proper implementation are identified as major challenges that are affecting the usability of the subsystems by differently abled individuals.

However, even though different systems are able to address different aspects like OCR-based information extraction, voice-based interaction, and accessibility, a comprehensive solution that can encompass all these aspects is still lacking. Moreover, different systems are not able to support user data storage and intelligent auto-filling; rather, manual input is still required.

To address these issues and create a comprehensive solution for intelligent and efficient information filling through a digital platform like Smart Form, this system has been designed by integrating different AI-based techniques like document understanding, intelligent data mapping, voice-based interaction, and user data storage.

III. PROPOSED SYSTEM

The system that has been proposed in the paper is named Smart Form, and it is a form-filling assistant that makes use of AI technology to automatically fill out digital and scanned forms by utilizing text extraction, intelligent automation, and voice interactions. When a user wants to fill out a form, the system processes the information that has been input by the user in the form of a scanned or digital image. For scanned images, the Gemini API has been used to extract the information, and PyPDF2 has been used for processing the information in the case of structured PDFs.

In order to further improve usability, a voice-assisted interaction module has been integrated into Smart Form, allowing for the completion of all remaining unfilled fields. For this purpose, Kokoro Text-to-Speech is used to read out prompts for each field, thus allowing the user to understand what is required without having to look at each field. The user's response is received in speech format, which is then converted to text format by using the Whisper Speech-to-Text model. These responses are then used to automatically fill in the respective fields, thus creating a seamless user experience. These processes are orchestrated by LangChain and LangGraph.

The frontend has been built using Streamlit. Streamlit is an open-source Python library for building web apps for machine learning and data science. The backend has been implemented built using Python. The Python code merges all the AI components. The execution of data processing and the data flow. Python takes care of these. Once done, users are given the option to review, edit and download the final form in text format. Smart Form, by incorporating artificial the introduction of intelligence and voice-driven automation has had a tremendous impact on the world of automation helped in reducing manual efforts and errors in form filling.

IV. METHODOLOGY

The proposed Smart Form system is based on a structured technique that involves artificial intelligence and document, Change the wording of sentences to enable automation of processing and voice interaction completion of forms. This system's workflow is overall Adds input acquisition and form field extraction, This is intelligent data mapping, interactive data collection, and more. This is intelligent data mapping, interactive data collection, and more. Validation and output generation. These steps are conceived to be accurate, efficient

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 10, Issue 9, pp 19-25, September-2026
https://doi.org/10.47001/IRJIET/2026.109002

and user interaction. The Smart Form system follows a well-defined and intelligent workflow – combining high-tech technologies to make it easier for the form to be filled. It begins with Collecting the input of users in documents, pictures or speech Search for form fields that are relevant to the command using automated techniques. The system then applies intelligent data mapping to match the user information with the required fields are filled in correctly. In cases of missing or it interacts with the audience by prompting them with interactions and questions that are not clear from the data. To collect the information that is needed. After collecting the data, validation mechanisms ensure correctness and Make sure that the form is complete before filling the final form. This systematic approach increases reliability, minimizes manual data entry and enhances quality.

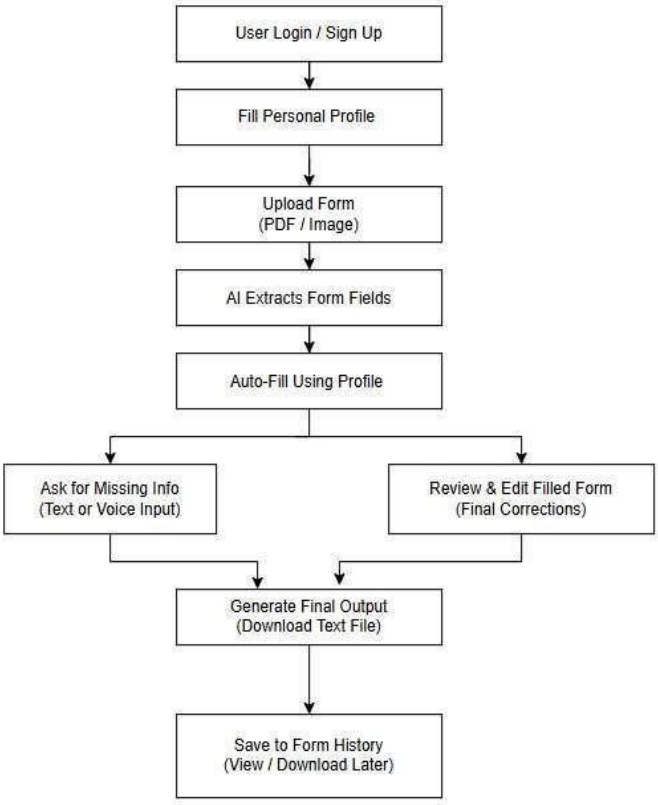

Figure 4.1: System Workflow

4.1 Input Acquisition and Preprocessing

The system can accept any type of a PDF form as long as it is in PDF format as images. The PDFs are converted into images using the pdf2image library to provide images in high resolution (200 DPI). Images are used as the standardized input for further analysis. In addition, the profile information provided by the user is stored in the MongoDB database, providing a centralized facility for the automation of the form filling process.

4.2 Form Field Detection and Extraction

The system accepts various forms in PDF format as well as images. The PDFs are converted into images using the pdf2image library to provide images in high resolution (200 DPI). Images are used as the standardized input for further analysis. In addition, the profile information provided by the user is stored in the MongoDB database, providing a centralized facility for the automation of the form filling process.

4.3 Intelligent Data Mapping and Auto-Filling

The extracted form field information is then matched with the existing information in the user's profile through natural language understanding. Variations in field names and semantic similarity are taken care of in the process. A confidence level is generated for each field based on the accuracy of the match. Fields with high confidence levels are automatically filled in, while uncertain fields are left for the user to fill in.

4.4 Interactive Data Collection using Voice and Text

In cases where the data cannot be automatically entered, the system makes use of an interactive question and answer process. In this process, the system provides prompts and presents them to the user in the form of text using text-to-speech functionality (KokoroTTS). Users can then provide the information in the form of text or speech. In the case of speech, the system makes use of the Faster Whisper speech recognition model to convert the speech into text. This information is then mapped to the fields, allowing the users to interact in a hands-free manner.

4.5 Data Validation and Error Handling

In the final step of the process, the system validates the correctness of the information entered in the various fields. This ensures the accuracy and consistency of the information entered in the various fields. In consistencies in the information are automatically corrected in this step.

4.6 Output Generation and Storage

Once validated, the completed form is displayed to the user in an editable format. Upon confirmation; a structured output is generated in text format. The completed forms are stored in MongoDB with timestamp values. This Enables History to be tracked downloadable copy of the user. Fill out their forms and be able to view past forms on demand.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 9, pp 19-25, September-2026

https://doi.org/10.47001/IRJIET/2026.109002

V. IMPLEMENTATION

The Smart Form system is developed on a modular and A design that is scalable and uses web technologies, AI systems, speech recognition and processing, cloud-based data storage, frameworks, and. The system has a multi-stage processing pipeline that allows every form is provided to a computer for computer-assisted completion of the form. The tools encompass user insight, analysis and intelligent data mapping input. It uses Streamlit for the front end and Pandas for the backend using Python for the backend. The system's overall workflow starts by authenticating and profiling and controlling users. A form is then uploaded in, In PDF or image format. Then, document analysis is done using artificial intelligence-based recognition techniques. Then, intelligent mapping is done between extracted information and user profile information. For information that cannot be automatically filled in, an interactive module is included with the ability for text and voice input. The filled out form is then checked for validity and shown to users for their review and saved in a database for future access.

5.1 Technologies Used

5.1.1 Frontend Framework: Streamlit

Streamlit is used for creating an interactive web interface for the application. Streamlit has native support for: Real-time file uploads, form input, audio recording, and live streaming updates. The ability of Streamlit to be reactive is beneficial for updating a user interface according to user action. The concept of session states is used to maintain the continuity at multiple points throughout the process. In addition, caching is used to enhance performance by removing redundant calculations.

5.1.2 Backend Language: Python

The language used for creating the backend is Python. For AI model integration and document processing libraries, as well as database operations, Python is utilized. User authentication is implemented with SHA-256 hashing, as well as interactions with the API and file uploading are done using Python. There are a lot of libraries available in Python to implement all the functionalities of the system efficiently.

5.1.3 AIModel: Gemini2.5FlashLite

The Gemini 2.5 Flash Lite AI model is used for intelligent analysis of the form and auto-filling of the data. This AI Use model to identify the form in text and images picks up fields and conducts structural analysis and semantic analysis Match-up of the user data.

5.1.4 Document Processing: pdf2image

The pdf2image library is used to convert the pdf documents into images with a resolution of 200 DPI. This ensures that the structure of the form is maintained and can be ready to be processed efficiently by the AI model.

5.1.5 Database: MongoDB

The MongoDB database is used to store user credentials, user information, and form history. MongoDB's data structure is best for storing hierarchical data, which in this application consists of form data and user information. The database is stored in MongoDB Atlas for secure and scalable cloud-based storage.

5.1.6 Speech Recognition: Faster Whisper

Faster Whisper is used for speech-to-text conversion. In this, speech is converted into text with high accuracy. It can also execute on both CPU and GPU. This text is directly mapped to corresponding form fields.

5.1.7 Text-to-Speech: KokoroTTS

KokoroTTS is used for text-to-speech conversion. In this, text is converted to speech. This is used to read out questions that are displayed on the forms. This is done to make the system more accessible. This way, users can interact with the system without using their hands.

5.1.8 Image Processing: Pillow (PIL)

Pillow is used for image preprocessing. In this, images are resized and compressed. This is done to reduce processing time. This does not compromise the accuracy of the image.

VI. RESULTS

In this section, the performance evaluation of the Smart Form system is discussed. In this case, the results of the experimental system are given. This highlights the efficiency of incorporating AI, document handling, and speech interaction for automatic form filling.

6.1 Performance Evaluation

The efficiency of the system was assessed in various ways, such as the accuracy of the information retrieved, the reliability of the auto-fill function, time taken to process, speech recognition accuracy and usability.

The AI-based form extraction component has proven to be highly accurate in the identification of the overall structure of the form, whether it was in the form of a PD For images. The system was also able to identify more than 90%of the

ISSN (online): 2581-3048
Volume 10, Issue 9, pp 19-25, September-2026
https://doi.org/10.47001/IRJIET/2026.109002

form fields, including text boxes, checkboxes, radio buttons, and drop-down menus. The intelligent auto-fill component has also proven to be highly accurate in the auto-filling of the form fields with the user's information, with the system achieving 85-90% accuracy in the auto-filling of the form fields with the user's information.

For high-confidence matches, it was possible to obtain a high degree of accuracy, reaching 95%. For uncertain fields, user input was required.

In terms of efficiency, it was seen that the system was able to execute all stages of processing quickly. For PDF conversion to images, it took about 1–2 seconds for each page, whereas for extracting information from forms, it took 3–5 seconds for single-page forms and 8–12 seconds for multi-page forms. For auto-fill, it took about 2–4 seconds, thus allowing a user to fill a typical form in 2 minutes, resulting in a saving of 70–80%.

The performance of the speech interaction module was impressive, as the Faster Whisper model resulted in 90-95% accuracy in quiet environments. However, the performance declined slightly in noisy environments. The audio output was clear and natural using the Kokoro TTS system, which improved the accessibility and interaction of the system. The MongoDB database ensured 100% data persistence, whereas the system performance remained stableevenafter20 concurrent users. The responses from the end-users reflected high satisfaction, as the system is intuitive and easy to use, as reflected by over 90% responses.

6.2 System Results and Interface Demonstration

This section describes the prominent interfaces and the output of the workflow developed in the Smart Form system.

6.2.1 User Authentication Interfaces

The registration interface and the login interface provide secure access to the system by using encrypted credentials.

Figure 6.1: User Registration Interface

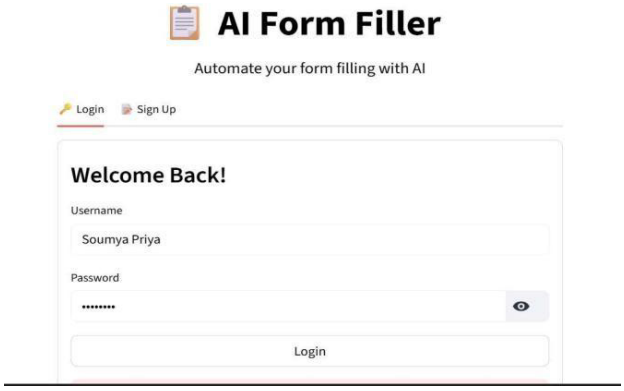

Figure 6.2: Login Interface

6.2.2 User Profile Management

The profile interface allows users to store personal and contact details, enabling automatic form filling.

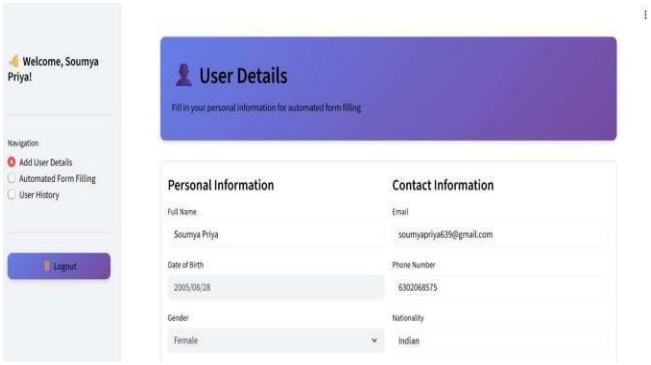

Figure 6.3: User Profile Management Interface

6.2.3 Form Upload and Processing

The system can be used to upload PDF and image forms for AI processing.

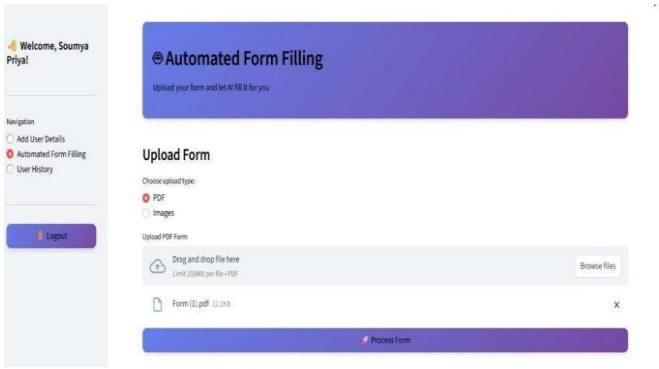

Figure 6.4: Form Upload Interface

ISSN (online): 2581-3048
Volume 10, Issue 9, pp 19-25, September-2026
https://doi.org/10.47001/IRJIET/2026.109002

Figure 6.5: Sample Form Document

6.2.4 Interactive Data Collection

The system is designed to gather information from missing data with both text and voice input methods.

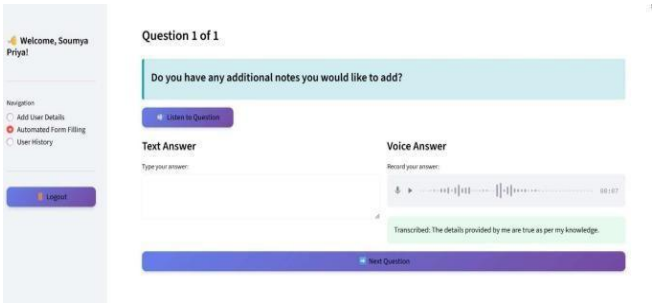

Figure 6.6: Interactive Question-Answer Interface

6.2.5 Form Review and Validation

Users can check and modify the data that is automatically entered.

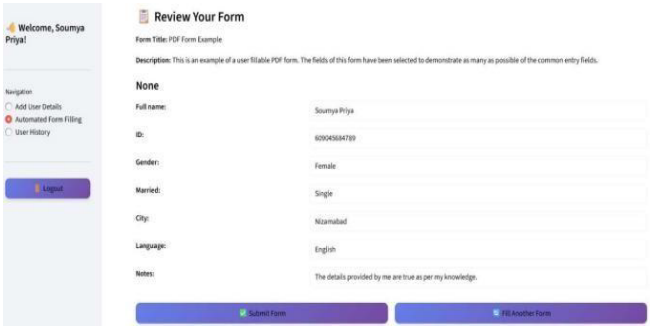

Figure 6.7: Form Review interface

6.2.6 Final Output and Storage

The system automatically produces downloadable output and saves the completed forms for future reference.

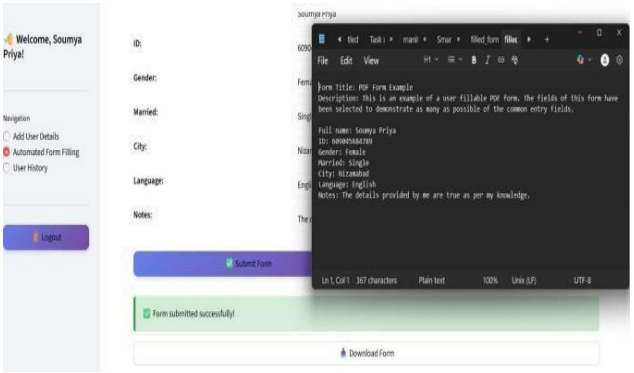

Figure 6.8: Form Submission and Download Interface

6.3 Comparative Analysis

Compared to the traditional and semi-automated form filling methods, the proposed Smart Form system has a great advantage. It typically takes 10-15 minutes manually to complete a form and there is repetitive information that can lead to high error rates (15-20%). The traditional auto fill systems of web browsers only auto complete 40-50% of the forms and do not have contextual understanding, validation, or voice interaction. The proposed Smart Form system can achieve a high level of automation with an accuracy of 85-90%, is multi-modal, and has intelligent validation, unlike the traditional form. Has a central profile-based system which minimizes repeat entry and persistent storage which makes previously completed forms easy to access.

VII. CONCLUSION

The AI Form Filler system is a good example of implementing advanced AI, NLP, and multimodal interaction, and delivering an automated form filling service quickly, precisely, and user-friendly. The application uses Streamlet's user interface and Python programming language to automate the fields in a form, intelligently match the last answer with the next field, generate questions from the context of the form, speak to the user, and validate the form. With the help of Google Gemini 2.5 Flash Lite, the accuracy of form field extraction can be improved to over 90% and the accuracy of intelligent form filling can be improved to 85-90%. The application also supports text and voice interaction, while minimizing repetitive tasks with profile management and form history. The form completion time can be reduced by 70-80% and the application will take less than 2 minutes for completion.

ISSN (online): 2581-3048
Volume 10, Issue 9, pp 19-25, September-2026
https://doi.org/10.47001/IRJIET/2026.109002

The current implementation has been tested and strengthened to be robust but there are opportunities for further strengthening with improved effectiveness and efficiency in the future. Examples of the possibilities encompass improvement of the OCR technology to be able to better process handwritten and un-structured forms, noise cancellation and multilingual speech recognition improvement. Plus, the use of digital signatures, direct form submission, mobile apps, and enterprise APIs will only further enhance the AI Form Filler in the future.

ACKNOWLEDGEMENTS

We are deeply grateful to Mr. M.Kiran, Assistant Professor in the CSE Department at GNITS, for his unwavering support, direction, and inspiration, which helped mold our project from beginning to end.

We would like to express our sincere gratitude to Dr. K. Ramesh Reddy, Principal of GNITS, providing us with the resources and environment needed to carry out our work smoothly.

We sincerely thank Dr. N. Kalyani, Dean of Innovation & Incubation, and Dr. M. Seetha, Dean of R&D, for the insightful advice and encouragement during our journey. We would especially like to thank Mrs. Sandhya Assistant Professor and Dr. Anil Kumar Assist. Professor, Dept. of CSE, Department of CSE, GNITS, project coordinators for their timely feedback, support, and direction at every turn.

REFERENCES

[1] Pasham, C., & Frazer, J. (2025). Real-Time Voice-Based Web Form Interaction using NLP. (Research study proposing a speech-driven form-filling system.).

[2] Li, B., *et al.* (2025). Form Factory: An Intelligent Platform for Automated Digital Form Creation and Completion. (Study introducing OCR + NLP-based automatic form detection and filling.).

[3] Syed, W. (2025). AI-Powered Multi-Modal Form Filling: Voice + Image Recognition. (Research integrating voice recognition and OCR for multimodal form automation.).

[4] Deshmukh, A., & Chalmeta, R. (2024). User Experience and Usability of Voice User Interfaces: A Systematic Literature Review. *Journal of Human–Computer Interaction Studies*, 12(3), 45–68. (Review on usability and accessibility of VUI systems.).

[5] Bercaru, V., & Popescu, N. (2024). A Systematic Review of Accessibility Techniques for Online Platforms: Current Trends and Challenges. *International Journal of Digital Accessibility*, 9(1), 1–22. (Review on accessibility tools and multimodal techniques for inclusive digital platforms.)

Citation of this Article:

Matti Kiran, P Kiran, Ch. Shivani, & G. Sreenidhi. (2026). AI-driven Smart Form. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(9), 19-25. Article DOI https://doi.org/10.47001/IRJIET/2026.109002

****

