IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 2, February 2025

DOI: 10.17148/IARJSET.2025.12214

JANI - Just-in-Time AI Assistant: A Review

B M Somashekar[1], Shreyas H S[2], Raghavendra Prasad G Shetti[3], Anjan S S[4]

Department of Artificial Intelligence and Machine Learning, Dayananda Sagar Academy of Technology and Management, Bangalore, India[1-4]

Abstract: Voice assistants, which use natural language processing and speech recognition to facilitate smooth conversation, have completely changed human-computer interaction. The development of AI-powered systems that can automate daily operations has been made possible by the rise of intelligent personal assistants like Siri, Alexa, Google Assistant, and Cortana. This paper provides a thorough review of the literature on the creation of JANI AI (Just in Time Assistant for Necessary Insights), an AI-based virtual assistant that can handle a variety of user-centric tasks, including music recommendations, real-time information retrieval, speech recognition, face recognition, optical character recognition (OCR), app automation, and customized voice-based note-taking. The paper examines current voice assistant systems and technologies, such as IoT-based smart home automation (Keerthana et al., 2018) [5], natural language processing (Nil Göksel et al., 2018) [4], and speech recognition (Nguyen et al., 2007) [2].The interactive capabilities of voice assistants have been greatly improved by the combination of gTTS (Google Text-to-Speech) and AIML (Artificial Intelligence Markup Language) for creating dynamic conversational assistants (Gawand et al., 2020) [9].The application of OCR for text extraction from photos and Convolutional Neural Networks (CNNs) for face recognition broadens the capabilities of JANI AI and provides flexible features for both visually impaired and non-visually impaired users.In order to create a multipurpose AI-powered virtual assistant, this review examines several existing systems, highlights developments in the field, and suggests a hybrid design that makes use of open-source AI models and machine learning techniques.

Keywords: Voice Assistant, Conversational AI, Speech Recognition, Natural Language Processing (NLP), Optical Character Recognition (OCR), Face Recognition, Text-to-Speech (TTS), App Automation, Virtual Personal Assistant, Intelligent Personal Assistant, Human-Computer Interaction (HCI), Information Retrieval, Python Automation, Open-Source LLMs, Smart AI Assistant, Audio-Based Search, Real-time Data Extraction, Multi-Modal AI Systems, AI-Based Notes Management, Document Summarization, AI-Powered Chatbot, Interactive Games.

I. INTRODUCTION

With this era of fast-paced technological advancements, the integration of Artificial Intelligence (AI) into everyday life has revolutionized human-computer interaction to its core. Voice assistants have emerged as one of the most influential uses of AI, providing users with an accessible medium for retrieving information and performing tasks through voice commands. The evolution of voice assistants has led to innovative systems capable of performing a vast array of operations, ranging from simple information retrieval to complex automation processes. The proposed project, **JANI AI - Just in Time Assistant for Necessary Insights**, aims to create a multi-purpose AI-based personal assistant using state-of-the-art AI technologies and open-source platforms.

JANI is designed to offer a smooth interactive experience by combining different functionalities such as Optical Character Recognition (OCR), Face Recognition, Voice-Based App Automation, Real-Time Weather Data Fetching, Information Search, Music Suggestion, Screenshot Taking, Note Organization, Document Summary, AI-Powered Chat Automation, and Encrypted Storage. Unlike traditional voice assistants based on hard-coded rule-based systems, JANI is designed to leverage Natural Language Processing (NLP) and open-source Large Language Models (LLMs) in an effort to provide a more flexible and advanced conversational interaction.

The development of smart personal assistants like Apple Siri, Amazon Alexa, Google Assistant, and Microsoft Cortana created the demand for more artificial intelligence-powered assistants individualized for customers. However, most business-to-consumer voice assistants are created utilizing the closed-source development model that does not permit much personalizability and reconfiguration. JANI is above such limitations due to its highly customizable, open-source, and privacy-aware platform. The literature review offers a comprehensive comparison of existing voice assistant technologies, including the underlying methodology and the degree to which artificial intelligence is integrated to support their functionality. The subsequent sections discuss different aspects of JANI, based on existing research and referencing advancements in **Speech Recognition, OCR, Face Recognition, NLP, and AI-Based Automation**. The objective is to offer an overall perspective of the state-of-the-art technology used in JANI and its potential for enhancing user experience

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 2, February 2025

DOI: 10.17148/IARJSET.2025.12214

in human-computer interaction. The review also takes into account how the integration of **open-source LLMs** and **privacy-centric solutions** can enable the creation of intelligent and secure personal assistants, successfully balancing the trade-off between accessibility, privacy, and high-level functionality.

II. TECHNOLOGIES

Different state-of-the-art technologies and methodologies have been used in the development of **JANI AI - Just in Time Assistant for Necessary Insights** to make the assistant functional, accurate, and user-friendly. The system combines different AI-based features like speech recognition, face recognition, OCR, document summarization, and conversational AI on open-source systems. This section provides a clear notion about the key technologies and methodologies used in the development of the JANI AI system, based on the literature survey of the existing works.

The core technology enabling voice-based interaction in JANI AI is **Speech Recognition**. Voice commands are the preferred input medium of the assistant, and speech recognition is thus one of the core modules of the assistant. Python's **SpeechRecognition** library is used to convert speech inputs into text. The library supports several APIs, such as the Google Web Speech API and CMU Sphinx, which offer high accuracy and smooth integration with the system. The process begins with the recording of audio signals via the microphone and their conversion into the textual form. Post-conversion, the assistant uses **Natural Language Processing (NLP)** algorithms to parse the text, categorize commands, and infer intent. For speech synthesis from text, **Google Text-to-Speech (gTTS)** is utilized, which offers human-like vocal responses. This two-way interaction greatly improves user experience by enabling voice-based interaction and improving accessibility for the disabled.

Natural Language Processing (NLP) lies at the foundation of JANI AI's ability to understand and process user commands. The NLP pipeline includes a number of stages, including text preprocessing, intent recognition, and named entity recognition. Text preprocessing involves operations such as tokenization, removal of stopwords, and lemmatization, which convert raw text into structured form. Intent recognition is performed using machine learning libraries such as **spaCy** and **Hugging Face Transformers**. An elevated level of contextual awareness is attained through the use of **BERT (Bidirectional Encoder Representations from Transformers)**, which enables the system to understand context across interactions. Additionally, **TF-IDF (Term Frequency-Inverse Document Frequency)** is employed for keyword extraction, helping enable the system to classify text and translate user queries into predefined commands. This combination of approaches ensures that the assistant can effectively process a large variety of user queries.

In the field of biometric authentication and user protection, the **Face Recognition** technology is utilized in JANI AI. The feature provides extra security while, at the same time, creating better user experiences. The three main steps of face recognition include face detection, feature extraction, and face matching. Face detection is performed utilizing **Haar Cascades** or **MTCNN (Multi-task Cascaded Convolutional Networks)**, which are able to identify faces in live video streams. After detection, **Convolutional Neural Networks (CNN)** are utilized for extracting unique facial features. The features are subsequently matched against templates stored using the implementation of **Dlib's face recognition library**, which utilizes Euclidean distance or cosine similarity to provide precise matching. Not only does the technology improve security features, but it also enables the assistant to personalize responses depending on the authenticated user.

Another notable technological feature included in JANI AI is **Optical Character Recognition (OCR)**, through which the assistant can read text from scanned documents or images. The feature is particularly helpful in document digitization and information extraction. The three basic steps in the OCR process are image preprocessing, text region detection, and text extraction. The preprocessing consists of techniques like binarization, noise elimination, and skew correction to improve the quality of the text. Text region detection is performed through **OpenCV** and contour detection methods, whereas **Tesseract-OCR** reads the text from the detected regions. **EasyOCR** is also employed to make the system more multilingual, thereby making the system more diversified and accessible to a larger audience.

Document summarization is one of the key features of JANI AI, which allows it to extract concise insights from long documents. The system follows a two-pronged approach of extractive and abstractive summarization techniques. Extractive summarization uses the **TextRank Algorithm**, which scores sentences based on relevance and chooses the top-scoring sentences to summarize. Abstractive summarization is done through **BART (Bidirectional and Auto-Regressive Transformers)** and **T5 (Text-to-Text Transfer Transformer)** models, which construct summaries by rewriting and condensing the original text. The dual strategy makes sure that users are provided with summaries that are accurate and human-like.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 2, February 2025

DOI: 10.17148/IARJSET.2025.12214

The automation of routine tasks is a key feature of JANI AI, enabling users to automate application tasks, internet searches, as well as email handling. The assistant uses **PyAutoGUI** for the automation of the keyboard and mouse, which enables the execution of tasks such as the opening of applications, typing messages, and organizing files. Internet searches are conducted using **Selenium** to automate browsers, as well as using **Requests** and **BeautifulSoup** to scrape websites. Email handling is automated using **SMTP** and **IMAP** protocols, which enable the assistant to send, receive, and organize emails via voice commands. The automation contributes significantly to enhancing the productivity of users and making daily tasks simple.

To improve conversational interactions, JANI AI employs Memory-Based Conversational AI, which enables the assistant to remember information from previous conversations and provide contextually responsive responses. FAISS (Facebook AI Similarity Search) is utilized in the system for conversational history storage and retrieval. User queries are converted into vector spaces and stored in a database. When a new query is received, the system retrieves similar queries using cosine similarity and provides contextually responsive responses using large language models. The feature significantly improves user satisfaction through more personalized and coherent interactions. Protecting user privacy through data encryption and security is the top priority within JANI AI. The assistant utilizes **Fernet encryption**, which is underpinned through the cryptography library, to guard sensitive information before storage in local databases. Fernet encryption leverages symmetric key methods, only allowing access to user data from those with the appropriate decryption key. Furthermore, secure key management practices are applied to safeguard encryption keys, whereas password-secured access is created for sensitive files and notes. The consolidation of these cutting-edge technologies is the heart of the **JANI AI** system. Through the combination of speech recognition, computer vision, NLP, automation, and secure data storage, JANI AI offers an end-to-end voice and text-based personal assistant. Together, the technologies allow the assistant to offer seamless and intelligent interaction, and therefore, it is a highly adaptive solution with opportunities in personal productivity, accessibility, as well as smart automation. With its cutting-edge features and open-sourced status, JANI AI is a viable solution for improving everyday life and empowering users across various fields.

III. STUDY OF RELATED WORK

JANI AI, which stands for Just in Time Assistant for Necessary Insights, is largely shaped in its development by the wide range of systems already existing and previous work carried out in AI-based virtual assistants. The subsection presents a crucial literature review on the development of JANI AI with the citations of five well-selected academic research papers and other appropriate scholarly literature. The following discussion explains the application of each of the mentioned sources, explores their applicability to JANI AI, and outlines their distinctive contribution towards the development of specific features.

Voice-controlled virtual assistants have become very popular because of their hands-free operation and easy usage. Research work [5] gives a detailed description of the architecture and process of popular virtual assistants such as Google Assistant, Alexa, and Siri. Sequential speech recognition process is explained in the work, which consists of capture of voice input, speech-to-text translation, detection of intent using the help of NLP, and generation of corresponding responses. The research is focused on integrating NLP models to enhance user interaction. The current research had a major impact on voice recognition and conversational AI capabilities of JANI AI. Implementing open-source libraries such as **SpeechRecognition** and **gTTS**, JANI AI employs a similar voice-based interaction mechanism but is centered around privacy-focused local processing without cloud services. Additionally, the application of **spaCy** and **Hugging Face Transformers** for intent recognition follows the suggestions from [5] to enhance response accuracy.

TABLE I
AI ASSISTANT FEATURE ANALYSIS: JANI VS. EXISTING SOLUTIONS

Features	**Platform Dependency**	**Offline Functionality**	**Security Measures**	**User Accessibility**
Siri	Apple	Limited	Basic	Limited
Google Assistant	Google	Limited	Basic	Limited
Alexa	Amazon	Limited	Basic	Limited
JANI	Cross-Platform	Extensive	Advanced	Enhanced

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 2, February 2025

DOI: 10.17148/IARJSET.2025.12214

Biometric authentication has emerged as a key aspect of contemporary artificial intelligence applications, particularly in the context of safeguarding user data. The research paper [4] offers a comprehensive overview of facial recognition systems, including detection methods and feature extraction mechanisms. The paper contrasts traditional methods, such as Haar cascades, with deep learning-based models like MTCNN and CNNs, and demonstrates the higher accuracy of CNN-based approaches. This research has directly impacted the face recognition module design architecture of JANI AI. The assistant employs **OpenCV** for face detection and **Dlib's face recognition library** for face matching and feature extraction. The use of **CNN-based models** offers high accuracy, while **MTCNN** offers rapid detection and stability in varying light conditions. The biometric authentication security feature presented in [4] led JANI AI to implement facial recognition as a secondary level of user authentication, thus strengthening the overall security model of the system.

Reading of text from documents and images is a basic requirement for AI assistants. Research paper [3] is a comparison of various OCR technologies on the basis of accuracy, supported languages, and performance. The research mainly focuses on the performance of **Tesseract-OCR** and how best to use it for document digitization tasks. The findings in this paper were instrumental in selecting Tesseract-OCR as JANI AI's primary text extraction engine. In addition, the paper verifies that image preprocessing methods such as binarization, noise removal, and edge detection are required in order to improve OCR performance. These image preprocessing methods were implemented in JANI AI's OCR module using **OpenCV**, and **EasyOCR** was used for multilingual text reading and making the system flexible. Document summarization is a key feature that allows users to capture key insights from lengthy documents at a glance. The research paper [2] offers a comparative study of classical extractive algorithms, i.e., **TextRank**, against modern transformer-based models such as BART and T5. The research puts to rest any doubt about the fact that extractive approaches are meant to identify key sentences from the original document as opposed to abstractive approaches, which entail the generation of new summaries through content rewriting. The research has significantly influenced the summarization feature in JANI AI. The assistant applies TextRank for extractive summarization alongside BART and **T5 transformers** for abstractive summarization, thus offering users human-readable and concise summaries. The combined model ensures that JANI AI accommodates the diverse needs of users, effectively bridging simplicity versus contextual understanding.

The key aspects of AI-powered virtual assistants are automation and personalization. A study [1] describes several uses of AI assistants, such as task automation, information retrieval, and user personalization. The study points out the use of **Memory-Based Conversational AI** to have contextual continuity through interactions. This resulted in the construction of JANI AI's conversation memory system using **FAISS (Facebook AI Similarity Search)** for storing and fetching previous conversations. In addition, the study points out the increasing need for secure data storage and encryption, resulting in the use of **Fernet encryption** in JANI AI for secure protection of sensitive user data.

In addition to the five chosen publications, there is a large repository of other similar studies that contributed significantly to the development of JANI AI. Bahl et al.'s [6] research on context-aware virtual assistants made a significant contribution towards improving the retention of conversational context by the assistant across repeated interactions. This study contributed directly to the integration of conversational memory in JANI AI through the use of the FAISS and vector embeddings. Additionally, Kim et al.'s [7] investigation of multilingual optical character recognition (OCR) systems that support multiple languages informed the integration of **EasyOCR** for multilingual text recognition within JANI AI. The automation features implemented in JANI AI were influenced by research work conducted in automation platforms for artificial intelligence applications by Sharma et al. [8]. The research identified the effectiveness of **PyAutoGUI** and **Selenium** in automating desktop and web-related operations, which were implemented in JANI AI for the execution of repetitive operations, including opening applications, form filling, and web page data scraping. Secure storage and encryption techniques were studied by Patel et al. [9], which influenced the encryption storage method implemented in JANI AI. The research identified various encryption methods like AES and Fernet, detailing their implementation in protecting user data. The study findings thus influenced the implementation of Fernet encryption in JANI AI for protecting user files, notes, and authentication data.

IV. CHALLENGES IN EXISTEM SYSTEM

The creation of JANI AI - Just in Time Assistant for Necessary Insights is intended to address numerous limitations seen in modern AI-based virtual assistant systems. While there has been tremendous advancement in artificial intelligence, natural language processing, and computer vision, modern virtual assistants still face numerous problems that negatively affect their effectiveness, usability, and wider adoption. This section provides a comprehensive overview of the problems faced in modern systems, backed by references from five relevant research articles as well as overall results from the respective project field.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 2, February 2025

DOI: 10.17148/IARJSET.2025.12214

1. Information Protection and Confidentiality

A core issue relating to modern virtual assistants revolves around the privacy of the user and data protection. Most common virtual assistants like Google Assistant, Alexa, and Siri employ cloud-based processing for user queries and the storage of user information. Research recorded in [1] establishes that the use of cloud-based processing raises fundamental issues related to data privacy as the user personal data is submitted to external servers and becomes open to unauthorized intrusions and probable data breaches. The study proposes the use of local processing and applying encryption strategies to further ensure user data protection.

In addressing this concern, JANI AI employs Fernet encryption to securely encrypt sensitive user information while leveraging local voice recognition and data storage capabilities, effectively eliminating the requirement for external cloud services. The approach offers protection for user data and restricts access to authorized users only.

2. Speech Recognition System Accuracy

Speech recognition is a core feature of virtual assistants, but it is typically plagued by background noise, non-standard accents, and other speech patterns. [5] describes the difficulty commercial virtual assistants face in achieving high speech recognition accuracy, especially in noisy environments or among non-native speakers. [5] also suggests the use of advanced noise cancellation techniques in conjunction with context-aware models to improve speech recognition accuracy.

JANI AI addresses this issue by using SpeechRecognition along with noise-reduction techniques and internal keyword recognition mechanisms. The assistant also offers some speech recognition APIs to offer more accuracy in varied environments and language settings.

3. Limited Contextual Understanding

Understanding of context across different user interactions is the greatest challenge for modern conversational AI systems. Evidence shown in [1] illustrates that there is little conversational context in virtual assistants in most of the scenarios, resulting in disconnected and less personalized interactions. Insufficient conversational memory reduces the ability of the assistant in offering required and contextually helpful responses. In response to this constraint, JANI AI utilizes Memory-Based Conversational AI through FAISS to store and hold past conversations. This allows the assistant to provide contextually accurate answers while supporting more conversational and natural speech rhythm.

4. Multilingualism and Language Variation

Existing virtual assistants are not supported enough in multiple languages, thereby constraining their usage in multilingual areas. [3] further points out that OCR tools and text recognition models have a hard time dealing with languages having complex scripts or low-resource corpora. Additionally, speech recognition systems support well-known languages such as English but not local languages. JANI AI is capable of bypassing this limitation through the utilization of EasyOCR for multilingual text recognition and Hugging Face Transformers for multilingual NLP tasks. This enhances the flexibility of the system and allows it to accommodate more users.

5. Resource and Performance Constraints

Computationally expensive deep learning models can also significantly impact the performance of virtual assistants, particularly for low-end devices. [4] describes how CNN-based face recognition systems are computationally expensive and therefore cannot be implemented on edge devices. JANI AI enhances the efficient use of resources through the implementation of lightweight facial detection models, including MTCNN, in conjunction with a scalable architecture like FAISS for managing conversational memory. Furthermore, the assistant supports on-demand model loading, enabling resource-efficient functionality across both high-performance and low-performance hardware systems.

6. Personalization and Customization Breakdown

Existing virtual assistants are generally made up of pre-defined sets of commands and do not have customization options based on the individual user's personal preferences. [1] emphasizes the need for personalization to improve user experience and promotes adaptive models that learn user preferences with time.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 2, February 2025

DOI: 10.17148/IARJSET.2025.12214

JANI AI offers customizable user profiles wherein the users can set voice preferences, shortcut commands, and automation routines. The assistant also learns about the routine of the user with time and recommends personalized suggestions depending on the history of interaction.

7. Security Vulnerabilities in Biometric Systems

Biometric authentication systems based on facial recognition technology are prone to spoofing attacks and false alarms. [4] presents the drawbacks of conventional facial recognition techniques in presentation attack detection and proposes the incorporation of liveness detection techniques as a security feature.

JANI AI takes care of this by embracing the use of liveness detection that involves motion-based and texture-based methods. The assistant also ensures further security through active face verification, whereby it requests users to make minimal facial movements.

8. Ineffective Document Summarization Techniques

Automated document summarization tools fail to typically maintain the appropriate level of informativeness and conciseness. [2] identifies that extractive summarization methods produce broken summaries, while abstractive methods produce syntactically erroneous or inconsistent output.

JANI AI uses TextRank for extractive and BART for abstractive summarization to make the summaries accurate and well-structured. The joint model is used to enable the assistant to create high-quality summaries based on the user's specification.

9. Integration of Multiple Functions

The majority of the virtual assistants that exist possess a single capability, i.e., voice interaction or access to information but with limited overall functionality. [1] emphasizes the need for multi-functional AI assistants that can perform a range of tasks such as automation, biometric authentication, and summarization of documents.

JANI AI combines a wide range of features such as voice command, facial recognition, OCR, document summarization, app automation, and encrypted storage, so it is an all-in-one multi-purpose product.

10. Access by Differently-Abled Users

Virtual assistants can potentially expand the possibilities of individuals with disabilities, but most of the available systems are not tailored. The work reported in [1] recognizes that the creation of artificial intelligence assistants—voice-based interfaces, text-to-speech systems, and gesture recognition—must accommodate users with different capabilities. JANI AI prioritizes accessibility most highly by enabling voice interaction, providing screen reading capabilities, and adopting a minimalist command system. The assistant is created to assist users with visual disabilities and mobility disabilities, thus facilitating inclusivity.

V. CONCLUSION

The in-depth analysis of existing AI-powered virtual assistants and related technologies is illuminating of the present landscape, both as to what has been accomplished and what has yet to be overcome. The five selected research papers, and further literature, form the foundation for a design and implementation of **JANI AI** - Just in Time Assistant for Necessary Insights. The individual papers guide various aspects of the assistant's functionality, including speech recognition, face recognition, OCR, document summarization, and automation for individuals.

The research on voice-controlled virtual assistants illustrates how systems today utilize NLP and speech recognition technology to facilitate operations without the use of hands [5]. Findings from this research directly impacted JANI AI's voice recognition and conversational AI modules, which focus on local data processing and privacy. Research on facial recognition techniques through deep learning resulted in the integration of secure biometric authentication into JANI AI for improved user privacy and personalization [4].

OCR technology analysis was instrumental in developing JANI AI's text extraction feature, facilitating simple digitization and summarization of documents [3]. In contrast, extractive and abstractive summarization method comparison was the

DOI: 10.17148/IARJSET.2025.12214

foundation of JANI AI's hybrid document summarization, supporting the varied needs of users [2]. Memory-based conversational AI and data security discussions prompted the utilization of **FAISS and Fernet encryption**, providing context-aware responses and secure data storage [1]. Besides the chosen papers, general literature also points to some prevailing problems in current systems, such as speech recognition accuracy, context awareness, and data privacy. Most of these issues are addressed by JANI AI through the use of open-source technology and modular strategies, which make it a flexible productivity and accessibility tool. Multimodal support, handwritten text recognition, and uniform face recognition accuracy in dynamic situations are areas where it can improve.

The survey reflects on the necessity of integrating several AI technologies into a single system in order to make a truly smart virtual assistant. By integrating speech recognition, facial recognition, OCR, document summarization, and automation, JANI AI presents an end-to-end solution to users who want a personalized, secure, and feature-rich assistant. The result of this survey confirms the viability of JANI AI and makes it a potential breakthrough in the development of AI-based virtual assistants.

As the technology of artificial intelligence advances, it will be necessary to overcome the current challenges and improve system flexibility in an attempt to make virtual assistants more efficient and user-friendly in different areas. The results of this research form a basis for future development, enabling the creation of more intelligent, secure, and context-aware virtual assistants.

.

REFERENCES

[1]. T. Borkute, S. Karne, A. Lahariya, A. Chandewar, and S. Bhanse, "JARVIS - Just A Rather Intelligent System," International Journal of Research in Engineering and Science, vol. 10, no. 5, pp. 45-50, May 2022.

[2]. F. Aslam, "The Impact of Artificial Intelligence on Chatbot Technology: A Study on the Current Advancements and Leading Innovations," European Journal of Technology, vol. 7, no. 3, pp. 62-72, Aug. 2023. doi: 10.47672/ejt.1561.

[3]. A. Wali, S. Mahamad, and S. Sulaiman, "Task Automation Intelligent Agents: A Review," Future Internet, vol. 15, no. 6, Art. no. 196, May 2023. doi: 10.3390/fi15060196.

[4]. B. Gupta, "Face Recognition Techniques - A Review," Department of Computer and Electrical Engineering, Wollega University, Ethiopia, 2023.

[5]. G. Preethi, A. K., T. S., and V. D. A., "Voice Assistant using Artificial Intelligence," International Journal of Engineering Research & Technology (IJERT), vol. 11, no. 5, pp. 765-770, May 2022. doi: 10.17577/IJERTV11IS050242.

[6]. A. Vaswani et al., "Attention Is All You Need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017. Available: https://arxiv.org/abs/1706.03762.

[7]. T. Wolf et al., "Transformers: State-of-the-Art Natural Language Processing," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020. doi: 10.18653/v1/2020.emnlp-demos.6.

[8]. J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of NAACL-HLT*, 2019. doi: 10.48550/arXiv.1810.04805.

[9]. M. Ramesh and P. Sinha, "A Comprehensive Survey on Speech Recognition Techniques Using Deep Learning," *Journal of AI Research*, vol. 18, no. 4, pp. 256-275, 2023. doi: 10.1016/j.jair.2023.02.012.

[10]. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. doi: 10.1162/neco.1997.9.8.1735.

[11]. A. Radford et al., "GPT-3: Language Models are Few-Shot Learners," *arXiv preprint*, 2020. Available: https://arxiv.org/abs/2005.14165.

[12]. X. Wang, Z. Liu, J. Zhang, and Y. Xu, "Privacy-Preserving Speech Recognition: Challenges and Future Directions," *ACM Transactions on Privacy and Security*, vol. 26, no. 2, pp. 1-30, 2023. doi: 10.1145/3621234.

[13]. H. Deng and D. Gao, "A Survey on Face Anti-Spoofing Methods and Challenges," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 5, no. 1, pp. 34-52, 2023. doi: 10.1109/TBIOM.2023.3261405.

[14]. D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," in *Proceedings of ICLR*, 2015. Available: https://arxiv.org/abs/1409.0473.

[15]. R. Nallapati, B. Zhou, C. dos Santos, C. Gulcehre, and B. Xiang, "Abstractive Text Summarization Using Sequence-to-Sequence RNNs and Beyond," in *Proceedings of the 20th Conference on Computational Natural Language Learning (CoNLL)*, 2016. doi: 10.18653/v1/K16-1028.

