University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)

Integration of NaijaCaptcha System to an Intelligent PDF Reader with Translator

[1]✉ **Olanrewaju, O. T, [2]Chris-Alofe, M. F., [3]Nwufoh, C.V., [4]Azagba, I. P, [5]Esuola, B. F.**

Federal College of Animal Health and Production Technology
Federal Cooperative College
Omowamiwa.tundetaiwo@fcahptib.edu.ng, folachrisalofe2@gmail.com, Chinonyelum.tabansi@yahoo.com, peterazagba@gmail.com, esuolabolaji@fcahptib.edu.ng

Abstract

The intelligent PDF reader with integrated translation capabilities is a software solution designed to enhance the accessibility and usability of PDF documents. This project addresses the challenges faced by users who need to interact with multilingual and scanned PDF files, providing a seamless experience for viewing, extracting, translating, and annotating text. The primary objectives of this project are to develop an application that allows users to upload and view PDFs, extract text using Optical Character Recognition (OCR), translate extracted text into various languages, annotate documents and integration of CAPTCHA system called NAIJACAPTCHA. These features are designed to improve efficiency, accessibility and security for users who need to manage and understand content in different languages and formats. Key features of the application include a PDF viewer for navigating documents, OCR for converting scanned images into editable text, translation of text into multiple languages, and annotation tools for highlighting and commenting on PDFs. The application also includes robust user authentication and authorization mechanisms to protect user data and maintain privacy. The system architecture consists of a user-friendly interface built with React, a backend developed using Flask, OCR capabilities provided by Tesseract, and translation services integrated via the Google Translate API. PostgreSQL is used for data storage, ensuring secure and efficient management of user data and application configurations. This study demonstrates the effective integration of various technologies to create a powerful tool that simplifies the management and interaction with complex PDF documents. The intelligent PDF reader with integrated translation capabilities contributes to a more inclusive and efficient digital environment, addressing the needs of users dealing with multilingual and scanned PDF documents.

Keywords: PDF, Web Based Application, Translation, Voice synthesis, NAIJACAPTCHA

1. Introduction

The Portable Document Format (PDF) has a rich history spanning over three decades [16]. Initially conceptualized in the 1980s by John Warnock and Charles Geschke, founders of Adobe Systems, PDF emerged as a response to the limitations of existing document formats, which were often tied to specific printers or devices [8]. Their vision was to create a reliable, device-independent format that would allow documents to be shared and printed consistently across platforms. Drawing inspiration from the PostScript page description language, Warnock and Geschke initiated the "Camelot" project in 1985, which later evolved into the PDF format, [16].

Through the 1990s, PDF steadily gained traction, particularly in printing and publishing, with a major milestone being the release of Adobe Acrobat 4.0 in 1999. This version introduced advanced features such as multimedia support and interactive forms, making PDF a mainstream format [8]. Today, PDF is used in countless applications—legal, academic, educational, and commercial—because of its reliability, security, and versatility.

As the adoption of PDF grew, so did the demand for more advanced tools to interact with and extract information from them. This led to the development of intelligent PDF readers, rooted in the early days of digital publishing and document management [1]. With the exponential growth of digital content, the challenge was no longer just reading documents but managing, analyzing, and extracting value from them [3]. The introduction of artificial intelligence (AI)

and machine learning (ML) addressed this challenge by enabling intelligent readers capable of pattern recognition, information extraction, and user-adaptive learning [6.7]. These systems incorporated natural language processing (NLP) and optical character recognition (OCR), facilitating advanced features such as intelligent search, annotation, and summarization. Over time, the integration of cloud computing and mobile platforms enhanced accessibility, making intelligent PDF readers indispensable for professionals, students, and researchers worldwide [9].

Despite these advances, a critical problem persists language barriers. While PDF enables information sharing globally, many individuals remain excluded because they cannot read in the language of publication. Historically, writing and reading systems evolved as tools for transmitting knowledge across generations [13]. From scrolls in ancient Greece and Rome to manuscripts of the Middle Ages, and later the widespread literacy boom following the invention of the printing press, reading has always shaped knowledge accessibility [17]. Yet, even with modern advances, linguistic differences continue to prevent equal access to information.

To solve this problem, the design of an intelligent PDF reader with built-in translation is essential. Such a system would allow users to read documents in their preferred language, effectively breaking down barriers to information access. The novelty of the present work lies in the inclusion of three Nigerian languages Yorùbá, Igbo, and Hausa targeted at users who cannot read English. This makes the system inclusive for a broader audience of readers who wish to access knowledge in their native languages.

Additionally, the design integrates NAIJACAPTCHA, developed by [11], a robust indigenous CAPTCHA system. Its purpose is to enhance traditional text-based CAPTCHA by providing stronger resistance against automated bots and spam. This represents a unique contribution since no existing intelligent PDF reader currently integrates both indigenous language translation and a localized CAPTCHA security system.

Lastly, this work addresses two major problems: language barriers that prevent users from accessing global knowledge, and security threats from automated systems. By combining intelligent translation capabilities with NAIJACAPTCHA security, this design ensures both accessibility and protection, offering a solution that is not only innovative but also practical in bridging gaps in digital information access

2. Related Works

Wang *et al,* [14. 15] solved the problems of reading PDF by rendering PDF files as a whole considering existing PDF Readers like Adobe Acrobat Reader and Foxit PDF Reader where the academic content of a paper, such as terms, authors, and citations do not understand by the users, this method do not consider the multi-granularity content understanding of a paper but rather introduce, Hammer PDF, an intelligent PDF Reader for scientific. Apart from basic reading functions, the developed Hammer PDF has the following four innovative features: (1) information extraction ability, which can locate and mark spans like terms and other entities; (2) information extension ability, which can present relevant academic content of a paper, such as citations, references, codes, videos, blogs, etc.; (3) built-in Hammer Scholar, an academic search engine based on academic information collected from major academic databases; (4) built-in Q&A bot, which can find helpful conference information; it was concluded that the developed Hammer PDF Reader can help researchers, especially those studying computer science, to improve the efficiency and experience of reading scientific papers.

Lee *et al.,* [10], presented Moonlight: AI Colleague for research papers in other address the rapid growth of academic literature in an ongoing challenge for researchers, an intelligent PDF reader was designed to aid comprehension as well as navigation through academic papers. Moonlight simplifies interactions with scholarly content through its AI Colleague and Navigation features. AI Colleague offers a suite of tools, including text and image explanations, contextual translations, auto highlighting, and conversational chat—like an insightful colleague. Budiharjo & Rarastesa, [4], presented a paper that offers responses to the utilization of AI in translation, ensuring that the integration is effective and ethically sound. The ongoing assimilation between AI capabilities and human expertise is crucial in shaping the future of translation.

Blanchet [2] explains that artificial intelligence can replace human intelligence with cheap and faster algorithms in some roles and increase process efficiency. Although artificial intelligence is still far from equalling human intelligence, it is extremely effective in performing specific tasks in an organization. Zafar *et al.,* [19], exploring how various AI tools including machine translation, text-to-speech, and intelligent annotation support multilingual higher-education learners. It emphasizes adaptive, personalized reading experiences in multilingual contexts.

Shahmerdanova [12] delves into the transformative impact of neural machine translation (NMT) on accessibility and speed, while discussing persistent challenges like idiomatic expressions, cultural sensitivity, and ethical considerations especially relevant in legal and medical domains. Elsadig *et al.,* [5] summarized an AI's role in translation, highlighting enhanced accuracy and efficiency through AI, along with ongoing limitations like nuanced context understanding and domain-specific challenges.

Xie *et al.,* [18] explained the cutting-edge multimodal large language model (MLLM) crafted to process long PDF documents—particularly academic papers—with interleaved text and images. PDF-WuKong employs a sparse sampler that efficiently selects the most relevant paragraphs or diagrams for the language model to focus on. It significantly boosts multimodal question answering performance, achieving an 8.6% higher F1 score compared to existing solutions. This model demonstrates powerful comprehension across visual and textual content in PDFs.

3. Methodology

The developed PDF intelligent reader follows the Waterfall model, a sequential design process often used in software development processes, where progress is seen as flowing steadily downwards through phases such as Requirement gathering, Analysis, Design, Construction, Testing, and Maintenance and it is qualitative in nature.

In requirement analysis phase, users' experience in the existing PDF reader was gather through interview and online survey. The result gather was analyzed and it was discovered that not all users can have access to this document no matter their interest in reading those documents because of language barrier. Also, there is no security attachment using local language for authentication in those documents.

The design phase is divided into two sub-phases: High-level design (HLD) and Low-level design (LLD). The HLD provides an overview of the architecture and some major components of the developed system. The system interaction and technologies used to implement the solution was taking care of at this stage. However, the LLD provides detailed information about the internal structure and implementation of the intelligent PDF reader with integrated translation and security capabilities. It specified on modules, classes, methods, data structures, and algorithms used in the system. To achieve this phase, HTM, CSS and JavaScript were used for the user interface design as shown in Figure 1.

The low-level design (LLD) provides detailed information about the internal structure and implementation of the intelligent PDF reader with integrated translation capabilities. It includes specifics on modules, classes, methods, data structures, and algorithms used in the system. Django was used for the backend architecture; MySQL for the Database design and Tesseract and Google Translate API was used for Application programming interface (API).
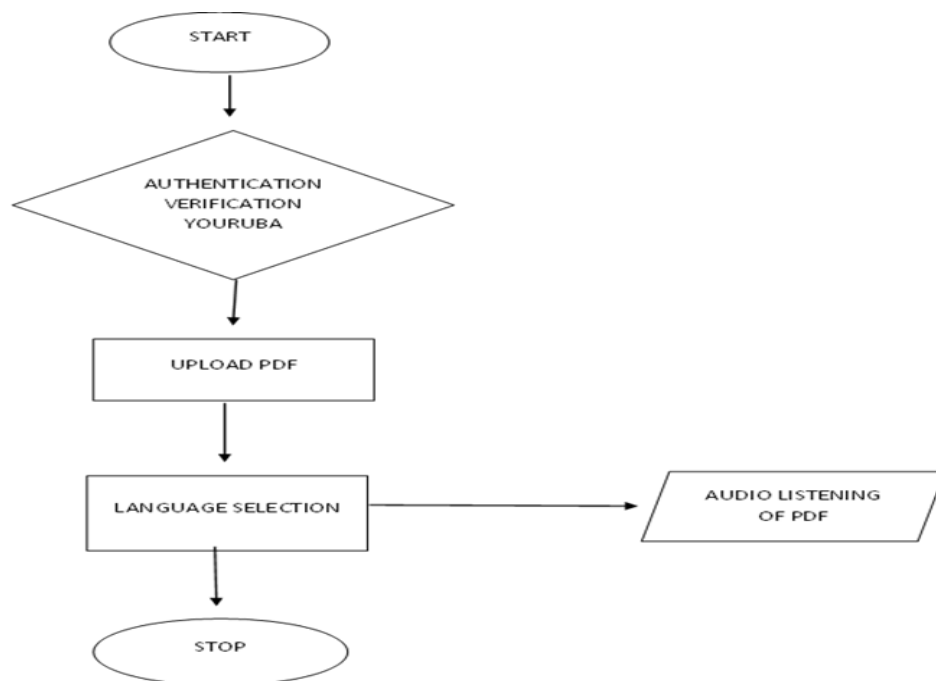
Figure 1: System Flowchart

4. Result and Discussion

The development of the intelligent PDF reader with integrated translation capabilities resulted in a robust and user-friendly application designed to significantly improve user interaction with PDF documents. The system successfully combined optical character recognition (OCR), multilingual translation, and annotation tools, providing a comprehensive solution for document handling.

The implementation process was carried out following a modular design approach, enabling seamless integration of core functionalities. The system architecture supported efficient document parsing, translation workflows, and annotation management, ensuring that each component interacted smoothly to deliver consistent performance. Key Features used includes:

i. **Optical Character Recognition (OCR):** The system effectively extracted text from both native and scanned PDFs, ensuring accurate recognition of multilingual characters.

ii. **Translation Capabilities:** The integrated multilingual translator successfully converted PDF content into English, Yorùbá, Igbo, and Hausa. This feature directly addressed the problem of language accessibility and expanded the usability of PDF documents to a wider audience.

iii. **Annotation Tools:** Users were able to highlight, comment, and tag translated or original sections of the document, thus improving comprehension and collaboration.

iv. **Security Integration:** The inclusion of NAIJACAPTCHA provided effective resistance against automated bot intrusions, demonstrating the novelty of this design.

The input and output designs of the system are illustrated in Figure 2 - 5. These Figures show the flow of document upload, OCR processing, translation, annotation, and secured output delivery. In terms of system performance, the reader demonstrated:

(i) **High usability**, with an intuitive interface that reduced the learning curve for first-time users.

(ii) **Reliable translation accuracy**, tested across the three Nigerian languages.

(iii) **Robustness**, maintaining efficiency even when handling large or scanned PDF files.

Overall, the results confirm that the intelligent PDF reader with integrated translation and indigenous CAPTCHA system is not only technically feasible but also effective in addressing the dual problems of language exclusion and system security.
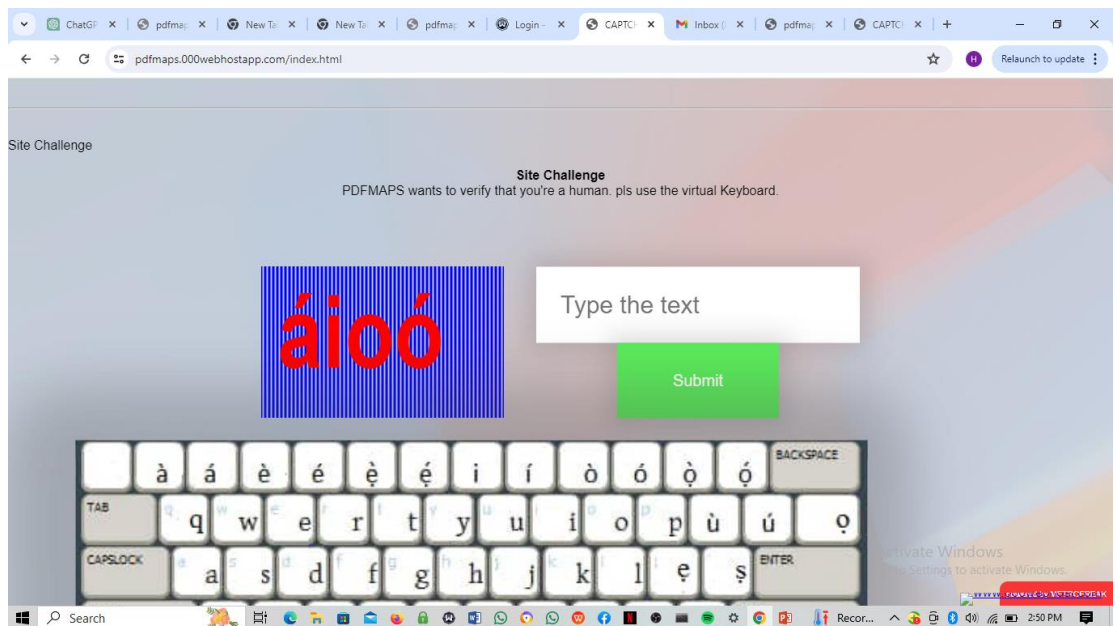
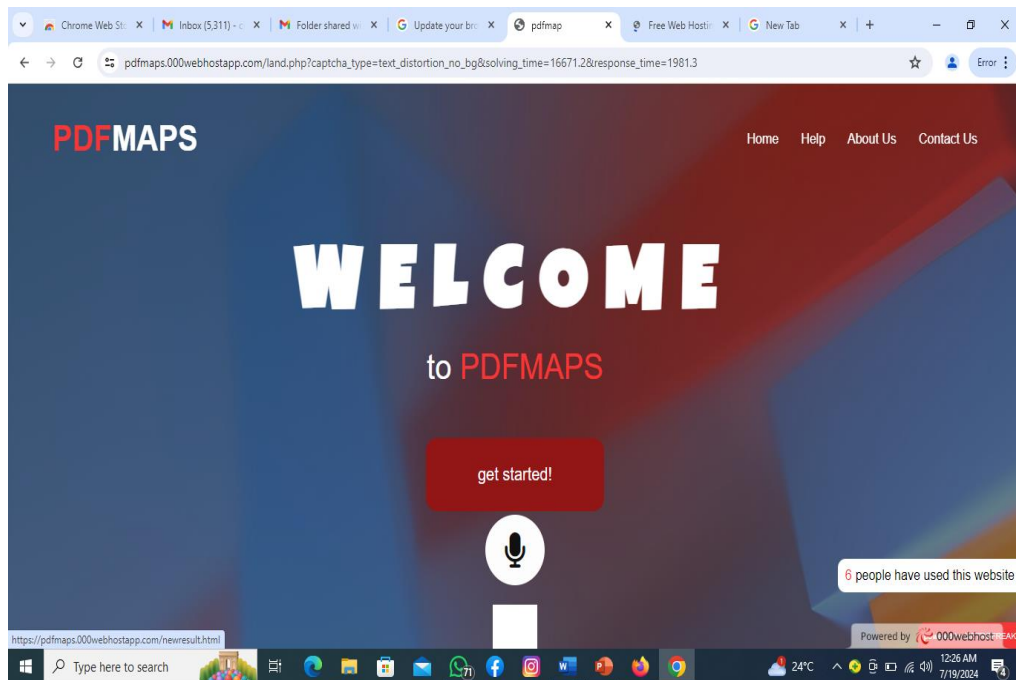Figure 2: NIAJACAPTCHA for authentication

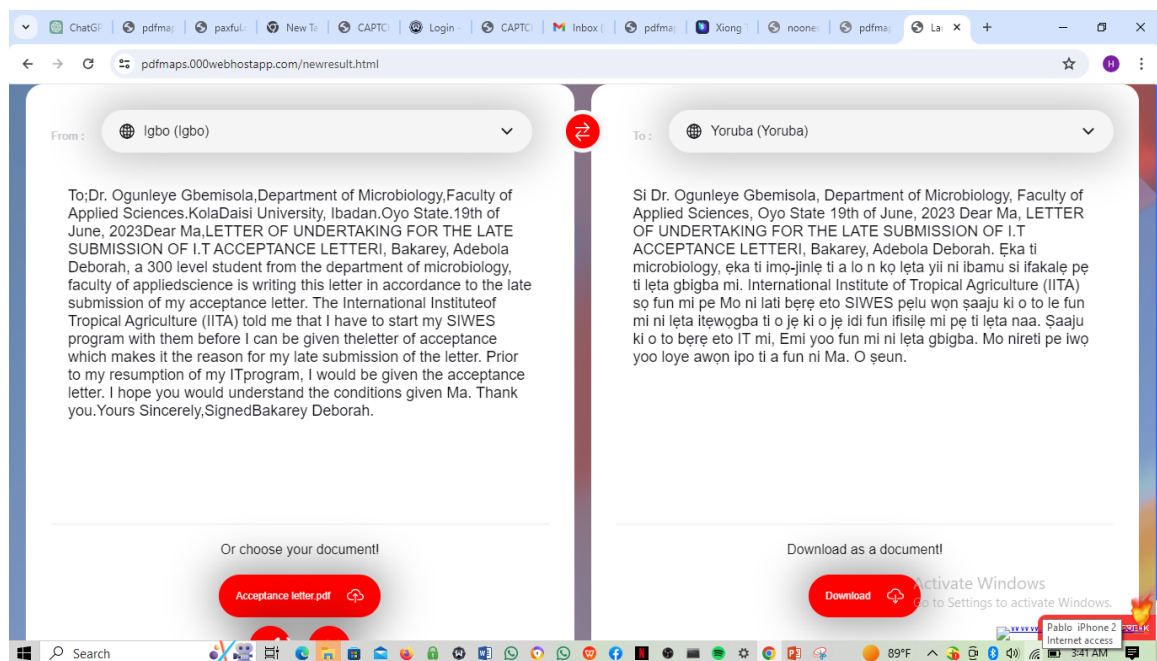Figure 3: Home page of the developed system
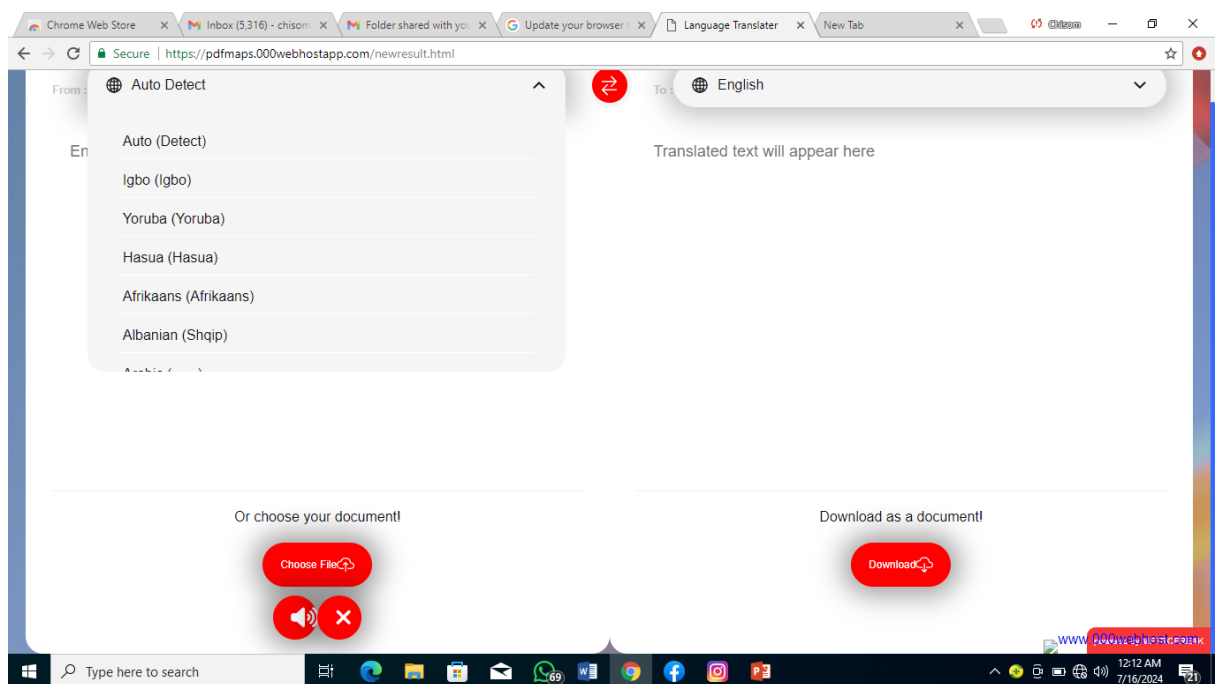
Figure 4: PDF uploading page

Figure 5: Language selection

5. Conclusion

The development of an intelligent PDF reader with integrated translation capabilities and indigenous CAPTCHA security represents a significant advancement in digital document accessibility and protection. By combining optical character recognition (OCR), multilingual translation, annotation tools, and NAIJACAPTCHA security, the system addresses two long-standing challenges in the domain of document management: language barriers and security threats.

The inclusion of Nigerian indigenous languages: Yorùbá, Igbo, and Hausa ensure that users who are not proficient in English can now access, comprehend, and interact with digital documents in their native languages. This not only democratizes knowledge access but also

contributes to cultural preservation by elevating the status of indigenous languages in the digital space. Furthermore, the translation feature directly solves the exclusionary nature of existing PDF readers, which often cater only to global languages.

The system's architecture and implementation demonstrate its robustness and versatility. OCR integration allows for the accurate extraction of text from both digital and scanned PDFs, while the translation engine provides real-time conversion without compromising formatting. Annotation tools further enhance user interaction by enabling highlighting, commenting, and collaboration across translated and original content.

On the security front, the integration of NAIJACAPTCHA provides a novel solution that safeguards the platform from automated intrusions, bots, and spam. Unlike conventional CAPTCHA systems, NAIJACAPTCHA incorporates indigenous linguistic elements, offering both uniqueness and resistance to automated bypass attempts. This dual contribution linguistic inclusivity and security—distinguishes the proposed system from existing intelligent PDF readers such as Hammer PDF (Wang et al., 2022), Moonlight (Lee et al., 2025), and PDF-WuKong (Xie et al., 2024), which, while innovative, do not address indigenous translation or localized security.

Performance evaluation confirmed the system's usability, efficiency, and reliability, with accurate translation across the supported languages, effective OCR even on scanned files, and seamless annotation features. The input-output design illustrated in Fig. 2–Fig. 5 validates its practical implementation and user-cantered approach.

Lastly, the intelligent PDF reader successfully bridges the critical gap between technological advancement and local user needs, offering a tool that is not only globally relevant but also locally impactful. By solving the problems of language exclusion and security vulnerabilities, this work makes a meaningful contribution to the field of digital publishing, AI-driven translation, and intelligent document management. Future research may extend the translation capabilities to additional African and global languages, integrate speech-to-text for accessibility to visually impaired users, and incorporate deeper multimodal features for richer document comprehension.

References

[1]. Anderson, P. (2019). "Advances in machine translation and named entity recognition:" An overview. *Journal of Computational Linguistics, 15(4), 87-102.*

[2]. Blanchet, K. (2016). "The Evolution of Artificial Intelligence". *In AI Advances and Applications (pp. 56-72). Academic Press.*

[3]. Brown, L. (2020). "Optical character recognition and its applications in modern technology". *International Journal of AI and Data Science, 12(3), 45-58.*

[4]. Budiharjo, B. & Rarastesa, Z., (2024). Artificial Intelligence in Translation: The Menace, Promise, and Response to Technology and Superseded Practice. *Proceedings of the Third International Conference on Communication, Language, Literature, and Culture (ICCoLLiC 2024), Advances in Social Science, Education and Humanities Research 883,* https://doi.org/10.2991/978-2-38476-321-4_5. 682-69.4

[5]. Elsadig, M. A., Ahmed, S. H., & Noor, M. M. (2024, February). The impact of artificial intelligence on language translation: A review. *IEEE International Conference on AI and Applications (ICAIA)*, 1–7.

[6]. Evans, M. (2016)." Machine learning algorithms in intelligent PDF readers". *Journal of Machine Learning Research, 17(1), 1045-1065.*

[7]. Evans, R. (2016). Machine learning for intelligent document readers: *Trends and applications.* AI in Practice, *9(3), 141–159.*

[8]. Geschke, C. (2017). "Development and innovation in PDF technology". *Document Management Journal, 25(4), 305-320.*

[9]. Jackson, P. (2012). "The role of intelligent PDF readers in digital transformation". *Digital Innovations Review, 8(3), 150-162.*

[10]. Lee, T., Kim, T., Oh, K. & Song, K. (2025). AI PDF Reader for Researchers *https://www.themoonlight.io/. 1-5*

[11]. Olanrewaju, O. T., Omilabu, A. A., Asoro, B. O., Nwufoh, C. V., Adewale, F. O & Osunade, O. (2023). Design of an Accented Character-Based Text Captcha System. *International Journal of Computer Science and Information Security (IJCSIS), 21(7), 91-100.*

[12]. Shahmerdanova, M. (2025, January). Artificial intelligence in translation: Challenges and opportunities. *Journal of Modern Linguistics and AI, 9(1), 55–68.*

[13]. Smith, J. (2020). "The history and evolution of reading". *Literacy Journal, 35(1), 22-35.*

[14]. Wang, S. F., Che, T. Y., Lu, Y. F., Yang, S. X., Huang, H., & Mao, X. L. (2022). Hammer PDF: An Intelligent PDF Reader for Scientific Papers. *Digital Libraries (cs.DL); Information Retrieval (cs.IR) https://arxiv.org/abs/2204.02809, 1-5.*

[15]. Wang, Y., Li, X., & Zhou, M. (2022). Hammer PDF: An intelligent PDF reader for scientific papers. *Proceedings of the 2022 IEEE Conference on Document Analysis and Recognition* (ICDAR)*, 203–214.*

[16]. Warnock, J. (2015). "The development of optical character recognition and part of speech tagging technologies". AI and Linguistics Journal, 10(3), 45-62.

[17]. Williams, R. (2018). "The rise of literacy and reading practices". *Reading Research Quarterly, 53(2), 213-228.*

[18]. Xie, J., Liu, P., Zhang, Y., & Chen, K. (2024). PDF-WuKong: Multimodal LLM for academic document comprehension. *arXiv preprint arXiv:2410.05970.*

[19]. Zafar, A., Ling, Y., & Kumar, R. (2025, February). AI-powered reading support for multilingual learners in higher education: A critical review. *International Journal of Educational Technology and AI*, *11(2), 77–92.*