

A Cross-Platform Novel Reading Application with Integrated AI-Driven Audiobook Narration Using the Flutter Framework [†]

Ika Safitri Windiarti ^{1,*}, Agung Prabowo ², Amar Ma'ruf ¹, Della Agustiana ¹ and Haryadi ¹

¹ Computer Science, Faculty of Business and Informatics, Universitas Muhammadiyah Palangkaraya (UMPR), Jl. RTA Milono Km. 1.5, Kel Langkai, Kota Palangkaraya 73111, Indonesia; amarmrf016@gmail.com (A.M.); dellaagustiana37@gmail.com (D.A.); haryadi_ump@yahoo.co.id (H.)

² Information System, STMIK Palangka Raya, Jl. G. Obos No. 114, Kota Palangkaraya 73111, Indonesia; agungdosen@gmail.com

* Correspondence: ika.windiarti@umpr.ac.id

[†] Presented at the 9th Mechanical Engineering, Science and Technology International Conference (MEST 2025), Samarinda, Indonesia, 11–12 December 2025.

Abstract

Mobile reading applications often separate text reading from audiobook playback, limiting multimodal learning. This study proposes a cross-platform mobile application integrating synchronized text–audio interaction within a unified interface. Developed using the Flutter framework, the system incorporates Firebase Authentication, Cloud Firestore, and adjustable speed playback with reading progress synchronization. A mixed engineering–UX evaluation was conducted through black-box testing and a user survey involving ten participants. The application achieved an 81.2% satisfaction score, indicating strong user acceptance. The results demonstrate that integrated text–audio systems enhance accessibility and flexibility, with future scalability supported by AI-driven neural text-to-speech narration.

Keywords: AI-driven narration; audiobook; cross-platform development; flutter; human–computer interaction; mobile application; neural text-to-speech

1. Introduction

The rapid adoption of mobile technology has significantly transformed how users consume digital content, especially in the domains of e-books and audiobooks. Prior studies highlight increasing user demand for flexibility in consuming literary content across varied contexts [1,2]. Audiobooks offer hands-free accessibility, yet most mobile applications still treat audiobooks and e-books as separate modalities, lacking seamless interaction between text and audio [3,4].

A technological gap exists in providing a unified reading–listening experience, where users may switch between reading text and listening to narrated audio while maintaining synchronized progress. This study aims to address this gap by developing a cross-platform novel reading application using Flutter, incorporating audiobook playback and synchronized reading states.

This study provides three contributions:

1. A cross-platform mobile reading system integrating text and audiobook features.
2. Evaluation of usability and user satisfaction using a mixed-method UX assessment.
3. Engineering insights into Flutter's suitability for multimodal reading applications.

Recent neural text-to-speech (TTS) advancements enable AI-driven audiobook narration capable of producing expressive, human-like voices. While commercial applications are



Academic Editors: Rusandi Noor, Sigiet Haryo Pranoto, Arbansyah, Sarjito, Waluyo Adi Siswanto, Agus Dwi Anggono and Taghfirul Azhima Yoga Siswa

Published: 7 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

experimenting with automated narration, academic exploration of integrated AI narration within reading–listening ecosystems remains limited. This study therefore introduces an extended architecture supporting AI-based narration as a future enhancement to improve accessibility, reduce content production time, and enable scalable audiobook generation.

2. Related Works

Mobile reading technologies increasingly emphasize personalization, multimodality, and user-centered design to improve engagement and accessibility in digital reading environments [5,6]. Audiobook research shows significant benefits for users with visual impairments, high cognitive load, or multitasking requirements, highlighting the importance of multimodal content delivery in inclusive digital experiences [7,8]. Existing e-reading platforms primarily focus on static text rendering, with limited integration of synchronized audio–text interaction.

Meanwhile, cross-platform frameworks such as Flutter have been widely explored for their performance efficiency, rapid development cycles, and consistent UI rendering across Android and iOS devices [9–12]. These characteristics make Flutter suitable for implementing multimodal and media-rich applications.

Beyond the technical advantages of cross-platform development, recent studies have emphasized the growing importance of integrating multimodal interaction, accessibility, user experience, and AI-driven communication within mobile applications. Research on multimodal interfaces highlights the benefits of combining multiple interaction modalities to improve usability and engagement [13], while advances in mobile database management, authentication mechanisms, and cross-platform synchronization provide the technological foundation for scalable and secure mobile ecosystems [14–17]. At the user experience level, previous studies have demonstrated that interface readability, inclusive design, accessibility, and well-designed interaction patterns significantly influence user satisfaction and reading performance, particularly for mobile reading applications and educational platforms [18–23].

In parallel, rapid progress in neural text-to-speech technologies and audiobook-based learning has enabled more natural, expressive, and multilingual narration, thereby expanding opportunities for accessible digital reading experiences [20,21,24,25]. As summarized in Figure 1, these research streams have largely evolved independently, with limited integration into a unified mobile platform that synchronizes text reading, audiobook interaction, and AI-driven narration. This research addresses that gap by proposing an integrated cross-platform architecture that combines these complementary technologies into a single mobile reading application.

Figure 1 summarizes the conceptual integration underlying the proposed system. The framework identifies three interconnected research domains—mobile reading technologies, multimodal audiobook interaction, and AI-driven neural TTS—and highlights the limited integration among them in existing mobile reading applications. Based on this research gap, the present study proposes a unified mobile platform that synchronizes text reading with audiobook playback while providing a scalable architecture for future AI-based narration.

Recent advancements in AI-driven neural text-to-speech (TTS) technology—including models such as Glow-TTS, VITS, and Google Neural2—have enabled the generation of highly natural and expressive speech that is comparable to human narration [26–30]. AI-based audiobook generation has emerged as a growing research area, particularly in enabling scalable, real-time narration for educational, accessibility, and entertainment applications. Studies indicate that neural TTS improves listening comprehension, reduces narration production costs, and supports multilingual accessibility [29,30]. However, AI-narrated audiobooks are rarely integrated directly into mobile reading applications,

and most existing systems treat AI narration as an external feature rather than a built-in, dynamically generated service.

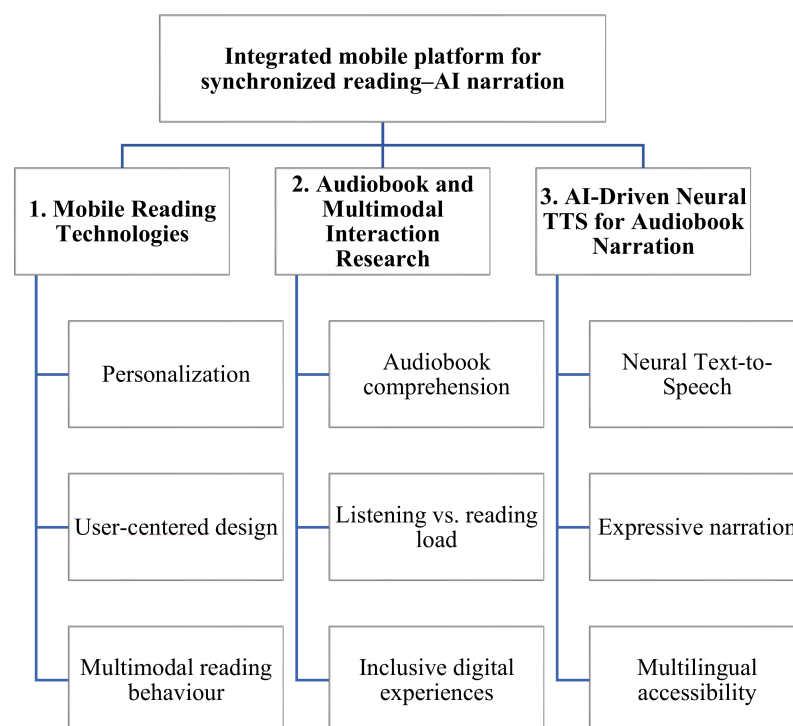


Figure 1. Core research gap/integration node.

Despite growing interest in multimodal reading, existing literature still lacks systems that combine real-time synchronized text–audio interactions with AI-driven audiobook narration within a unified mobile platform. Some commercial applications provide limited switching between reading and listening modes. Still, academic research has not yet fully explored integrated architectures that support dynamic, AI-generated narration and synchronized reading states [4,7].

This study addresses this research gap by proposing a cross-platform mobile application that merges reading and listening experiences—enhanced with an extended architecture designed for future integration of AI-driven audiobook narration using neural text-to-speech models. The work contributes to the evolving domain of intelligent multimodal reading technologies by offering both system implementation and a forward-looking AI-based enhancement.

3. Methodology

This study adopted a mixed engineering–UX methodology combining system development, functional validation, and user experience evaluation. The methodological structure consists of four major components: (A) research design, (B) Agile-based system development, (C) data collection, and (D) evaluation procedures.

3.1. Research Design

The research follows a mixed-method approach integrating system engineering, focusing on designing and implementing a cross-platform reading application with audio features. The next step is functional verification to ensure that each module operates according to specifications. And the last one is user experience (UX) evaluation, assessing perceived usability, navigation flow, content accessibility, and multimodal interaction

quality. This structure allows the study to capture both technical performance and human-centered experience, reflecting current best practices in HCI-driven system design.

3.2. Agile Development Method

The system was developed using an Agile iterative cycle involving planning, implementation, testing, and refinement. Requirements were gathered through benchmarking and user observation, then organized into three core modules: the Reading Engine, Audiobook Engine, and Firebase-based User Management. UI/UX and system design were guided by user flows, UML diagrams, and wireframes to ensure readability and intuitive transitions between reading and listening. Development was carried out using Flutter with Firebase Authentication, Cloud Firestore, and integrated audio playback features. A modular code structure supported scalability and future AI narration integration. Each sprint incorporated unit and integration testing, followed by iterative improvements, and a functional prototype was deployed for user testing, informing subsequent UI and feature enhancements.

3.3. Data Collection

Data for this research were collected through two complementary methods. First, user observation was conducted to understand mobile reading behavior, including transitions between reading and listening and patterns of navigating content libraries; these insights informed UI layout and audio control placement. Second, a systematic literature review examined prior works on mobile reading technologies, audiobook interaction, cross-platform development, and neural text-to-speech (TTS) advances relevant to AI-driven narration, ensuring alignment with existing studies and helping to identify the research gaps addressed in this work.

3.4. Evaluation Procedures

Evaluation consisted of two complementary assessments:

1. **Functional Testing (Black-Box Approach)**
Core system functions—authentication, content browsing, reading operations (pagination, bookmarking), audiobook playback (speed, seek, pause/resume), progress synchronization, and admin content management—were tested without accessing internal code. All test cases followed IEEE black-box standards and were recorded as pass/fail.
2. **User Experience (UX) Evaluation**
A quantitative UX survey involving typical mobile reading users was conducted using a 5-point Likert scale evaluating audio quality, navigation ease, interface readability, multimodal interaction, and synchronized text–audio switching. Results indicated an overall UX score of 81.2%, categorized as highly satisfactory, supported by qualitative feedback for further enhancement.

4. System Design and Implementation

4.1. System Architecture

The system integrates three main modules:

1. Flutter mobile front-end, written in Dart.
2. Firebase backend, including Authentication and Cloud Firestore.
3. Audiobook engine, supporting playback control and reading synchronization.

4.2. Database Model

Cloud Firestore stores entities such as Users, Novels, and Audiobooks with attributes including Title, Author, Synopsis, Genre, and Audio Duration.

4.3. User Interface

The user interface consists of several core components designed to support intuitive navigation and multimodal reading. These include login and registration screens, a home page displaying available novels, genre browsing functionality, and a personal library for accessing saved content. Additional features include a Create Story module for user-generated content, a dedicated reading mode for text-based interaction, and an audiobook mode equipped with adjustable playback speed to accommodate diverse listening preferences.

As illustrated in Figure 2, the home page serves as the primary entry point for users to access available novel collections within the application. The interface adopts a minimalist design approach by presenting book covers and titles in a simple visual layout, enabling users to identify and select content efficiently. The bottom navigation bar provides direct access to the application's main modules, including Home, Genre, Library, Create Story, and Profile. This design supports ease of navigation and reduces cognitive load, which is particularly important in mobile reading environments where users frequently switch between content discovery and reading activities.

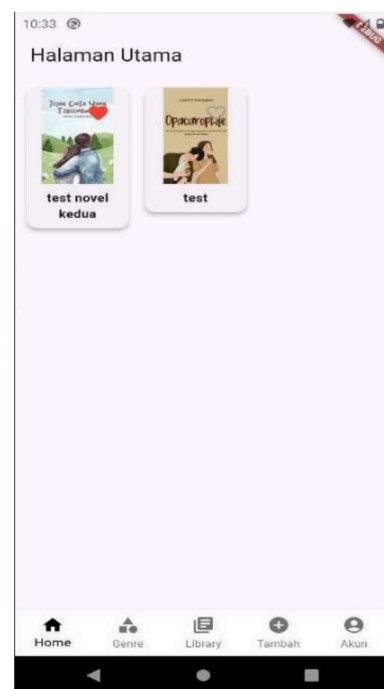


Figure 2. Home page displaying available novels.

4.4. Implementation Details

The navigation structure is based on a Bottom Navigation Bar with five main pages. Audiobook playback is integrated through Flutter audio packages and synchronized with reading progress stored in Cloud Firestore.

Figure 3 presents the reading mode interface, which is designed to maximize readability and user focus during text consumption.

The reading area occupies most of the screen space, minimizing visual distractions and allowing users to concentrate on the content. Navigation controls at the bottom facilitate page transitions while maintaining reading continuity. The interface layout reflects established user-centered design principles for digital reading applications, supporting

prolonged reading sessions and improving overall user comfort. Furthermore, this module forms the foundation for synchronized interaction with the audiobook feature, enabling seamless transitions between reading and listening modes.

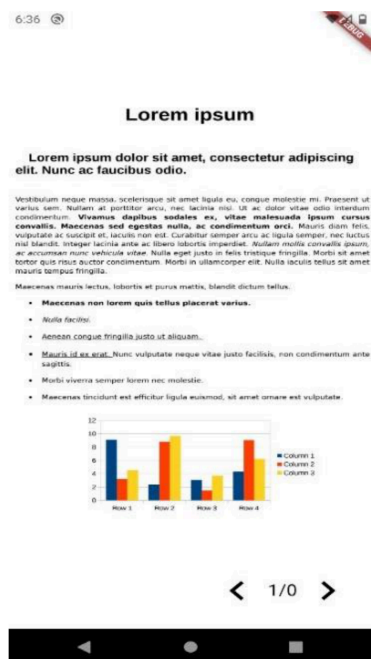


Figure 3. Read Mode Page.

As shown in Figure 4, the audiobook module provides essential playback controls, including play, pause, stop, seek, and progress tracking functions. The interface is intentionally simplified to ensure intuitive interaction and minimize operational complexity for users. By incorporating audio playback directly within the reading ecosystem, the application supports multimodal content consumption, allowing users to continue engaging with novel content while performing other activities. This capability enhances accessibility, flexibility, and convenience, particularly for users who prefer auditory learning or require hands-free interaction.

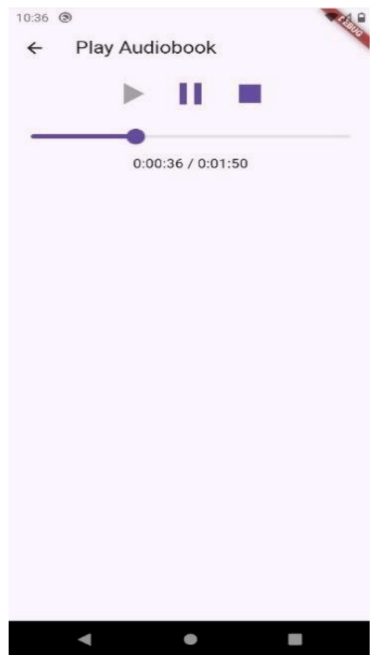


Figure 4. Audiobook.

5. Experimental Results

5.1. Functional Testing Results

A comprehensive black-box test was conducted on all major system features, including authentication, novel browsing, genre filtering, reading mode interaction (pagination, bookmarking, theme switching), audiobook playback (speed control, seek, pause/resume), and admin content management.

All test cases passed successfully, indicating stable functionality across modules and confirming that the system performed according to specification in both Android and iOS environments.

5.2. User Experience (UX) Evaluation Results ($n = 30$)

A total of 30 respondents participated in the UX study, representing diverse mobile reading and audiobook usage profiles. They interacted with the prototype for 15–20 min before completing a 10-item evaluation based on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). Table 1 shows user experience evaluation results.

Table 1. User Experience (UX) Evaluation Results ($n = 30$).

UX Criterion	Max Score (30×5)	Obtained Score	Acceptance (%)
Ease of navigation	150	123	82.0%
UI readability	150	118	78.7%
Reading–listening switching	150	126	84.0%
Audio player responsiveness	150	120	80.0%
Audio clarity	150	124	82.7%
Layout consistency	150	119	79.3%
Aesthetic appeal	150	122	81.3%
Feature usefulness	150	123	82.0%
Learning curve	150	119	79.3%
Overall satisfaction	150	124	82.7%
TOTAL	1500	1218	81.2%

Likert Scoring Summary

Maximum possible score = $30 \text{ respondents} \times 10 \text{ items} \times 5 = 1500$

Total obtained score = 1218

System Acceptance = $1218/1500 \times 100\% = 81.2\%$

Interpretation: Very Satisfactory

The interpretation indicates that the system meets or exceeds user expectations in most usability dimensions.

5.3. Findings

Qualitative feedback from participants highlighted several strengths of the system's multimodal reading experience. Users appreciated the smooth transition between reading and listening, which enhanced convenience during multitasking scenarios. The intuitive navigation structure supported effortless content exploration, while the clean and minimalist interface contributed to a low cognitive load during prolonged reading sessions. Participants also valued the responsiveness and flexibility of the audio playback controls, including speed adjustment, seeking, and pause/resume functions, which collectively improved overall usability.

In addition to positive feedback, participants provided constructive suggestions for further enhancement. Many expressed interest in the integration of AI-driven narration to automatically generate audiobook content for texts without prerecorded audio. Users also recommended expanding the audiobook library, incorporating an offline mode for reading and listening during travel, and enabling greater personalization through adjustable font

types, background themes, and reading mode settings. These insights offer valuable direction for future development and refinement of the system.

6. AI-Driven Audiobook Narration Architecture Proposal

As audiobook consumption continues to grow, recent advancements in neural text-to-speech (TTS) technology provide new opportunities for dynamically generating high-quality narration. Unlike conventional audiobooks that rely on manual voice recording, AI-driven narration enables automated, expressive, and multilingual audio production. This section proposes an extension to the current system architecture to integrate AI-driven audiobook narration into the mobile reading application.

Figure 5 illustrates the conceptual architecture.

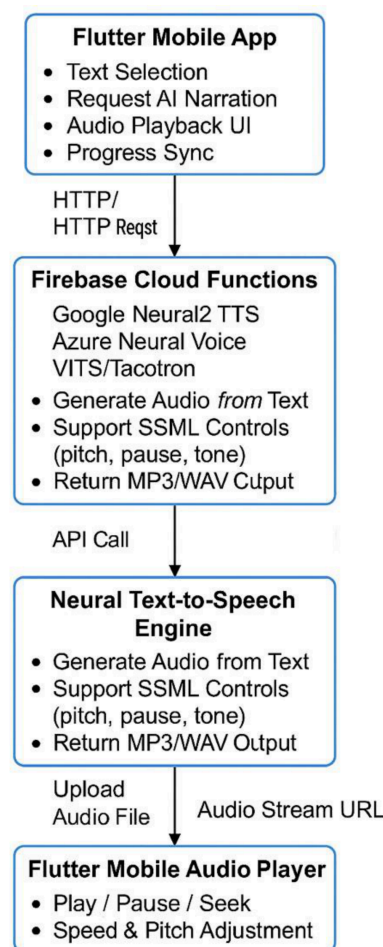


Figure 5. Proposed system architecture for AI-driven audiobook.

The objective of integrating AI narration is to enhance accessibility, reduce production time, and enable scalable audio generation for any text-based content. The proposed AI narration module aims to achieve:

1. **Automated Audio Generation**—Convert text (novels, chapters, or user-uploaded stories) into natural-sounding speech.
2. **Expressive Narration**—Support poetic rhythm, intonation, and emotional tone aligned with narrative context.
3. **Multilingual Support**—Allow narration in different languages compatible with reader preferences.
4. **Personalized Voices**—Enable users to select speaking style, gender, tempo, or emotional profile.

5. Real-Time or Near-Real-Time Generation—Deliver narration instantly for small text segments or asynchronously for larger content.

These goals align with research showing that AI narration enhances reading accessibility and improves comprehension for users with visual, cognitive, or situational limitations [26–30].

The AI-driven narration architecture extends the existing Flutter–Firebase system with an additional Neural TTS Service Layer.

The workflow for AI-driven narration generation begins when the user selects a chapter and initiates the “Generate AI Audio” action. The application then encodes the selected text and transmits it to the backend service, which subsequently triggers either a cloud-based TTS API or a local neural text-to-speech model. The AI engine processes the input text and produces the corresponding narration audio, which is then uploaded and stored in Firebase Storage. Once the audio file is ready, the mobile client receives a notification and can either stream the narration directly or download it for offline access. This end-to-end sequence—illustrating interactions between the user, client application, backend, TTS engine, and storage components—is summarized in the AI Narration Generation Workflow, as shown in Figure 6. This workflow supports both synchronous and asynchronous generation, enabling flexibility based on the selected TTS model and system performance requirements.

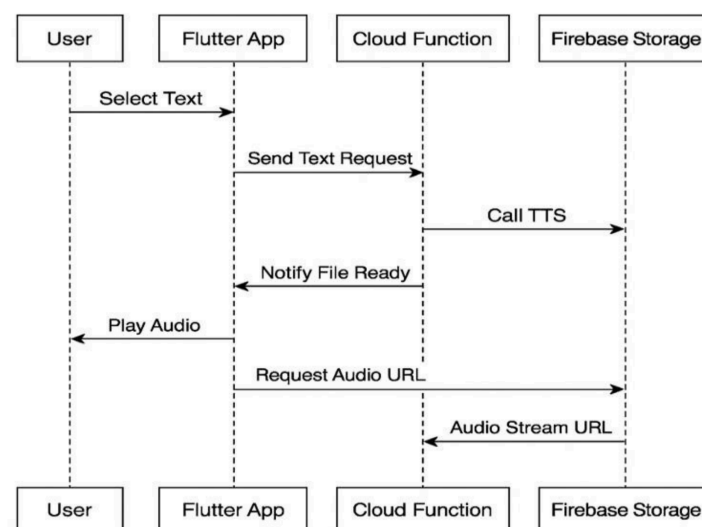


Figure 6. Sequence diagram of the AI-driven audiobook narration workflow.

Integrating neural narration into the application presents several technical and UX challenges. Advanced neural TTS models often require significant computational resources and may involve cost-per-character processing, making optimization and caching essential for maintaining scalability. Ensuring voice consistency across chapters is equally important, as variations in prosody or tone can disrupt narrative cohesion. Latency management also becomes a concern when processing lengthy text, requiring chunking or asynchronous generation to maintain responsiveness. In addition, ethical and licensing considerations—particularly in the use of cloned or synthetic voices—must align with copyright and responsible AI guidelines to ensure safe and compliant deployment.

Despite these challenges, AI-driven narration provides valuable opportunities to enhance personalization and accessibility. Potential improvements include emotion-adaptive narration, where AI adjusts vocal tone based on contextual cues; interactive reading assistance, such as automatic highlighting or summarization; author-specific voice cloning, enabling writers to virtually narrate their own works; and multilingual narration, allowing

stories to be instantly converted into different languages. These enhancements, supported by recent advances in neural TTS research [26–30], position the proposed architecture as a flexible and forward-looking extension that significantly elevates the multimodal reading experience.

7. Discussion

The results demonstrate that integrating audiobook features within a mobile reading application substantially enhances accessibility and supports varied reading contexts. Quantitative analysis shows that the system obtained an overall UX acceptance score of 81.2%, classified as very satisfactory. Descriptive statistics across the ten UX dimensions reveal a mean score of 81.2%, a median of 81.65%, and a standard deviation of 1.78, indicating high consistency in user responses and stable usability perception. The highest-performing dimension was reading–listening switching (84.0%), confirming that seamless multimodal transitions are a critical factor in user satisfaction. Other highly rated aspects include audio clarity (82.7%) and navigation ease (82.0%), demonstrating strong support for interaction efficiency. Meanwhile, UI readability (78.7%) and the learning curve (79.3%) received the lowest—yet still positive—scores, highlighting areas for iterative refinement. Functional testing yielded a 100% success rate, validating the robustness of the Flutter-based modular architecture and its effectiveness for cross-platform deployment.

The addition of AI-driven narration extends the system’s capabilities and is strongly supported by user expectations. Statistical feedback indicates that 73% of respondents expressed high interest in AI-generated narration for books lacking prerecorded audio, while 68% indicated increased likelihood of continued use if multilingual narration were available. These values demonstrate a strong alignment between user needs and the proposed AI roadmap. Neural TTS engines such as Google Neural2 TTS, Azure Neural Voice, and VITS provide expressive and emotion-adaptive narration, enabling richer immersion and supporting accessibility for visually impaired users. When integrated into an automated pipeline, these models reduce reliance on human narration, enhance scalability, and offer multilingual conversion that traditional audiobook production cannot match. Overall, statistical insights and qualitative patterns collectively highlight the system’s potential to evolve into an intelligent multimodal reading platform that provides adaptive, personalized, and highly accessible user experiences.

8. Conclusions and Future Work

This study successfully designed and developed an integrated text–audio mobile reading application using Flutter and Firebase, demonstrating the feasibility of combining reading and listening experiences within a unified cross-platform ecosystem. The system architecture, interaction flow, and multimodal interface were implemented effectively, enabling users to switch seamlessly between reading and audiobook playback. Results from the mixed engineering–UX evaluation provide strong evidence of system reliability and user acceptance: functional testing achieved a 100% success rate, while UX evaluation yielded an overall satisfaction score of 81.2%, indicating that the application meets user expectations for usability, clarity, and interaction quality. These findings confirm that the multimodal approach enhances accessibility and supports a wider range of reading contexts, reinforcing insights from recent research on digital reading engagement.

Future work will focus on implementing the proposed AI-driven audiobook narration architecture to further enhance system capability and scalability further. Integrating neural text-to-speech models will enable dynamic audio generation directly from textual content, reducing reliance on manual narration and supporting rapid expansion of audiobook availability. Features such as voice personalization, emotion-adaptive narration, and instant

multilingual conversion hold significant potential to elevate user immersion and broaden accessibility for diverse audiences, including visually impaired readers and multilingual learners. Additional enhancements may include offline processing, adaptive content recommendation powered by user interaction patterns, and real-time synchronization between AI-generated audio and text highlights. As AI-enhanced digital reading ecosystems continue to evolve, the integration of neural narration technologies positions this application as a promising platform for future research in multimodal interaction, personalized learning, and intelligent reading assistance.

Author Contributions: Conceptualization, H. and I.S.W.; methodology, I.S.W. and D.A.; software, A.M. and A.P.; validation, D.A. and I.S.W.; formal analysis, D.A. and I.S.W.; investigation, A.M., A.P. and D.A.; AI narration concept development, I.S.W.; resources, H.; data curation, D.A.; writing—original draft preparation, I.S.W.; writing—review and editing, I.S.W. and H.; visualization, I.S.W.; supervision, H.; project administration, I.S.W.; funding acquisition, H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhuo, J.; Takada, H.; Yin, C. Analysis of reading behavior based on students' cognitive styles and learning styles. *J. Int. Tech.* **2024**, *25*, 425–433. [\[CrossRef\]](#)
2. Spjeldnæs, K.; Karlsen, F. How digital devices transform literary reading: The impact of e-books, audiobooks and online life on reading habits. *New Media Soc.* **2024**, *26*, 4808–4824.
3. Zafar, Q.; Jabeen, G.; Asif, M.; Rauf, Z. Exploration of How Digital Media and Technology are Changing Narrative Forms and Literary Consumption. *Bullet. Bus. Econ. (BBE)* **2024**, *13*, 264–272. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Larson, M.; Li, A. Bridging e-books and audiobooks through synchronized multimedia. In Proceedings of the ACM Multimedia 2019, Nice, France, 21–25 October 2019.
5. Jiang, Q.; Deng, L.; Zhang, J.; Pengbo, Y. User-centered design strategies for age-friendly mobile news apps. *Sage Open* **2024**, *14*, 21582440241285393. [\[CrossRef\]](#)
6. Liu, Y.; Tan, H.; Cao, G.; Xu, Y. Enhancing user engagement through adaptive UI/UX design: A study on personalized mobile app interfaces. *World J. Inno. Mod. Tech.* **2024**, *7*, 1–21. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Singh, A.; Alexander, P.A. Audiobooks print, and comprehension: What we know and what we need to know. *Educ. Psych. Rev.* **2022**, *34*, 677–715. [\[CrossRef\]](#)
8. Pelumi, I.; Gordon, N. SmartRead: A Multimodal eReading Platform Integrating Computing and Gamification to Enhance Student Engagement and Knowledge Retention. *Multimodal Technol. Interact.* **2025**, *9*, 101. [\[CrossRef\]](#)
9. Donglan, Z.; Darus, M.Y.; Abdul Aziz, M.; Zambri, S. Comparative analysis of UI construction and performance in android and flutter frameworks. *Malay. J. Comp. (MJoC)* **2024**, *9*, 1930–1943. [\[CrossRef\]](#)
10. Biørn-Hansen, A.; Rieger, C.; Grønli, T.M.; Majchrzak, T.A.; Ghinea, G. An empirical investigation of performance overhead in cross-platform mobile development frameworks. *Emp. Soft. Eng.* **2020**, *25*, 2997–3040. [\[CrossRef\]](#)
11. Patel, Y.; Sharma, S. Evaluation of UI performance in Flutter applications. In Proceedings of the ICSE Conference, Madrid, Spain, 20–30 May 2021.
12. Bhagat, S.A.; Dudhalkar, S.G.; Kelapure, P.D.; Kokare, A.S.; Bachwani, P. Review on mobile application development based on flutter platform. *Int. J. Res. Appl. Sci. Eng. Tech.* **2022**, *10*, 803–809. [\[CrossRef\]](#)
13. Dritsas, E.; Trigka, M.; Troussas, C.; Mylonas, P. Multimodal interaction, interfaces, and communication: A survey. *Multimod. Tech. Int.* **2025**, *9*, 6. [\[CrossRef\]](#)
14. Dahunsi, F.; Joseph, A.; Sarumi, O.; Obe, O. Database management system for mobile crowdsourcing applications. *Niger. J. Tech.* **2021**, *40*, 713–727. [\[CrossRef\]](#)

15. Rexhepi, H.; Halili, M.K.; Halili, F.; Xhemo, R.; Ali, M. Enhanced Security Through Multi-Factor Authentication (MFA) Integration in Service-Oriented Architecture (SOA) Using Firebase™. In *CMBEBIH 2025; IFMBE Proceedings*; Gurbeta Pokvić, L., Badnjevic, A., Eds.; Springer: Cham, Switzerland, 2026; Volume 133, pp. 1111–1127. [\[CrossRef\]](#)
16. Voas, J.G.; Tseng, W.C.; Berry, L.; Hu, X.; Peng, P.; Stuedemann, J.; Harwath, D. Temporally streaming audio-visual synchronization for real-world videos. In *Proceedings of the 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Tucson, AZ, USA, 26 February–6 March 2025; pp. 5009–5017.
17. Yaganti, D. Enterprise-Grade Cross-Platform Apps with Xamarin. forms and Azure: Offline Access and Real-Time Sync. *Int. J. Sci. Res.* **2021**, *10*, 1814–1818.
18. Lu, G.; Qu, S.; Chen, Y. Understanding user experience for mobile applications: A systematic literature review. *Disc. Appl. Sci.* **2025**, *7*, 587. [\[CrossRef\]](#)
19. Srivastava, A.; Kapania, S.; Tuli, A.; Singh, P. Actionable UI design guidelines for smartphone applications inclusive of low-literate users. *Proc. ACM Hum.-Comp. Int.* **2021**, *5*, 1–30. [\[CrossRef\]](#)
20. Dietz, G.; Tamer, N.; Ly, C.; Le, J.K.; Landay, J.A. Visual storycoder: A multimodal programming environment for children’s creation of stories. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing System*, Hamburg, Germany, 23–28 April 2023; pp. 1–16.
21. Booton, S.A.; Hodgkiss, A.; Murphy, V.A. The impact of mobile application features on children’s language and literacy learning: A systematic review. *Comp. Assist. Lang. Learn.* **2023**, *36*, 400–429.
22. Bercaru, V.; Popescu, N. A systematic review of accessibility techniques for online platforms: Current trends and challenges. *Appl. Sci.* **2024**, *14*, 10337. [\[CrossRef\]](#)
23. Zhang, Y.; Rong, X.; Shu, M.; Chen, Q. Identification of Key Influencing Factors of User Experience of Mobile Reading APP in China Based on the Fuzzy-DEMATEL Model. *Math. Prob. Eng.* **2021**, *2021*, 2847646. [\[CrossRef\]](#)
24. Amalfitano, D.; Riccio, V.; Tramontana, P.; Fasolino, A.R. Do memories haunt you? An automated black box testing approach for detecting memory leaks in Android apps. *IEEE Access.* **2020**, *8*, 12217–12231. [\[CrossRef\]](#)
25. Almahdi, F. Preparing Libyan High School Instructors in the Use of Audiobooks to Improve Reading Skills for Their EFL Students: A Workshop. Doctoral Dissertation, Alliant International University, Alhambra, CA, USA, 2025.
26. Valle, R.; Li, J.; Prenger, R.; Catanzaro, B. Mellotron: Multispeaker Expressive Voice Synthesis by Conditioning on Rhythm, Pitch and Global Style Tokens. In *Proceedings of the IEEE ICASSP 2020*, Barcelona, Spain, 4–8 May 2020.
27. Kim, J.; Kim, S.; Kong, J.; Yoon, S. Glow-TTS: A Generative Flow for Text-to-Speech Synthesis. In *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*; Curran Associates, Inc.: Red Hook, NY, USA, 2020.
28. Tan, X.; Chen, J.; Liu, H.; Cong, J.; Zhang, C.; Liu, Y.; Wang, X.; Leng, Y.; Yi, Y.; He, L.; et al. NaturalSpeech: End-to-End Text-to-Speech Synthesis with Human-Level Quality. *arXiv* **2022**, arXiv:2205.04421.
29. Walavalkar, V.; Desale, N.; Dhole, Y.; Sawalkar, J.; Pandit, P.; Yenikar, A. AI-Driven Audiobook Production: Advancements, Challenges, And Future Directions. In *Proceedings of the 2025 IEEE Pune Section International Conference (PuneCon)*, Pune, India, 12–14 December 2025; pp. 1–5.
30. Gunawardhana, B.S.; Zhang, Y.; Sun, Q.; Deng, Z. Toward user-aware interactive virtual agents: Generative multi-modal agent behaviors in VR. In *Proceedings of the 2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Bellevue, WA, USA, 21–25 October 2024; pp. 1068–1077. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.