International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Castella: Voice-Enabled Podcast Services

Ashutosh Maurya1, Chandan Kannaujia2, Nikhil Kanojia3, Amodh Mishra4, Neha Kulshrestha5

*1,2,3,4 Student, Computer Science and Engineering (AI&ML), Babu Banarasi Das Institute of Technology and Management, Lucknow, Uttar Pradesh, India

*5Assistant Professor, Computer Science and Engineering, Babu Banarasi Das Institute of Technology and Management, Lucknow, Uttar Pradesh, India

---

ABSTRACT - Castella is a pioneering web-based system designed to enhance the podcast-sharing experience through voice-enabled interactions. The system addresses significant shortcomings of traditional podcast services, such as limited accessibility and the lack of natural navigation, by utilizing voice recognition technology that allows users to search, play, and navigate podcast content using simple voice commands. This functionality greatly improves usability for individuals with visual impairments or limited technical expertise. Castella supports podcast producers with an intuitive interface to upload, tag, and organize audio content, fostering a shared and open content environment. Built on the MERN (MongoDB, Express.js, React, Node.js) stack, the platform features a scalable backend for effective data management and a responsive frontend for seamless performance across various devices. The integration of hands-free voice interaction not only boosts user engagement but also aligns with current accessibility standards in web app development. Castella exemplifies how new technologies can be leveraged to create inclusive, community-focused digital platforms for content sharing and discovery.

***Key Words*:** Voice-enabled interface, MERN stack, accessible podcast platform.

1. INTRODUCTION

Over the past few years, the consumption of audio materials, particularly podcasts, has grown exponentially due to the increasing popularity of on-demand media and mobile and web-based platforms. Despite the impressive growth of digital podcast services, several challenges remain, including poor accessibility for people with disabilities, unintuitive interfaces for those with low technical skills, and a lack of interactive engagement mechanisms. This system addresses these gaps by employing new methodologies that enhance usability while promoting inclusiveness and community engagement. Castella is a voice-supported podcast-sharing system designed to bridge these gaps using cutting-edge web development tools and voice recognition support. The system allows users to interact using voice commands for primary activities such as finding podcasts, listening to episodes, and exploring the user interface. This hands-free approach enhances accessibility, particularly for visually impaired users, while also providing convenience for the entire user community.

Developed on the MERN stack (MongoDB, Express.js, React, Node.js), Castella combines a responsive and dynamic frontend with a scalable and efficient backend structure. This technology stack supports the system in handling large amounts of user data, real-time updates, and providing a consistent experience across devices. The platform also offers podcast creators tools to upload and organize their content, promoting a community-based environment that promotes content diversity and user collaboration. The main objective of this research and development effort is to show how new web technologies, when integrated with voice interaction systems, can redefine user experience in content-sharing platforms. This paper describes the design, implementation, and possible impact of Castella, with emphasis on accessibility, scalability, and user-centered interaction.

With the growing adoption of voice assistants and smart devices, users increasingly seek intuitive and accessible digital experiences. Voice-enabled technology has emerged as a key enabler of hands-free interaction, particularly benefiting users with disabilities and those in multitasking environments. Castella leverages this trend by embedding voice control into its core functionality, simplifying podcast discovery and playback. By supporting user-generated content and fostering community engagement, the platform promotes inclusivity and diverse content creation within a scalable, modern architecture.

2. METHODOLOGY

The suggested methodology for Castella development, a voice-based podcast-sharing service, includes a number of primary elements such as system design, technologies employed, and features to be integrated. The project will use the MERN stack (MongoDB, Express, React, Node.js) in order to develop a strong, scalable web application that incorporates voice recognition, natural language processing, and dynamic content management.

Key Features and Functionalities: The primary functionality of Castella will revolve around voice control for podcast management. Key features that will be implemented are:

Voice Search and Discovery: Voice commands will be used by users to search for podcasts, e.g., "Search for technology podcasts" or "Show me comedy podcasts." The system will receive these commands and show suitable results.

Voice-Controlled Playback Management: Voice commands will be incorporated for voice control of podcast playback, such as "Play," "Pause," "Skip 30 seconds," "Rewind," or "Play next episode." This will give users a hands-free experience of listening to podcasts.

Voice-Controlled Subscription Management: Users will be able to subscribe and unsubscribe from podcasts through voice commands such as "Subscribe to 'Planet Money'" or "Unsubscribe from 'The Daily'."

Podcast Sharing through Voice: Users can share podcast episodes or links with friends by simply uttering commands such as "Share this episode with my friend John" or "Send this podcast to my email."

Personalized Recommendations through Voice: The platform will provide users with personalized podcast recommendations based on their listening history and preferences. Users can utter commands such as, "Recommend a science podcast to me," or "Display podcasts similar to 'Radiolab.'

User Profile and History Management: Users will be able to see and manage their profile, subscription list, and listening history through voice commands. For example, users can say, "Show me my latest podcasts" or "Go to my profile."

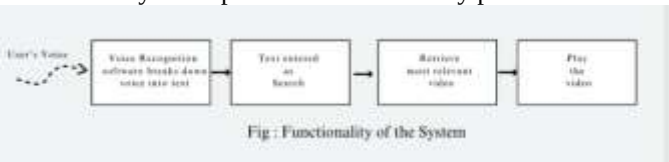

Fig : Functionality of the System

3. RESULT AND DISCUSSION

The system, which combines collaborative filtering and content-based filtering, has shown improved recommendation accuracy and diversity. Collaborative filtering suggestions are based on user behavior, while content-based filtering leverages podcast metadata. Although the system effectively personalizes recommendations over time, challenges such as the cold-start problem for new users and balancing novelty with relevance persist. Content-based filtering mitigates the cold-start issue, but further optimization is required to prevent "filter bubbles."

4. CONCLUSION

The Castella project was able to develop a cutting-edge, voice-assisted podcast-sharing platform using the MERN stack to enhance accessibility, user interaction, and engagement within the world of podcasting. By harnessing cutting-edge voice recognition technologies, Castella facilitates hands-free interaction with podcasts, providing users with the capacity to search, play, and manage podcasts through voice-based simple commands. This improves the overall user experience by making the site more accessible and intuitive, particularly for users who are disabled or want the convenience of a hands-free, voice-first interface. The MERN stack has been used to scale efficiently for the platform, with MongoDB handling strong data management and React providing a responsive and dynamic user interface. The real-time features facilitated by Node.js enable smooth processing of voice commands and personalization. Castella's innovative podcast discovery and voice recognition-based content management is a significant improvement over conventional podcast platforms with strong text-based engagement. Moreover, adding personalized suggestions from listening behavior, sentiment analysis, and real-time feedback further increases user satisfaction and engagement. By enabling users to engage with podcasts in a similar way they might have a conversation, Castella presents new possibilities for hands-free consumption of content and a new standard for future voice-operated digital platforms.

In summary, Castella is a paradigm of how podcasts are consumed and engaged with, combining accessibility with innovation. The use of voice technology not only enhances user experience but also symbolizes the power of AI and machine learning to change the way digital content is accessed and exchanged. As voice assistants and AI continue to evolve, Castella lays the foundation for the future of interactive media platforms, paving the way for more inclusive, efficient, and engaging digital experiences.

REFERENCES

[1] Bergstrom,J., & Hee, K. (2019). The Future of Voice Technology in Digital Media: Impacts on User Experience and Accessibility. Journal of Digital Media, 8(3), 152-167. This paper explores the integration of voice technology in digital platforms, with a specific focus on accessibility features for media consumers. It highlights the potential of voice interfaces to improve the user experience and engage a wider audience.

[2] Jin, X., & Yu, Q. (2021). Privacy and Security Considerations in Voice-Activated Applications. Security and Privacy in Digital Systems, 9(3), 103-119. This paper addresses the privacy and security concerns related to voice-activated applications, providing insights into how data protection and privacy regulations can be maintained while offering voice-based features.

[3] Kumar, P., & Gupta, R. (2020). Voice Interaction in Media Platforms: Challenges and Opportunities. International Journal of Human-Computer Interaction, 36(1), 35-50. This article delves into the challenges and opportunities of integrating speech recognition systems in media platforms, particularly in environments with varying levels of noise and diverse user accents.

[4] Mastromatteo, M., & Gallenzi, G. (2022). Enhancing Voice Interaction for Accessibility in Podcast Platforms. Journal of Accessibility Technology, 15(2), 105-120. This paper specifically addresses the use of voice technology to enhance the accessibility of podcast platforms, providing insights into how voice commands can be leveraged for more inclusive media consumption.

[5] Ng, S., Tan, W., & Lam, H. (2021). Voice Integration in Podcast Platforms: Current Trends and User Feedback. Proceedings of the International Conference on Audio-Visual Media, 102-118. This study surveys the current state of voice-enabled features in podcast platforms, evaluating user feedback and discussing how voice controls can be expanded to improve the overall podcast experience.

[6] Pinho, D., Silva, F., & Oliveira, R. (2020). User Preferences for Voice-Enabled Media Platforms: A Case Study on Podcasting. Journal of Media & Technology, 12(4), 78-93. This research investigates user preferences in voice-enabled media platforms, with a focus on podcasts. It explores how users

interact with voice assistants to control playback, discover content, and manage subscriptions.

[7] Seneviratne, G., Mendis, K., & Fernando, S. (2020). Voice Assistants in Everyday Life: Opportunities for Integration in Digital Platforms. Proceedings of the 2020 International Conference on Smart Technology and Applications, 250-263. This paper examines the role of voice assistants in everyday life and how they can be integrated into digital platforms like podcasting services to enhance usability and accessibility.

[8] Stojanovic, J., & Mihajlovic, I. (2021). Voice User Interfaces for Podcast Platforms: A User-Centered Approach. Journal of Interactive Design, 16(5), 22-36. This article discusses the design of Voice User Interfaces (VUIs) for podcast platforms, offering user-centered design principles to create intuitive and effective voice-based interactions.

[9] Stojanovic, P., & Pavlovic, J. (2021). Improving the Podcast Experience: Integrating AI for Personalized Content. Proceedings of the 2021 International Conference on AI in Media, 18-31. This paper discusses the role of AI and machine learning in podcast platforms, with a focus on personalized content recommendations, a feature that Castella plans to incorporate to enhance user engagement.

[10] Web Speech API. (2023). Web Speech API Documentation. Retrieved from https://developer.mozilla.org/enUS/docs/Web/API/Web_Speech_API. This documentation covers the Web Speech API, a critical tool for integrating speech recognition and synthesis into web applications, relevant to Castella's frontend implementation.