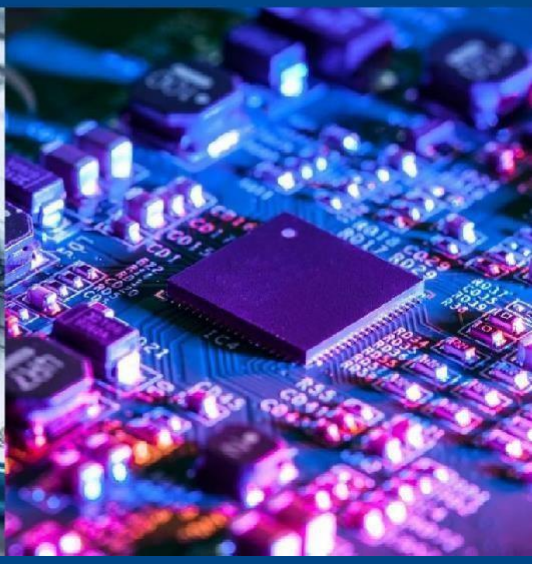
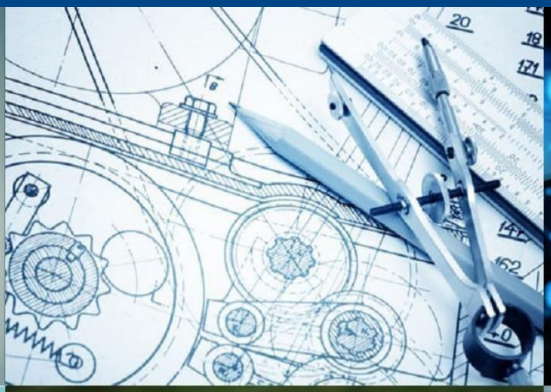
ISSN: 2582-7219

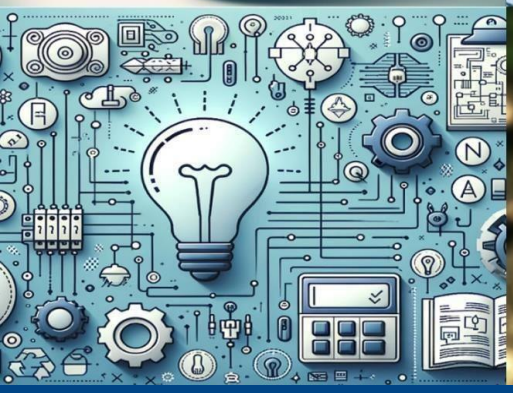




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 8, August 2025

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808076

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

SONICSCRIPT

Dr. Charles Arockiaraj M, Jayalakshmi

Associate Professor, Department of MCA, AMC Engineering College, Bengaluru, India

Student, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: SonicScript is a highly sophisticated transcription software platform that enables users to convert video and audio content to text with a high degree of accuracy, appealing to educators, content producers, and professionals alike. This platform is built on a modern tech stack (React, Next.js, Tailwind CSS, and Firebase) to provide a secure mobile/responsive and simple to use platform that eliminates headaches of submitting transcriptions and calculate time between podcast recordings and finalized transcript. Users can easily upload files or submit links to outside content (YouTube) or video and audio files on their desk or mobile devices. SonicScript is not only excellent for its front-end solution, but utilizes an in-house trained machine learning model to improve upon the quality of each and every transcription without access to existing third party APIs. Sonic Script has excellent authentication, secure data storage with built-in credits to manage billing without exposing the user to customer service, and per-access charging. SonicScript pushes for innovation with new features such as automatic feedback on transcription, download options as pdfs, specialized podcast style transcriptions, ability to submit other web links, and automatic text summarization

I. INTRODUCTION

The Sonic Script design embodies a forward- looking recap platform that aims to grease the audio processing and recap experience for stoners in the digital space. As a Next.js predicated frame(a Javascript- predicated frame erected on React) known for its performance and capability to gauge as demand is demanded, Sonic Script is designed to give a no- quiescence recap experience.

The Sonic Script architecture does n't only encompass features that are user- facing, it also enables secure data operation with Firebase Firestore as its core database. The design executed a document- acquainted schema, for the user data is stored into individual collections with customized access rule predicated on the clearances established within the Firestore security rules. As well, the codebase supports modularity that separates the operation into factors, APIs and services that contribute to maintainability, scalability and growth of features.

II. LITERATURE SYRVEY

[1] The transcription landscape has undergone transformative changes over the last few years, driven by advancements in Artificial Intelligence (AI), Natural Language Processing (NLP), and Cloud Computing. With the continued proliferation of audio content from interviews, podcasts, webinars, and conferences, the ability to find a reliable, inexpensive, and scalable transcription solution has never been more important. While there continues to be many commercial transcription platforms, most of these systems suffer from issues of scalability, cost, workflows, and transparency.As a result, academic and industry researchers have been exploring new open-source models, cloud-based architecture, and modern DevOps

practices to build platforms that are technically and monetarilyfunctional This literature survey will investigate the present status of transcription technologies, transcription architectural paradigms, open-source models, and best practices for building transcription platforms—most notably related to adopting Next.js in addition to Docker and Firebase. Transcription technologies have come a long way from transcribing manually to post- editing of automatic transcription. The first generation systems were template-based recognition with small vocabularies which made them impractical for real world use Deep learning and AI-based models like OpenAI's Whisper have been ever-evolving, and transcription accuracy, for a number of languages and dialects, is surpassing levels we have come to expect. Supported by other studies, we have shown that models based on neural networks are always more accurate than models that are informed primarily by language which includes when there is a large amount of noise or the vocabulary is based heavily on domain specific knowledge Sonic Script is using Whisper, and has relied on this advanced technology to

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

deliver high quality transcription at no licensing costs attributable to commercial engines.

[2] Open-Source vs .Proprietary Solutions: considerations while building modern transcription platforms is the question of which path to follow... open source frameworks or proprietary systems? There are clear advantages of potential accessibility and transparency by leveraging open source solutions such as Whisper, which could be useful for startups and independent developers Proprietary solutions typically come with an excessive price tag with very few customization options and, sometimes not but risks of vendor lock-in. The decision to use Whisper for Sonic Script is aligned with the movement toward open-source solutions seen in the industry, partly because of the financial barriers that are sometimes experienced by early developers and startups. Additionally, an open-source approach encourages a community of interactive development where knowledge is shared and improvements/ customization is contributed back to the community, which again is helpful for early-stage developers

II. LITERATURE SYRVEY

[1] The transcription landscape has undergone transformative changes over the last few years, driven by advancements in Artificial Intelligence (AI), Natural Language Processing (NLP), and Cloud Computing. With the continued proliferation of audio content from interviews, podcasts, webinars, and conferences, the ability to find a reliable, inexpensive, and scalable transcription solution has never been more important. While there continues to be many commercial transcription platforms, most of these systems suffer from issues of scalability, cost, workflows, and transparency.As a result, academic and industry researchers have been exploring new open-source models, cloud-based architecture, and modern DevOps

practices to build platforms that are technically and monetarilyfunctional This literature survey will investigate the present status of transcription technologies, transcription architectural paradigms, open-source models, and best practices for building transcription platforms—most notably related to adopting Next.js in addition to Docker and Firebase. Transcription technologies have come a long way from transcribing manually to post- editing of automatic transcription. The first generation systems were template-based recognition with small vocabularies which made them impractical for real world use Deep learning and AI-based models like OpenAI's Whisper have been ever-evolving, and transcription accuracy, for a number of languages and dialects, is surpassing levels we have come to expect. Supported by other studies, we have shown that models based on neural networks are always more accurate than models that are informed primarily by language which includes when there is a large amount of noise or the vocabulary is based heavily on domain specific knowledge Sonic Script is using Whisper, and has relied on this advanced technology to deliver high quality transcription at no licensing costs attributable to commercial engines.

[2] Open-Source vs .Proprietary Solutions: considerations while building modern transcription platforms is the question of which path to follow... open source frameworks or proprietary systems? There are clear advantages of potential accessibility and transparency by leveraging open source solutions such as Whisper, which could be useful for startups and independent developers Proprietary solutions typically come with an excessive price tag with very few customization options and, sometimes not but risks of vendor lock-in. The decision to use Whisper for Sonic Script is aligned with the movement toward open-source solutions seen in the industry, partly because of the financial barriers that are sometimes experienced by early developers and startups. Additionally, an open-source approach encourages a community of interactive development where knowledge is shared and improvements/ customization is contributed back to the community, which again is helpful for early-stage developers.

[3]. Data Privacy&Security
Due to the sensitive nature of audio data, the application must incorporate substantial security measures and transparent data governance. Research regarding sensitive data has established that unauthorized access, data leaks, and regulatory implications [GDPR, CCPA, etc.] are all major risks associated with handling sensitive data .Sonic Script can perform authentication with Firebase Authentication and management of digital assets/secure database with Firestore security rules.

This uses Google's infrastructure and allows for very fine-grain access, and will provide data at rest encryption, and it will also provide the user with a clear claim to data ownership and options regarding retention policy helping to develop trust further and assist with compliance. Research indicates that use of systems that exhibit a good privacy manner are considered significant by users who are willing to pay significantly for these systems, and non-compliance represents a

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808076

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

risk in and of itself develop an intuitive and efficient interface that enables users to interact seamlessly with a virtual mouse, Before SonicScript, the only alternative for performing audio and video transcriptions was to do so manually, or possibly using an online tool that had very limited functionality and accuracy. For manual transcriptionit could be fairly accurate in some aspects, but it was very slow and costly, and cannot be scaled for larger jobs or 'real time' requirements.Most of the online tools were mostly 'cheap [or free]' and faster, but in -accuracy related to any functional protocol related to variables such as noisy environments, multiple speakers, and accents - with multiple languages, that would often require

intervention. The same type of limitations existed with the online transcription tools, with most limits being related to the ability to input the file as a video file, as well as real time transcriptions Also neither online or manual had any ability to support special or ancillary services around the verification of users, the payment process, or create any sort of credit system. There are also many significant detected issues with how user "data" file privacy, file security, and if or how any platform could have any form of user controls or files that might be in upload storage on an "asset". Transcription has consisted of a fragmented traditional style legacy services that have proven to be slow and ineffective than it can be expect for the modern day, scalable, secure needs - thus the whole reason for an integrated and holistic solution like SonicScript.

III. SYSTEM ARCHITECTURE

SonicScript consists of a system architecture that is modular, cloud-based, and utilizes modern web technologies. This enables the implementation of AI processes, such as OpenAI's Whisper model for speech-to-text transcription, in scalable backend service modules. SonicScript's front end (client interface and admin interface) is built with next.js web framework technologies which enable responsive interface designs, while Firebase Authentication enables user authentication and user roles (admin and client) and session management, while also being able to sign users out for added security.

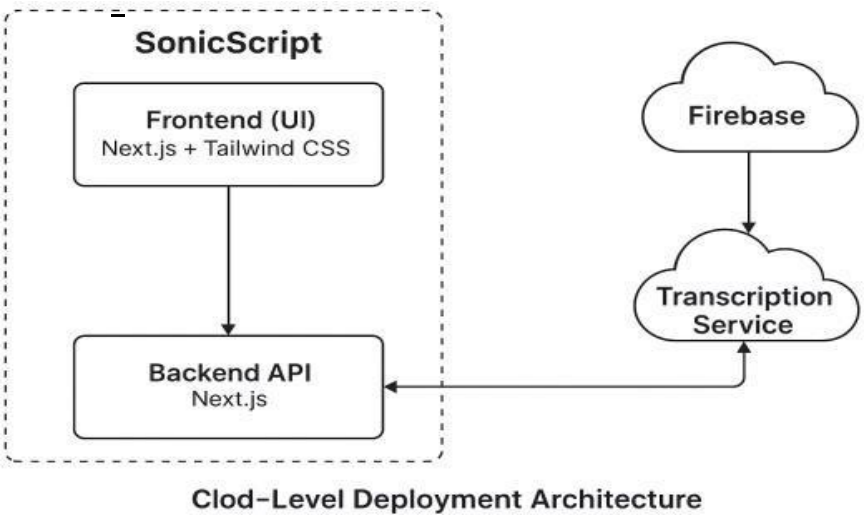


Fig 3.1 System Architecture

IV. METHODOLOGY

SonicScript was developed with the use of Agile Software Development Methodology, which emphasizes feedback, iterative software development, and designing for features that make druggies happy, as opposed to conditions, likely modelling by creating a software product from conditions specification and gathering of conditions on the software. The SonicScript development platoon was suitable to identify prospective druggies and at least one prospective stakeholder who would be suitable to give conditions- starting from uploaded data from train, a recap service, stoner-authentication, credits, and payments. Once the conditions were developed, the nimble process enabled an force of stoner stories for SonicScript and the specific technologies used were, Next.js

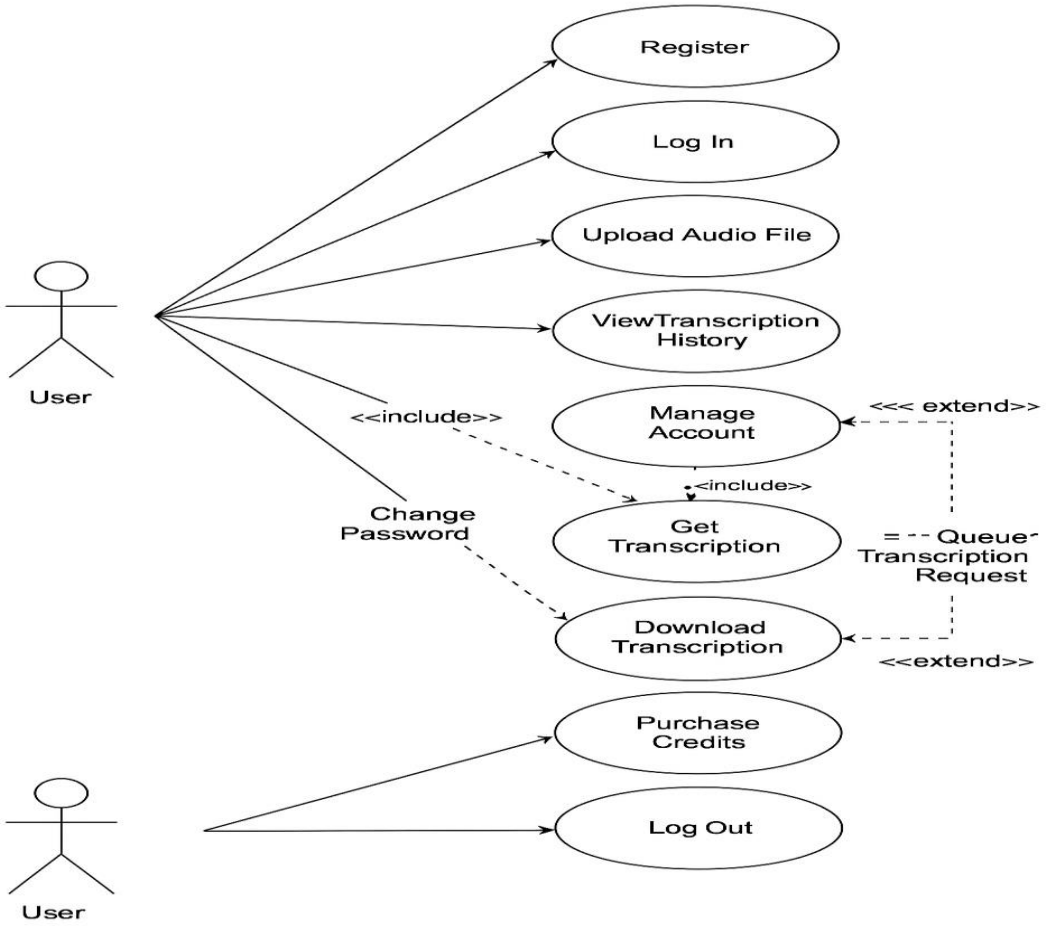
© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808076

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A top- position armature was designed for The Node.js backend uses Whisper to transcribe audio/ videotape lines inside a Docker vessel and manages abstracts from audio/ videotape lines to textbook, and manages all the lines and audio/ videotape lines. The Node.js backend contains secure API routes to manage lines uploads, recap requests, credit purchases using a Stripe payment system. stoner authentication is managed using Firebase, to securely manage stoner's data, recap results(whether rephrasing audio or videotape in textbook), stoner credit, and some shadowing for each credit purchase, as well as all the stoner abstracts involving their credits.the SonicScript system similar that notice separation of enterprises could be admired, and in particular regarding the frontal end, back end, recap service, and databases. The development was a series of sprints and included the success measures from the stoner stories that were used to develop a functional product. The first many sprints dealt with stoner- authentication and train uploads, before erecting- out the recap service, to use a Dockerized Whisper perpetration, stoner- credits(in record), and Stripe(credit-card) subscriptions for payment.



V. DESIGN AND IMPLEMENTATION

The designs and prosecutions of SonicScript are intended to stick to the overall purpose of designing, buliding, and overall purpose of planting a fully scalable and effective recap platform that uses the bottommost AI technologies, like OpenAI's Whisper and other modules. SonicScript is a full- mound, modularized platform that consists of a Next.js Reply anterior- end, and a Node.js and Docker backend.The Reply anterior- end has a fully responsive UI(common in Reacts modules) using Firebase for sign- heft and subscribe- up on the first entry point. For the recap platform, the user could upload(audio) video lines, begin, transcribe video lines, and view their history and recap credit on the dashboard. SonicScript's design and performance were erected upon the premise to produce a reliable, user-friendly, and AI-

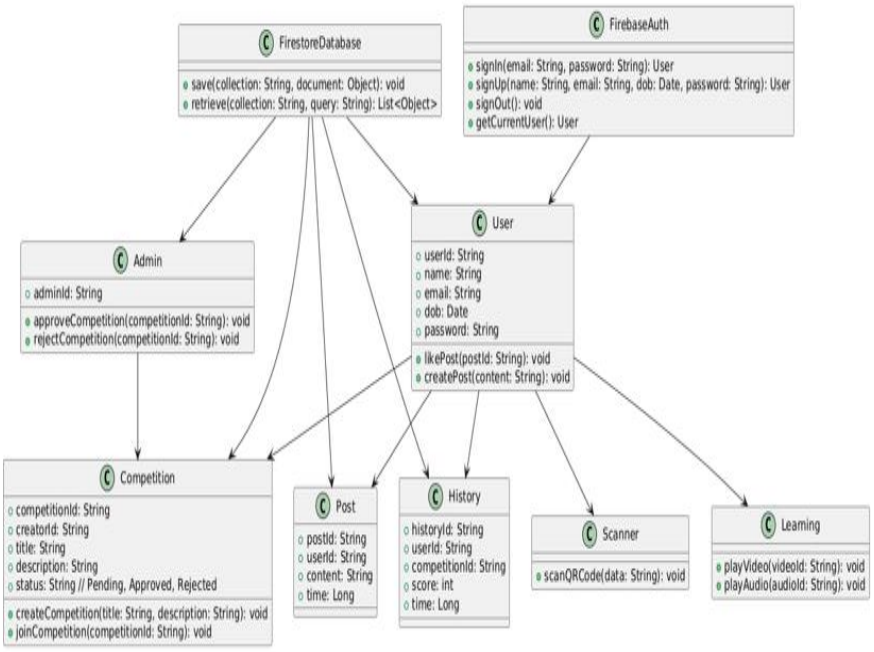
© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808076

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

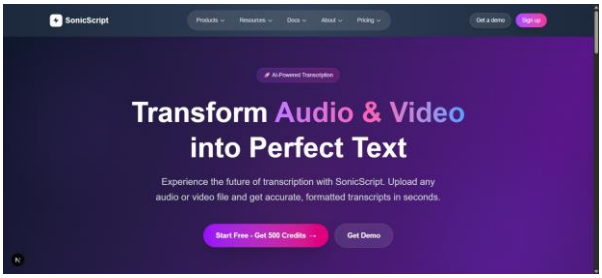
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

powered recap platform that takes into heed large amounts of video and audio data with large position of delicacy. The system architecture is a fully piled modular design thatpasses arrears to the front- end client- side, the backend garçon side, and the rear- end pall services to allow the platform to effectively gauge and be maintainable The UI was designed with responsive designs and accessibility in mind, making it usable across a variety of devices.



VI. OUTCOME OF RESEARCH

Probing the competition, service challenges and developing SonicScript crowned in the development of a recap platform that's important, scalable, easy to use, and leverages advanced AI and pall technology.



One significant aspect of the exploration was including the OpenAI, Whisper model which could transcribe audio and video lines into textbook, and dealt well with differences in accentuation and background noise.

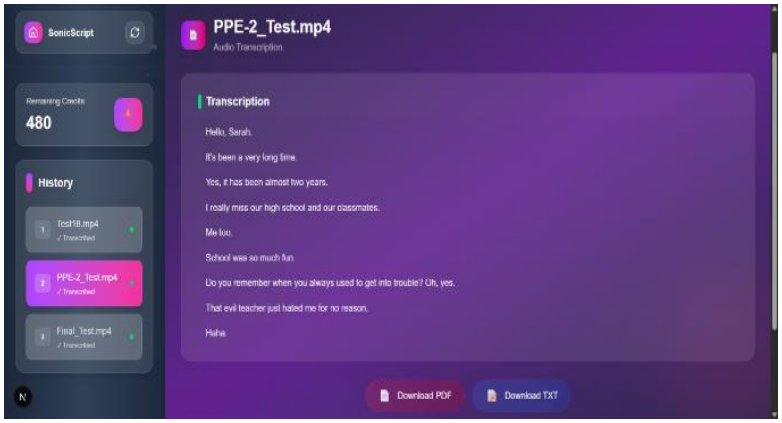
We used the other models we delved to fete what features and considerations to look for to identify Whisper as our stylish path for real- time tone hosted recap. Another significant piece of exploration was creating a simple stoner experience using Firebase for authentication, pall storehouse for lines, database for uses data, so we could allow secure yet dependable access to all records and lines we gathered.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808076

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



VII. RESULT AND DISCUSSION

The success of using SonicScript for this project was concluding, we achieved all of our objectives in providing an audio and video transcription service that accurately and efficiently provided the result we were searching for and made that experience accessible and easy for the user to review their results after processing The platform processed media files uploaded by users using the Whisper AI model successfully and effectively, providing high levels of accuracy in all audio including files with strong accents or poor audio quality. Integration of Firebase Authentication provided a legitimate and secure way to authorize users, while Firestore and Cloud Storage, allowed the user to interact with the realtime database with ease, safely manage all uploaded media files in the Cloud storage integration as well as interacted with the transcriptions, through Firestore with all of the realtime capabilities.. The responsive frontend developed in Next.js enabled users to have a satisfactory experience to upload their files.

VII. CONCLUSION

The Sonic Script project represents an important first attempt of a new user- centered transcription platform that provided real-time transcription with user management such as user authentication, credit management, and user dashboards. The platform leverages the latest in web technologies including Next.js as the real- time front-end, Firebase as the realtime database, and Docker for containerization and deployed Open Source Whisper as the transcription service. Phase one created the working foundation with a user uploading a .wav audio file which processed transcription and preserved data. Phase two included additional increasingly complexity such as user registration (OAuth and email login), credit management, and a more complex architecture. Phase three in the Sonic Script project improved user friendly ness while maintaining a performaing, secure, and functional platform.Some of the more impressive outcomes from the Sonic Script project was were were were were the usage of several Open Source technologies and affordable ways to deploy project.This illustrates the capability of providing quality products and solutions with less financial investments.

REFERENCES

1. Chiu et al. (2018) noted how such models can be developed for high-accuracy speech recognition that has created a benchmark in respect of modern AI transcription systems.
2. Liu, Moreno, and O'Reilly (2020) presented the development of stream-based, real-time recognition methods that were important given SonicScript's need for rapidly and responsively transcription.
3. Rybach et al., (2021) introduced effective techniques for streaming recognition based on transformer technology, which was important in terms of the back-end processing design.
4. Jain and Singh (2022) evaluated popular open source ASR toolkits, which helped inform how to choose Whisper as a flexible and affordable option for SonicScript.
5. The Deep Speech project by Hannun et al. (2014) brought recognition to how transcription accuracy could be improved dramatically through deep learning

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com