Voice Controlled News Web Application with Speech Recognition using Alan Studio

Aaditya Chaprana
Department of Computer Science and Technology
Meerut Institute of Engineering and Technology

Ranjeet Kumar, PhD
Department of Computer Science and Technology
Meerut Institute of Engineering and Technology

Ajay Saini
Department of Computer Science and Technology
Meerut Institute of Engineering and Technology

Akash Kumar
Department of Computer Science and Technology
Meerut Institute of Engineering and Technology

ABSTRACT
In this paper, a very fascinating proposition is presented, a new web-based service that is a fusion of the revolutionary new Alan Studio, News API and React. The proposed idea of the Voice Controlled Web Application provides a simplistic approach and ease to the user. The service provides all the components required for a user to be able to use voice and speech as a medium to find and look for news about his/her choice and the option to go through the news in a very concise or in very detailed manner. A voice-controlled system embedded in a web application can enhance user experience and can provide voice as a means to control the functionality of e-commerce websites. The system also offers the ability to search news on the basis of preferred location, source, theme, and interest. The system provides a very user-friendly, easy to use, dynamic, and informative user experience. The system enables the user to have the ability to not only stay informed and updated but that too by using one of the most sought-after technologies in the world right now. The system will reduce the amount of human effort required by the user to perform previously and will offer a more exciting way of getting informed.

Figure 1 Working Diagram of the System

Keywords
Alan Studio, News API, React, Voice Controlled Web App.

1. INTRODUCTION
To make life easier, advancements are being made in technology, and voice control is one of the leading technologies which are gradually being implemented on more and more devices. Voice control has become one of the most in-demand skills. More and more applications and services provide voice control capabilities. Voice control capabilities provide a very easy and hands-free experience for the user and enables him to use the service without having to physically use the device with his/her hands.

The modern lifestyles do not allow us to take time from busy schedules to sit down and read newspapers, magazines etc to get us informed about the latest events happening all around the world. With each successive generation, the society struggle even more to stay

Informed and educated about the world. Reading and attaining the latest news was a well-practiced habit that was followed by a huge part of the society that seems to get forgotten gradually. With the involvement and introduction of technology, it is observed that new ways of news reading are getting replacing old methods. News is manually customized to grab the attention and notice of its reader and is even available in a short, to the point format. There are various major, big news apps that give e-papers and a gist of news as notifications or updates for their readers and users. With every new story or article that is read, the reader gains knowledge about the events taking place across the globe. News reading is the quickest and most effective way to gather knowledge about state and global news. Since news agencies cover all subjects of interest like Fashion, Lifestyle, Fashion, Politics, Sports, Entertainment, and more, the reader is constantly updated about all these facts.

The traditional methods of data entry fail the requirements to support all type of users. Thus, it is necessary to develop systems and applications with enhanced usability for all users. Most of the current applications lack the accessibility features that may hinder some users for example the visually impaired users.

The hands-free approach provided by the system goes to a great length and makes the user interact more often as the user usually prefers to use voice command rather than giving commands by typing. One of the biggest advantages of the proposed system is that the voice recognition is not limited to

just mobile phones, laptops or computers but voice recognition is being installed in all type of devices that users interact with like smart televisions, smart watches etc.

Voice recognition is also being installed in vehicles. Cars and other automobiles are finding the benefits of voice recognition and voice command. The system is very easy to use, very user friendly and the user has the choice of language he wants to interact with the system.

The voice control technology is reaching new levels day by day and the system will also help in tackling one other problem that the normal user faces every day. Apart from some, most people face trouble while typing in a regional language. Most keyboards do not provide enough compatibility and comfort while typing in a regional language. By using the system, the user has the free choice of using the language he prefers. This helps in faster result searching, better user experience and better user satisfaction.

The system also takes care of the user's privacy. The system does not gets activated on its own and only gets activated when it's gets prompted by the user. The privacy of the user's data is one of the key concerns. The data stored in the system is entirely safe; it only records the news sources, news type etc, only to make the system feel more personal and more customized to the needs and liking of the user. After the development of the proposed idea, the user will be able to tackle the above-mentioned problems and would also be able to get the benefits from one of the fastest developing technologies in the technical industry at this moment of time.

According to various research groups and studies, 2019 saw about 111.8 million people used the voice assistant at least one time in a month in USA alone. The projected figure of the number of people using the voice assistants in USA by 2021 is around 132 million. The ever-increasing number of people using this technology just gives us the advantage and also a need for the likes of the proposed system.

Web apps equipped with voice enabled systems can provide flexibility in terms of users' choice of web interaction and can also increase the usability of the app for the general users when they are unable to interact with the system in the traditional methods.

2. LITERATURE SURVEY

Technically speaking, speech recognition goes way back to 1877 when Thomas Edison invented the phonograph, the first device to record and reproduce sound. Voice Recognition is used by everyone using a smart phone or a smart device. It has limitless scope and can be used in different ways. For example:

- Design of an Intelligent Voice Controlled Home Automation System by Sonali Sen, Shamik Chakrabarty, Raghav Toshniwal, Ankita Bhaumik [6]
- Voice controlled surgical suite by David F. McCall, Leslie M. Logue, Francis J. Zelina, Matthew V. Sendak, Julie R. Hinson, Ward
- L. Sanders, Steve Belinski, Brian E. Holtz [7]
- Preventing false wake word detections with a voice-controlled device by Ian W. Freed, William Folwell Barton, Rohit Prasad [8]
- Voice controlled wireless communication device system by Stephen S. Burns, Mickey W. Kowitz [9]

The above stated improvements show the ability and future of voice recognition and control. In the proposed system, voice recognition is used in another field and it stretches its scope into a new dimension. The proposed system takes a set of well-established and new upcoming technologies and it enables user to set up a combination that provides a useful, exciting and user-friendly experience which helps user save time as well as physical work.

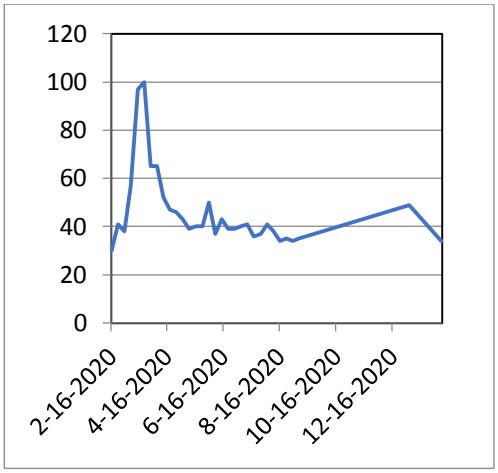

Figure 2 Google Trend Report for the term "Breaking News"

The above graph provided by Google Trends showcases that the term "Breaking News" is always among the most searched term worldwide. The term hit its peak in initial days of the Covid-19 pandemic as people wanted to know more about the latest news around the globe and within their own country. This showcases how reliant people are on their devices to get the knowledge of what's happening around them.

Table 1 Google Trend Report for the term "Breaking News"

January,2020	38
February,2020	57
March,2020	65
April,2020	43
May,2020	37
June,2020	40
July,2020	41
August,2020	38
September,2020	35
October,2020	49
November,2020	50
December,2020	35
Januray,2021	34

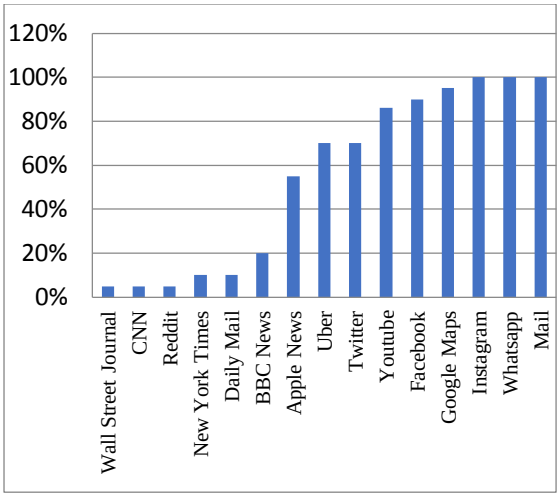

Figure 3 Apps Found in Mobile Phones

Reuters study of people between the ages of 18 and 35 in New York and London showcases that the amount of people that have a NEWS APP installed in their mobile phones is very less. People often rely on searching the news on their web browser rather than having a single dedicated app installed on their mobile phones. The system keeps track of this feedback from the users and does not force them to install anything to use the system.

Table 2 Apps Found in Mobile Phones

Wall Street Journal	5%
CNN	5%
Reddit	5%
New York Times	10%
Daily Mail	10%
BBC News	20%
Apple News	55%
Uber	70%
Twitter	70%
YouTube	86%
Facebook	90%
Google Maps	100%
Instagram	100%
WhatsApp	100%
Mail	100%

Table 3 Most downloaded apps in January 2021 (Non game)

	Overall Downloads	App Store Downloads	Google Play Downloads
1	Telegram	TikTok	Telegram
2	TikTok	YouTube	Signal
3	Signal	Zoom	TikTok
4	Facebook	Telegram	MX
5	WhatsApp	Facebook	Facebook
6	Instagram	Instagram	WhatsApp
7	Zoom	WhatsApp	Instagram
8	MX	Messenger	Moj
9	Snapchat	WeChat	Snapchat
10	Messenger	Signal	Zoom

The above table provided by Sensor Tower, shows the most downloaded apps in the two most popular app markets. The data clearly indicates that no News app has managed to get into the top 10, thus showing that users do not like to install news app and rely on web searches. This helps the system, as it does not ask the user to install anything in order to use the system.

Many studies have been conducted where the researchers have used the SRS functionality with the applications to make them usable by visually-impaired users. In an early study, Kevin Christian et al. [10] compared the performances of voice controlled and mouse-controlled web browsing. Based on the ratings of the 18 test subjects who used both methods concluded that the textual links are preferable to numbered links. Bajpai at el [11] proposed a system where a SRS was incorporated in a browser, the system would accept users' voice as input and based on the predefined set of operations, the browser would perform the operation.

The idea of the voice operated browser was extended by Han et al [12] such that it could be used in a smart TV. The research focused on navigating and controlling a dynamically generated hierarchical menu on a webpage with voice keywords. Sagar et al. [13] proposed an application where email services are combined with speech to text services and text to speech thus enabling users to write, send, and read their emails.

3. METHODOLOGY

The following figure depicts how the system works. It shows how the system interacts with the user and how and which processes take place in various cases.

The user provides a voice input to the web app. The app sends the signal to the Speech to text service. The STT service coverts the voice signal and converts it into text which is then sent for the keyword matching. If the keywords are matched then the functionality that the keyword is related to is performed but if the keyword is not matched then an error message is sent to the web app and then to the user. After the task is performed, the results are then shared to the web app which then replies to the user with a voice reply.

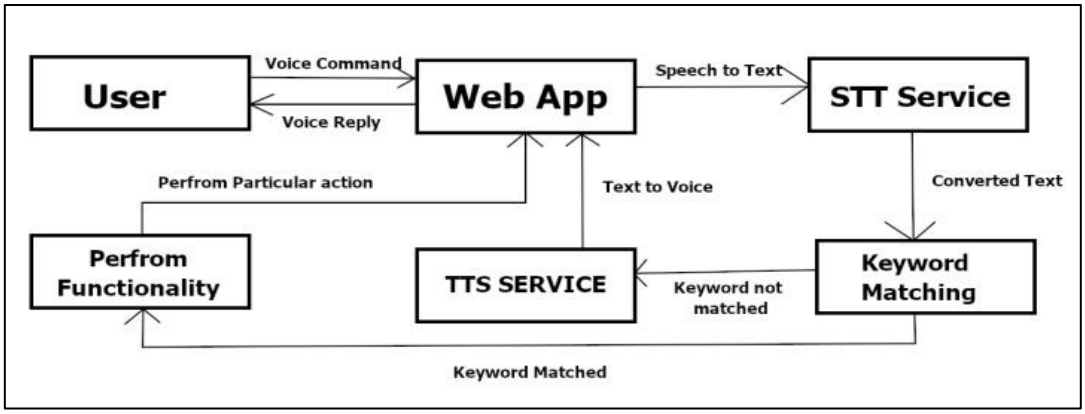

Figure 4 Architecture of Web Application

3.1 Alan Studio

Alan adds an entire serverless environment to construct complex and trustworthy in-app voice assistants and chat bots. Alan studio is a great tool that helps in seamless voice recognition. There is no need to make spoken language models, train the speech recognition software, launch and host voice parts — the Alan AI backend does the most of the work. The voice experience for any app can be constructed and made by a single developer, rather than a team of Machine Learning and Dev Ops experts.

With Alan, you can go beyond the abilities of touch and type interfaces and voice enable any complex workflow or function in any app. Voice scripts are constructed in JavaScript, which makes them heavily modifiable and flexible. The following flowchart depicts how Alan Studio works.

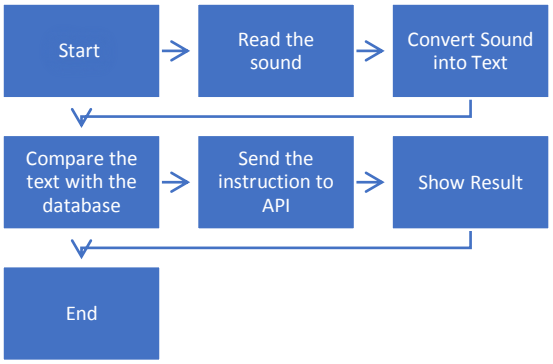

Figure 5 Working of Alan Studio

The performance of a SRS depends on the size of the vocabulary it supports for a certain level of accuracy, complexity, and information processing rate. Alan Studio has a very large sized vocabulary and thus STT and TTS generate great results. The other important aspect of a great SRS is that it is speaker independent that is the application can support voices and the pronunciation of words on a global level thus making it useful for a greater range of audiences.

The speaker independent recognition is more difficult because the internal representation of the speech must be global so that it can cover all types of voices and all possible ways of voice pronunciation and yet specific enough to find the differences between various words.

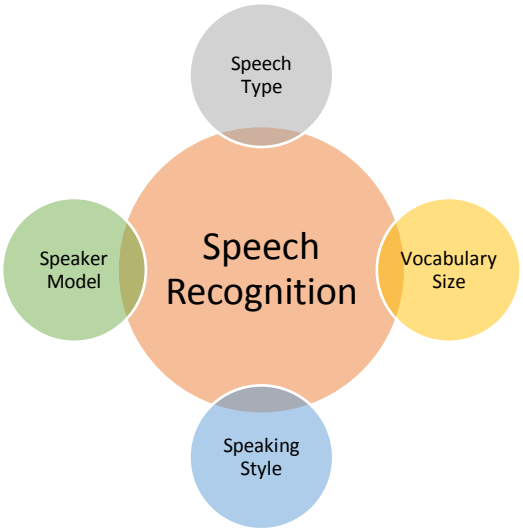

Figure 6 Taxonomy of SRS

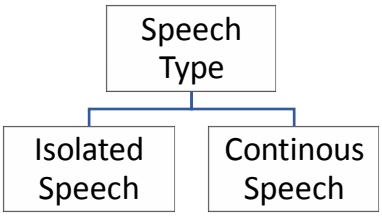

Figure 7 On the Basis of Speech Type

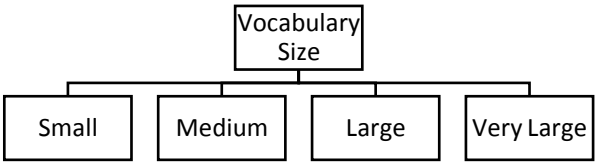

Figure 8 On the Basis of Vocabulary Size

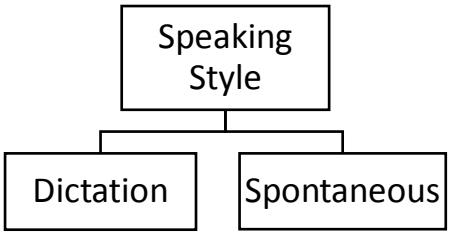

Figure 9 On the Basis of Speaking Style

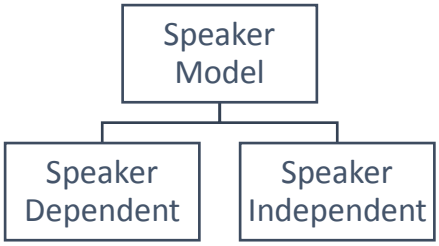

Figure 10 On the Basis of Speaker Model

3.2 React

React is a JavaScript library for developing user interfaces. React makes it effortless to create responsive UIs. Design basic views for every state in any application, and React will accurately update and render just the correct parts when user's data updates. React able user to construct encapsulated parts that manage their own state, and then combine them to make dynamic UIs.

Since component logic is constructed in JavaScript instead of templates, you can easily pass detailed data to any app and keep state out of the DOM.

React has been made from the start for continuous adoption, and you can use as few or as much React as you need**.** Whether you want to get a gist of React or include some features to a basic HTML page, or start a React-powered app, the links in this section will help you start.

3.3 News API

News API is a basic HTTP REST API for finding and collecting live articles from all over the web. It can help any simple queries like:

- What famous stories are TechCrunch currently?
- What latest articles were published about the next iPhone?
- Has my company or product been mentioned or reviewed by any blogs recently?

You can find articles with any combination of the following criteria:

- Keyword or phrase. Eg: find all articles containing the word 'Apple'.
- Date published. Eg: find every article published today.
- Source name. Eg: find every article by 'Mi'.
- Source domain name. Eg: find every article published on thenextweb.com.
- Language. Eg: find every article written in English.

You can sort the results in the following orders:

- Date published
- Relation to search keyword
- Popularity of source

News API is ample as a data source for news tickers and other applications where the user wants to display his live headlines. It keeps track of headlines in 7 departments across over 50 nations, and at over a thousand best publications and blogs, in near real time.

4. DESIGN IMPLEMENTATION

The above-mentioned technologies combine together and offer the user a perfectly working system that enables him to search and browse through the latest updates, news and articles from various sources, locations and topics. The system works very precisely, offering the user news from the most reliable and trusted sources.

Alan studio offers the voice recognition and delivery capabilities and analyses the vocal instructions provided by the user and identifies the keywords and the task that the user has asked. The selected keywords are passed through to the NEWS API that searches for the news articles about the selected keyword and returns the json file containing all the textual data that the articles contain.

The Material UI helps to extract the data from the json file and it passes the extracted data to the Alan studio which once again reads and analyses the data and converts into voice and reads the articles to the user.

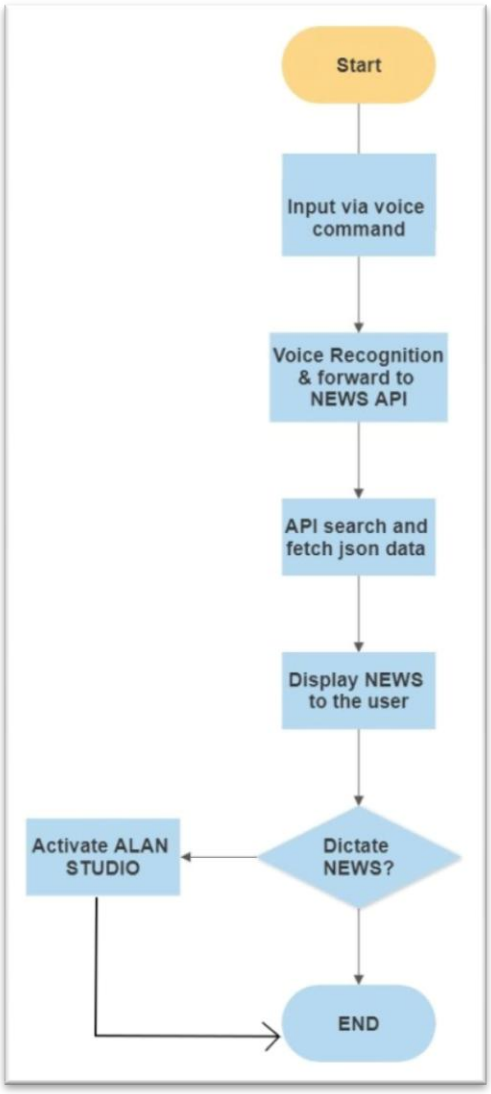

Figure 11 Working Flowchart of the System

5. EXPERIMENTAL RESULTS

The proposed project offers a new, faster, reliable, easy to use and a user-friendly experience that helps the user to stay informed and stay updated with the events happening around the world. The system is able to identify between the various keywords being said by the user and thus able to return the news according to the query provided by the user. The system distinguishes various sources, locations, topics and themes of the news that might be interesting for the user.

The above-mentioned data showcases how the system performs when various lengths of search quires put into the system. The system identifies number of keywords and shows the results according to the keyword identified.

As the number of words increase the number of keywords also increases and the system also fetches news articles according to the keywords. Thus, a variety of news articles are fetched which may not be related to the overall search query but related to the particular keyword identified by the system. But still the system manages to fetch a greater number of related articles as compared to the number of unrelated articles fetched.

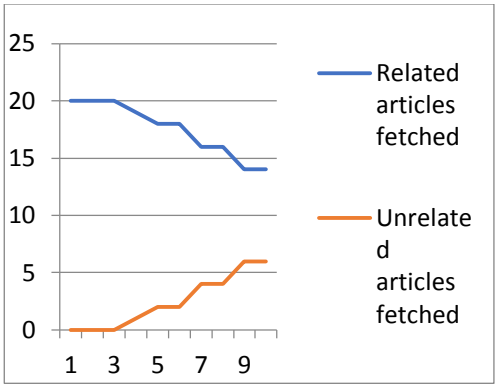

Figure 12 Graph Articles fetched vs. Number of words searched

The above graph also indicates that despite the number of related articles decreases slightly as the user increases the number of words searched and keywords identified by the system, the number of related articles fetched by system is more than the number of un-related articles fetched by the system.

6. SCOPE

Table 4 Experimental Results

Number of words searched	Number of keywords identified	Related articles fetched	Unrelated articles fetched
1	1	20	0
2	1	20	0
3	1	20	0
4	1	19	1
5	2	18	2
6	2	18	2
7	3	16	4
8	3	16	4
9	3	14	6
10	4	14	6

The system fits the needs for any organization or individual usage.

As discussed earlier, the amount of the people using voice command is increasing day by day. The number of devices supporting voice command is increasing day by day. The system is one of the examples of how this technology can be used to its maximum benefit and shows how brilliant and useful the voice command technology truly is.

As the number of users increase, the needs of the users will also increase. The users would want more and more apps and systems to be activated and used by the voice command. The system provides with one of the most sought after demands of the users. It will provide news on the go, with minimal physical interaction with the device, faster response and more user-friendly feel and environment.

This system can be used in a variety of ways in a different way in different devices. The system will easily help people who have less time and want to keep themselves updated.

The system could be installed, used and modified in a variety of ways and could be put to use for a variety of purposes depending upon the needs and requirement of the person using the system.

7. FUTURE WORK

When voice recognition began to emerge in the early 2010's with the introduction of what now is one of most recognized devices, Siri, no one was expecting this would become such a driving force for future technical innovations.

It is estimated that $1/6^{th}$ of the United States' population now owns a smart speaker. This just gives us a mere glimpse of the reach of this technology and the vast land of opportunity the system has.

One of the biggest estimations in the field of voice recognition for 2020 was that the search behaviours of the people are going to change. Brands are now experiencing a shift in which touch points are converting to listening points. Visual Interface with voice assistants is missing. The system will provide both visual and voice interaction to the user. Thus, the user will be able to get an entire user-friendly experience in which he is able to see and listen to the content that he has asked for and thus getting more engaged with the system and technology.

The technology is emerging all over the world and in the coming years, voice recognition is expected to grow and will create enormous revenues. As the system is working on most of the requirement for now, the system will be added number of functionalities and capabilities.

The system would support a greater number of languages supported by the system, thus adding even more comfort and ease for our user.

CSE 2019 proved that voice and visual displays are coming together into one seamless experience. The need to bring visual and voice capabilities together has been in the market for a long time and the system is a stepping stone in that direction, making use of the voice and display capabilities of the device in a unique and interacting manner and way.

The system would also become more and more user friendly while keeping track of the most favoured news source and news category.

8. CONCLUSION & DISCUSSION

The Proposed system enables a wide range of users to stay informed and updates while using as less time as possible. It makes getting informed and knowledgeable easy and very interesting. People with limited time now can easily get up to date just with the help of a few vocal commands.

It also helps physically challenged people to make use of the latest advancements in the technical fields and enables them to stay updated and informed without their health condition hampering them.

The system also enables user to listen to the articles that grab our interests and those that it thinks that are important for the user to know. The proposed system is a fine example of how one of the most sought-after features of the modern device can make our lives much easier and help us save both time and the physical work to stay informed.

The system will continue to develop and more and more functionalities cam be added to it. The number of efforts going into the development of voice recognition and voice command adding with the continuously increasing demand of the users, the system will never get outdated. The system is flexible, user interactive and compatible to almost every device that has the ability of voice command.

9. REFERENCES

[1] Advantages of Voice Recognition Technology Retrieved from https://www.insidetelecom.com/advantages-and-drawbacks-of-voice-recognition-technology/

[2] reading news is beneficial Retrieved from https Why://www.paperboy.com/blog/top-5-benefits-importance-of-reading-news/

[3] Alan Studio: https://alan.app/

[4] React.: https://reactjs.org/

[5] NEWS API : https://newsapi.org/docs

[6] Sonali Sen, Shamik Chakrabarty, Raghav Toshniwal, Ankita Bhaumik. Design of an Intelligent Voice Controlled Home Automation System. International Journal of Computer Applications 121(15):39-42, 2015.

[7] David F. McCall, Leslie M. Logue, Francis J. Zelina, Matthew V. Sendak, Julie R. Hinson, Ward L. Sanders, Steve Belinski, Brian E. Holtz. Voice controlled surgical suite. Google Patents US6591239B1, 1999-Present.

[8] Ian W. Freed, William Folwell Barton, Rohit Prasad. Preventing false wake word detections with a voice-controlled device. Google Patents US9368105B1, 2014.

[9] Stephen S. Burns, Mickey W. Kowitz. Voice controlled wireless communication device system. Google Patents US7957975B2, 2016-Present.

[10] K. Christian, B. Kules, B. Shneiderman and A. Youssef, "A Comparison of Voice Controlled and Mouse Controlled Web Browsing," Proceedings of the Fourth International ACM Conference on Assistive Technologies, no. ACM, pp. 72-79, 2000.

[11] . B. BAJPEI, M. S. SHAIKH and N. S. RATATE, "VOICE OPERATED WEB BROWSER," *International Journal of Soft Computing and Artificial Intelligence*, vol. 3, no. 1, pp. 30-32, May-2015.

[12] S. Han, G. Jung, B.U. C. Minsoo Ryu and J. Cha, "A Voice-controlled Web Browser to Navigate Hierarchical Hidden Menus of Web Pages in a Smart-tv Environment," Proceedings of the 23rd International Conference on World Wide Web, pp. 587-590, 2014.

[13] S. Sagar, V. Awasthi, S. Rastogi, T. Garg, S. Kuzhalvaimozhi, "Web application for voice operated e-mail exchange",

[14] https://blogs.oracle.com/marketingcloud/advantages-and-disadvantages-of-voice-assistants-for-marketers

[15] https://clearbridgemobile.com/7-key-predictions-for-the-future-of-voice-assistants-and-ai/

[16] https://trends.google.com/trends/explore?cat=16&q=%2Fm%2F052j8s

[17] https://www.fipp.com/news/chart-week-do-young-people-news-apps/#

[18] https://startuptalky.com/most-apps-downloaded/#