Notes Keeper Application Using Speech Recognition

Priya Maurya[1], Thomas Abraham[2], Dr. Sudeept Singh Yadav[3]

[1,2,3] Department of Computer Science and Engineering, Galgotias University, INTEGRATED BCA+MCA Greater Noida

Abstract— **Notes Keeper is an awesome internet based scratchpad, meant to permit you draw your thoughts via audio and voice.**
The cause of the mission is to produce web software to help with taking and coping with notes through the use of speech or audio or via simple textual content. The utility will allow individuals from all domain names inclusive of studies, lecturers, commercial enterprises, and others to take notes successfully and manipulate all factors of note-making.
The utility can be introduced using JavaScript and Web APIs to decorate the ease of taking notes, the usage of audio transformed text or easy text. It will increase the rate of taking notes. The software has loads of destiny scope because the time could be decreased substantially from taking notes manually to audio to textual content conversion.

Index Terms: **Notes keeping, javascript, speech recognition**

I. INTRODUCTION

Notes keeper application is a single-page web-based compact application to save notes and take notes and recover them for some time later.

Notes taking and keeping used to be very hectic and treacherous tasks. Originally, we used to make notes on the paper that would usually get lost or get torn in the process with time.

With the development of the computer, people started inventing things and notepads got into the making and usage. People had to decide and choose from different kinds of type typewriters and different keyboard layouts.

Applications used to be confusing and heavy. With the development of front-end technologies like JavaScript and React it became easier to make a single page and lightweight applications.

JavaScript (JS) is an insubstantial, interpreted, or just-in-time compiled coding language with the finest capabilities. While it is famous because of the scripting language for Web pages, many non-browser environments additionally use it, together with Node.Js, Apache CouchDB, and Adobe Acrobat. JavaScript is a prototype-based, multi-paradigm, unmarried-threaded, dynamic language, helping item-orientated, vital, and declarative (e.g. Functional programming) patterns.

The organization of JS modules is often achieved by using asynchronous module definition, a JS API applying the "inversion of control" pattern.[1]

Speech to text or speech recognition is the way of converting the speech or voice into the text form.

Speech recognition, additionally called automatic speech recognition (ASR), computer speech reputation, or speech-to-textual content is a functionality that allows the software to technique human speech into a written layout. While it's usually stressed with voice recognition, speech recognition focuses on the interpretation of speech from a verbal layout to a textual content one while voice popularity simply seeks to pick out a person's voice.

Speech recognition is done in several ways. Using web speech API is one of them. A Web API interface is available assuming it answers HTTP demands utilizing one of the predefined HTTP status codes[13].

Applications typically consume both, these services and backends, using web APIs – application programming interfaces invoked over networks that rely on web technologies like HTTP as a transport protocol or XML and JSON as data formats. In practice, these APIs are often"REST-like", in that they adhere to some of the constraints imposed by the Representational State Transfer (REST) architectural style[8].

II. LITERATURE SURVEY

An excellent be aware-taking app should be speedy. It has to sync your notes through a couple of gadgets quickly and appropriately. It ought to be available regardless of what computing platform you operate

now or may transfer to within the destiny. It ought to have lightning-rAPId seek throughout all of your notes. It shouldn't be slowed down with too many perplexing capabilities — however, it can also be so bare-bones that it won't meet your needs. It needs to shield the safety and privateness of your notes. It has also been inexpensive, or even loose.

Note-taking packages (additionally referred to as observe-taking apps) permit us to:

Store all notes and important facts digitally, generally in a cloud-based garage gadget.

Type, write and draw notes at the device of desire just as one could the use of pen and paper.

Add files, multimedia, and live recordings to their notes to enrich the meaning and context.

Collaborate and percentage notes with others right away and in real-time.

Studies advocate those texts produced by means of speech-to-text recognition (STR) generation from speech are beneficial for college kids, allowing them to attain a better understanding of content material.

Speech-to-Text acknowledgment (STR) innovation has acquired consideration as of late in the research literature. STR synchronously generates text from a lecturer's speech and displays it for students on a whiteboard or computer screens. Late proof recommends that utilizations of STR innovation are valuable for learning (Huang, Liu, Shadiev, Shen, and Hwang, 2015). For instance, during addresses, STR-text assists understudies with beating troubles in perusing, composing, and spelling (Nisbet, Wilson, and Aitken, 2005), helps support of hard of hearing understudies (Leitch, 2008; Wald and Bain, 2008), works with perception of learning content (Ranchal et al., 2013),especially of non-local speakers (Ryba, McIvor, Shakir, and Paez, 2006; Shadiev, Hwang, Huang, and Liu, 2016; Wald and Bain, 2008) and understudies in web-based coordinated learning conditions (Hwang, Shadiev, Kuo, and Chen, 2012; Kuo, Shadiev, Hwang, and Chen, 2012)[3]

In recent times, technology has also made existence a good deal less complicated for humans dwelling with disabilities to use computer systems. One such innovation is software primarily based on voice popularity, which has made it viable to dictate and compose files completely palms-free. It offers users the choice to use a speech-to-text device because these users communicate to the laptop and their enter is interpreted and transformed into electronic text. It

is a particularly dazzling product because it was first developed for those who would instead use their voice to operate a laptop. Therefore, no longer the most effective software assists the visually impaired, but additionally human beings with mobility issues and other bodily disabilities.

Overall, voice recognition has the capacity to be a robust method for anyone who cannot use a keyboard because of a visual or physical incapacity. Web Socket APIs innovation is utilized to make a two-way correspondence meeting where program gets occasion driven reactions without surveying the server for an answer[15].

Google's Speech API is often used to add speech recognition and synthesis to language learning activities. Nowadays, language learning activities are frequently developed using Google's Speech API. Application developed using Google's web speech API includes automated oral assessment applications, voice recorder and voice and audio shadowing using speech recognition .

While Web APIs indicate a rundown of properties a stage article ought to give, following what properties a given stage object gives is nontrivial[12].

Main findings show that two thirds of the companies use external services, but also about two thirds of the services used are internally developed and operated. REST and HTTP are quite present, while SOAP rarely appears. JSON is more common than XML and the most often used programming languages for implementing the companies' internal REST services are Java, JavaScript, C#, Python and Ruby[9].Web API descriptions can define the set of supported keys and the expected types of values[10].

A qualitative analysis of the results from checker on 6575 requests shows that most of reported inconsistencies were due to errors in the client code (calls to deprecated APIs, errors in the URLs, errors in data payload definitions) and incomplete Swagger specifications, as opposed false positives.[6].

III.DESIGN OF PROPOSED SYSTEM

The proposed system is automated and has been developed using the latest language therefore it gives more competence than the present system. It can take notes in the form of text or it takes speech input and transforms it into the text using web speech API

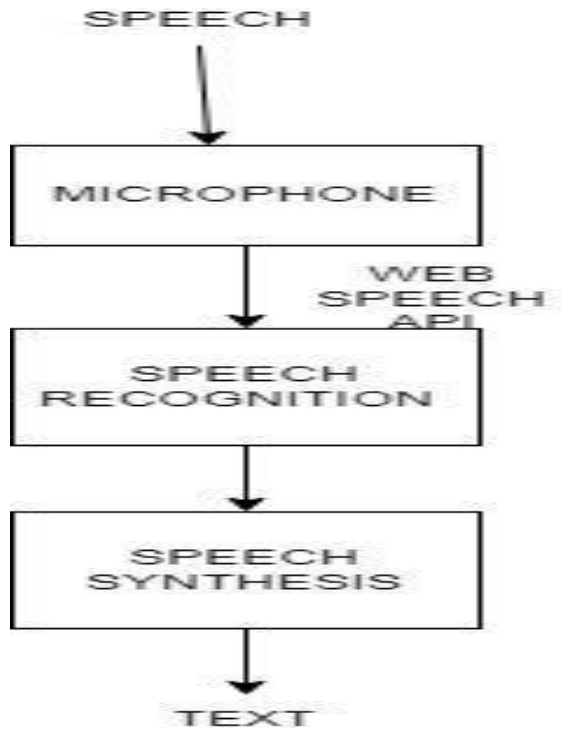integrated into the JavaScript application and stores it into the local storage.

Fig 1 : flowchart
Figure 1 flowchart shows the flow of audio input

Technology Used
We have used a JavaScript engine that's used to support the JavaScript scripts so as for them to be painted properly. CSS deals with the designing aspect of the online site. We'll use the analogy that HTML is the noun of web development whereas CSS is the adjective for HTML. JavaScript (JS) is an insubstantial, interpreted, or just-in-time compiled coding language with the finest capabilities.

A. Front End (HTML5 & CSS3)
Hypertext Mark-up Language is the essential, fundamental language to develop the structure and outline of the online site. CSS deals with the designing aspect of the online site. We'll use the analogy that HTML is the noun of web development whereas CSS is the adjective for HTML.

B. Back End (JavaScript, Web Storage (SQLite))
Each browser has its own JavaScript engine that's used to support the JavaScript scripts so as for them to paint properly. The simple activity of a JavaScript engine is to take the JavaScript code, then convert it into a fast, optimized code that may be interpreted by means of a browser. Below are the names of the JavaScript engines utilized in some of the most famous browsers obtainable.

As web technologies like JS become ever more important for the frontend layer of enterprise applications, these decisions must also be evaluated for the use of libraries and frameworks to provide fast and highly interactive applications on the web.[1]The JavaScript code responds to client occasions and asynchronously sends HTTP requests to the server for refreshing or recovering information[11].

The Web Storage API is complex and quick with instruments that permit programs to keep key-cost matches. It is intended to be substantially more amiable than the use of treats.

The key-value matches address capacity devices, which can be similar as articles with the exception of they stay in salvageable shape all through page stacks, and are consistently strings. The files are stored in SQLite files in a user's profile making it easier to store and retract files on the go.

C. Speech Recognition API
The Web Speech API grants us to involve voice records into net applications. The Web Speech API has parts: Speech Synthesis (Text-to-Speech), and Speech Recognition (Asynchronous Speech Recognition).

The info signal is first handled by a fringe unit like a range analyzer. It then goes through decrease in two examination by-blend circles, and the phoneme grouping shows up at the right. In order to simplify the diagram, the group of components performing the functions of storage, preliminary analysis, comparison, and control have been combined in a single block labeled strategy.[7]
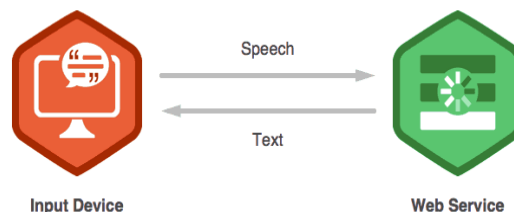

Fig 2: Working of Web Speech API
(source-https://blog.teamtreehouse.com/accepting-speech-input-html5-forms,online)

Figure 2 provides an overview of the working of the application. It depicts the conversion of audio input to text by web service.

D.   IDE:(Integrated Development Environment)
An IDE normally comprises a source code editor, compiler, interpreter, debugger/ Tester. Some examples are Visual Studio Code, Sublime, Atom, etc

IV. PROPOSED METHODOLOGY

The proposed system is computerized and has been developed using advanced language therefore it gives more facilities than the present system. It can take notes in the form of text or it can take speech input and convert it into the text using web speech API integrated into the JavaScript application and store it into the local storage of the browser.

FEATURES OF THE PROPOSED SYSTEM
This portion provides highlights of the system, its core components, and interfaces. They include:
1)   Starting the microphone and starting the speech recognition: The microphone is accessed using the Start recognition button.
Speech Recognition is utilized through the connection point for Speech Recognition, which gives the possibility to perceive voice input from a sound info (normally through the instrument's default discourse ubiquity administration) and answer precisely. For the most part, you can utilize the connection point's constructor to make a pristine Speech Recognition object, which has various occasion overseers accessible for distinguishing when discourse is input through the device's amplifier. The Speech Grammar interface addresses a case for a chose set of sentence structure that your application should perceive. Grammar is defined using JSpeech Grammar Format (JSGF).
Speech synthesis is accessed via the Speech Synthesis interface, a textual content-to-speech element that lets applications read out their text content material (typically through the tool's default speech synthesizer.) Different voice types are represented by using Speech Synthesis Voice items, and specific parts of the text that you want to be spoken are represented by using Speech Synthesis Utterance items. You can get these spoken with the aid of passing them to the Speech Synthesis. Speak() approach.
2)   Disabling the microphone: The microphone can be disabled after taking the input from the user. This will stop the further usage of a microphone.
3)   Taking text notes: Text notes can be taken using the text area and saved using the save and can be used further.
4)   Saving the converted data: The converted data can be stored in the local storage of the browser. The local Storage read-simplest belongings of the window interface permit you to get entry to a Storage object for the Document's origin; the saved facts are saved throughout browser classe.
The keys and the values stored with local Storage are always inside the UTF-sixteen DOM String format, which makes use of two bytes in step with the person. As with items, integer keys are robotically converted to strings. Local Storage facts are particular to the protocol of the document. In particular, for an internet site loaded over HTTP, local Storage returns a distinct item than local Storage for the corresponding internet site loaded over HTTPS.
5)   Playing the saved text as audio: The saved text can be saved as audio whenever needed.
6)   Deleting the existing notes: The existing notes can be deleted using the delete options from the list of saved notes that are of no further use or the user wishes to delete them.

V.COMPARISON BETWEEN OLD METHODOLOGIES

There are different applications present in the web environment for the sole purpose of notes taking and making the task easy. But these products are developed by multinational companies having thousands of employees dedicated to a single product.
These products are basically created to generate revenues and make a profit.
The products currently available were written in primitive languages such as -
Evernote, a very popular application, is written in C++.
However, C++ programs are regularly weighty if it's don't watch out, C++ program permits classes and accordingly the capacities with an identical name and over-burden works in this manner the image

damaging framework should be utilized, it can without much of a stretch be enveloped by C capacities.

C++ program has no thought of being quick and it's not utilized for stage subordinate applications any more extended than C or anything is. As a matter of fact, given the personality of the tool chain, it's presumably less reliable than others. Complex in an extremely enormous significant level program.

It is utilized for stage explicit applications regularly.

For a particular OS or stage, the library set is commonly picked that locks.

Having all the primitive features, it lacks innovation, creativity, and vision.

The application overcomes all the existing issues. It has a visionary approach of providing people with disabilities with an application that they can easily access and use on the go.

The application is written in the latest technology using the latest trends and needs.

There are heaps of JS libraries and systems with various extensions as can be seen plainly. The presentation of new non-standard usefulness in JS applications can typically be accomplished by incorporating a reasonable library. The libraries can be consolidated subjectively, e.g., distinctive gadget libraries can be utilized together. In comparison, an abundance of tools exist for statically typed languages, including sophisticated integrated development environments and specialized static analyses. Transferring such technologies to the domain of JavaScript web applications is challenging [5].

As of late, there likewise has been a shift of the use of JS to the backend side of the product engineering bringing about the chance of full-stack JS-based web applications. Practically speaking, a famous blend with the abbreviation "Signify" is arising as a common JS-based web application stack: MongoDB (information base), Express (backend system), AngularJS (frontend structure), Node.js (backend stage).

The application uses the latest Speech recognition feature for the conversion of audio input into the text and save them into the memory using the local storage of the browser.

The speech recognition API works on natural language processing.

Speech Recognition is being integrated into more devices and devices to be able to make existence less difficult as voice inputs are some distance more efficient than typing. Voice and Speech recognition era is improving each day and as in keeping with The University of Stanford, it has progressed to the extent where it can be a great deal quicker and extra correct with textual content outputs (e.g. Dictation on a cell device and so on.) then a person will be while typing on a keyboard. If such generation is applied, agencies can streamline management approaches and mitigate the weight of typing and different comparable tasks at the same time as allowing personnel to be conscious of greater complex components of the task.

VI. EXPERIMENTAL ANALYSIS

JavaScript: Unlike many different languages, where you want to move to download the language in your gadget with the intention to get admission to all of its features and create a development environment, with JavaScript everyone with a web browser all at once has an improvement surrounding proper in front of them. JavaScript is supported by every browser.

Speech Recognition using Web Speech API: Web Speech API is a JavaScript function that allows you to add speech recognition to any html web page. This API works in Chrome version 25 and later.[4]. The Web Speech API aims to permit net developers to provide, in an internet browser, speech-input and from text-to-speech output capabilities that are normally not to be had when the usage of famous speech-recognition or display screen-reader software program application. The API itself is cynical of the underlying speech reputation and synthesis execution and may guide each server-primarily based and purchaser-based totally definitely/embedded recognition and synthesis. The API is formed to allow each short (one-shot) speech to enter and continuous speech enter. Speech popularity outcomes are furnished to the internet web page as a listing of hypotheses, collectively with other relevant information for every hypothesis. Web API reports are physically produced and refreshed, in contrast to the documentation for neighborhood APIs which are ordinarily auto-produced[14].
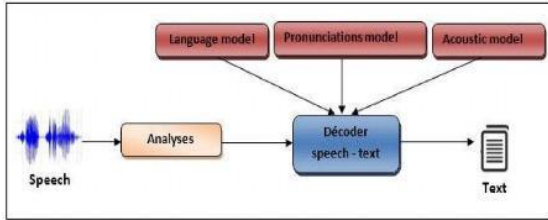
Fig 3: Figure 1. General Architecture of Automatic Speech Recognition Systems[2]

Usage of Local Storage: The Web Storage is modern and quick of components that constrain programs to shop key-cost matches. It is made to be a truckload more natural than the use of treats. The key-cost matches establish capacity devices, that are very much like articles other than they stay in one piece throughout page hundreds, and are generally strings. You can get to those values like an item or the use of the getItem() strategy (more prominent on that later).

In Google Chrome, web carport realities are saved in a SQLite archive in a subfolder inside the client's profile. The subfolder is situated at AppDataLocalGoogleChromeUser DataDefaultLocal Storage on Windows machines and ~/Library/ Application Support/Google/Chrome/ Default/Local Storage on macOS Firefox saves capacity objects in a SQLite record called webappsstore. Sqlite, which is in like manner put in the individual's profile organizer.

VII. CONCLUSION

Advanced voice generation will quickly be ubiquitous and we are able to be able to do herbal conversation with our clever cell phone. In the very near future, we can have natural spoken conversations with our clever Smartphone. Siri will rise from being an assistant to being a friend, truth seeker and guide. Now make this to a connected world with thousands and thousands of gadgets around us.

The API is integrated and is able to produce text from the speech using the microphone of the user.

Speech Recognition is the main interface that does all the work of ,it recognizes the audio input ,then a grammar interface Speech Grammar checks the grammar and asynchronous speech synthesis is done.
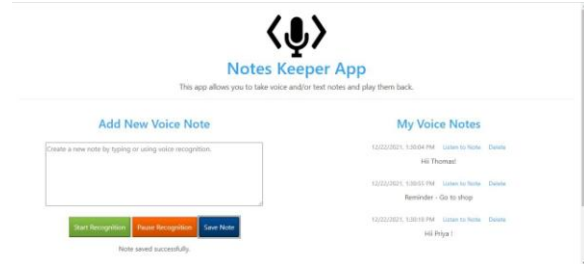
Fig 4 : Application Demonstration

Figure 4 demonstrate the working of the Javascript web speech API integrated in the application

Soon we are able to be speaking to our appliances – telling the dishwasher to show on, the fridge to re-order milk, the espresso device to brew stronger espresso. Voice era could make powerful contributions to society particularly wherein arms loose generation is wanted in hospitals and care centers. If a voice activated tool can name 911, it could store a lifestyle! The real pleasure in voice popularity is while it's blended with system mastering. Your smart cell phone will pay attention and study and harness that statistics to "examine your mind". Soon the linked household and workplace will start to expect your movements and put together itself in advance.

VIII. ACKNOWLEDGMENT

REFERENCE

[1] André Nitze, Andreas Schmietendorf "Modularity of JavaScript Libraries and Frameworks in Modern Web Applications" September 2014 Conference: User Conference

for Software Quality, Test and InnovationAt: Klagenfurt, ATVolume: 12

[2] Saliha Benkerzaz, Youssef Elmir, Abdeslam Dennai "A Study on Automatic Speech Recognition" Journal of Information Technology Review Volume 10 Number 3 August 2019

[3] Wu-Yuin Hwang,Yung-Hui Li ,RustamShadiev "Exploring effects of discussion on visual attention, learning performance, and perceptions of students learning with STR-support" ELSEVIER JOURNAL Computers & Education Volume 116, January 2018

[4] Paul Daniels "Using Web Speech technology with language learning applications" JALT CALL Journal, v11 n2 p177-187 2015

[5] Simon Holm Jensen , Magnus Madsen , Anders Møller "Modeling the HTML DOM and browser API in static analysis of JavaScript web applications" Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering September 2011

[6] Erik Wittern, Annie T.T. Ying, Yunhui Zheng, Julian Dolby, Jim A.Laredo "Statically Checking Web API Requests in JavaScript" 2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)

[7] M. Halle, K. Stevens "Speech Recognition: A Model and a Program for Research" IEEE JOURNAL -IRE Transactions on Information Theory (Volume: 8, Issue: 2, February 1962)

[8] Erik Wittern, Annie T.T. Ying, Yunhui Zheng, Julian Dolby, Jim A.Laredo, Christopher C.Young, Aleksander A. Slominski "Opportunities in Software Engineering Research for Web API Consumption" 2017 IEEE/ACM 1st International Workshop on API Usage and Evolution (WAPI) Conference: 23-23 May 2017

[9] Andy Neumann, Nuno Laranjeiro;,Jorge BernardinoAn "Analysis of Public REST Web Service APIs" IEEE JOURNAL, IEEE Transactions on Services Computing (Volume: 14, Issue: 4, July-Aug. 1 2021)

[10] P. Suter and E. Wittern, "Inferring Web API Descriptions From Usage Data," in Proceedings of the 3rd IEEE Workshop on Hot Topics in Web Systems and Technologies, 2015.

[11] C. S. Jensen, A. Møller, and Z. Su, "Server Interface Descriptions for Automated Testing of JavaScript Web Applications," in Proceedings of the 9th Joint Meeting on Foundations of Software Engineering, 2013, pp. 510–520

[12] S. Bae, H. Cho, I. Lim, and S. Ryu, "SAFEWAPI: Web API Misuse Detector for Web Applications," in Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering, 2014, pp. 507–517.

[13] David Bermbach, Erik Wittern "Benchmarking Web API Quality" Conference: International Conference on Web Engineering , 2016

[14] S.M. Sohan, Craig Anslow, Frank Maurer "A Case Study of Web API Evolution" Conference: 2015 2nd ACM International Conference on Mobile Software Engineering and Systems (MOBILESoft)

[15] Hari Krishna S M , Rinki Sharma "Survey on Application Programming Interfaces in Software Defined Networks and Network Function Virtualization" Global Transitions Proceedings Volume 2, Issue 2, November 2021, Pages 199-204