**SMART AI ASSISTANT**

Name: - Shivesh Kumar , Nikhil Kumar

Reg. no: -11802697,11802792

Advisor: - Sagar Pandey Sir

Department of Computer Science and Engineering, Lovely Professional University, India

**INTRODUCTION**

An **Smart AI Assistant** is a [software agent](https://en.wikipedia.org/wiki/Software_agent" \o "Software agent) that can perform tasks or services for an individual based on commands or questions. Sometimes the term "[chatbot](https://en.wikipedia.org/wiki/Chatbot" \o "Chatbot)" is used to refer to virtual assistants generally or specifically accessed by [online chat](https://en.wikipedia.org/wiki/Online_chat" \o "Online chat). In some cases, online chat programs are exclusively for entertainment purposes. Some virtual assistants are able to interpret human speech and respond via synthesized voices. Users can ask their assistants questions, control home automation devices and media playback via voice, and manage other basic tasks such as email, to-do lists, and calendars with verbal commands. A similar concept, however with differences, lays under the [dialogue systems](https://en.wikipedia.org/wiki/Dialogue_system" \o "Dialogue system)

**Acknowledgement**

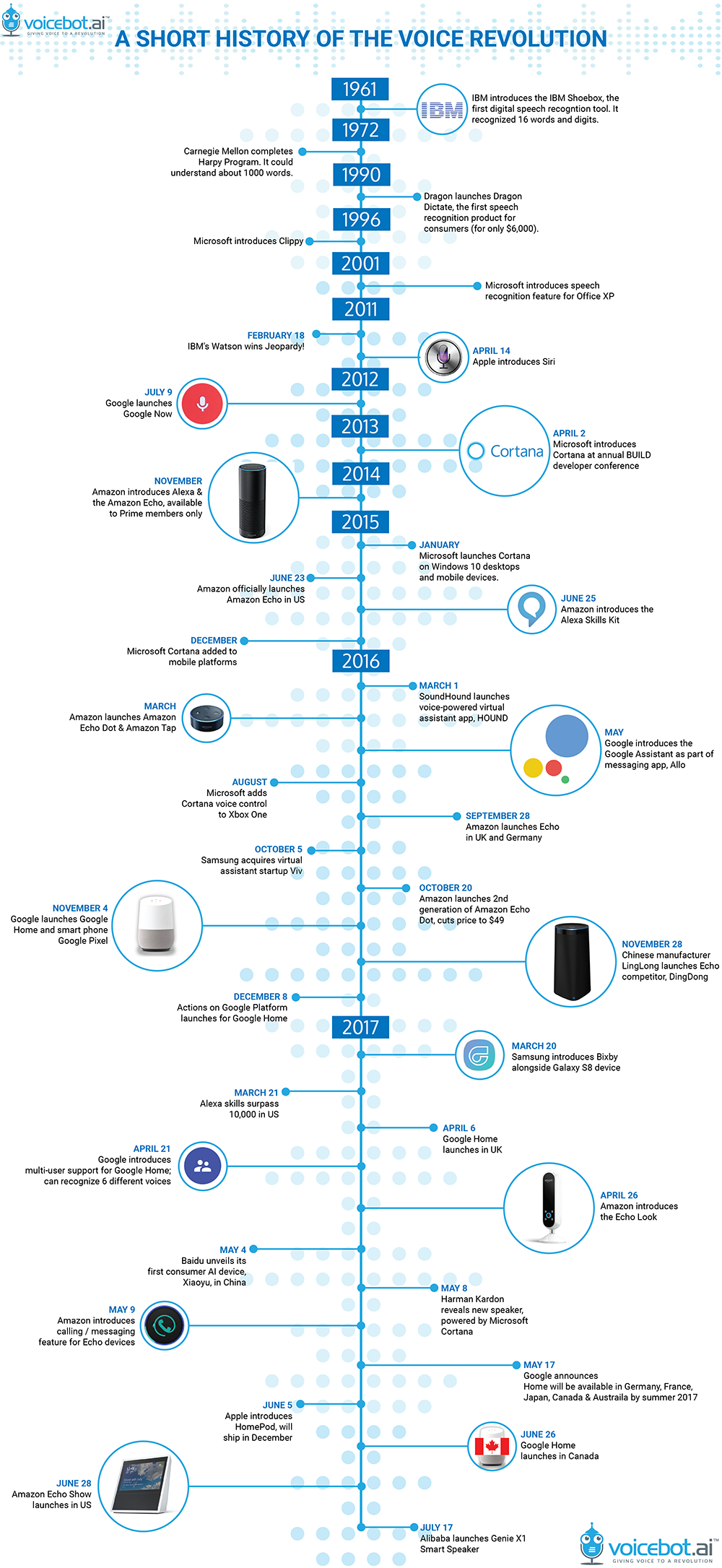
We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. We would like to extend my sincere thanks to all of them.

We are highly indebted to Sagar Pandey Sir for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project.

We would like to express our gratitude towards our parents & member of Lovely Professional University for their kind co-operation and encouragement which help us in completion of this project.

**HISTORY**

Another early tool which was enabled to perform digital speech recognition was the [IBM Shoebox](https://en.wikipedia.org/wiki/IBM_Shoebox" \o "IBM Shoebox), presented to the general public during the [1962 Seattle World's Fair](https://en.wikipedia.org/wiki/1962_Seattle_World%27s_Fair" \o "1962 Seattle World's Fair) after its initial market launch in 1961. This early computer, developed almost 20 years before the introduction of the first [IBM Personal Computer](https://en.wikipedia.org/wiki/IBM_Personal_Computer" \o "IBM Personal Computer) in 1981, was able to recognize 16 spoken words and the digits 0 to 9. The next milestone in the development of voice recognition technology was achieved in the 1970s at the [Carnegie Mellon University](https://en.wikipedia.org/wiki/Carnegie_Mellon_University" \o "Carnegie Mellon University) in [Pittsburgh](https://en.wikipedia.org/wiki/Pittsburgh" \o "Pittsburgh), [Pennsylvania](https://en.wikipedia.org/wiki/Pennsylvania" \o "Pennsylvania) with substantial support of the [United States Department of Defense](https://en.wikipedia.org/wiki/United_States_Department_of_Defense" \o "United States Department of Defense) and its [DARPA](https://en.wikipedia.org/wiki/DARPA" \o "DARPA) agency. Their tool "Harpy" mastered about 1000 words, the vocabulary of a three-year-old. About ten years later the same group of scientists developed a system that could analyze not only individual words but entire word sequences enabled by a [Hidden Markov Model](https://en.wikipedia.org/wiki/Hidden_Markov_Model" \o "Hidden Markov Model).[]](https://en.wikipedia.org/wiki/Virtual_assistant" \l "cite_note-spracherkennung1-5) Thus, the earliest virtual assistants, which applied speech recognition software were [automated attendant](https://en.wikipedia.org/wiki/Automated_attendant" \o "Automated attendant) and medical [digital dictation](https://en.wikipedia.org/wiki/Digital_dictation" \o "Digital dictation) software.[[6]](https://en.wikipedia.org/wiki/Virtual_assistant" \l "cite_note-6) In the 1990s digital speech recognition technology became a feature of the [personal computer](https://en.wikipedia.org/wiki/Personal_computer" \o "Personal computer) with [Microsoft](https://en.wikipedia.org/wiki/Microsoft" \o "Microsoft), [IBM](https://en.wikipedia.org/wiki/IBM" \o "IBM), [Philips](https://en.wikipedia.org/wiki/Philips" \o "Philips) and [Lernout & Hauspie](https://en.wikipedia.org/wiki/Lernout_%26_Hauspie" \o "Lernout & Hauspie) fighting for customers. Much later the market launch of the first [smartphone](https://en.wikipedia.org/wiki/Smartphone" \o "Smartphone) [IBM Simon](https://en.wikipedia.org/wiki/IBM_Simon" \o "IBM Simon) in 1994 laid the foundation for smart virtual assistants as we know them today. The first modern digital virtual assistant installed on a smart phone was Siri, which was introduced as a feature of the [iPhone 4S](https://en.wikipedia.org/wiki/IPhone_4S" \o "IPhone 4S) on October 4, 2011.[[Apple Inc.](https://en.wikipedia.org/wiki/Apple_Inc." \o "Apple Inc.) developed Siri following the 2010 acquisition of [Siri Inc.](https://en.wikipedia.org/wiki/Siri_Inc." \o "Siri Inc.), a [spin-off](https://en.wikipedia.org/wiki/Corporate_spin-off" \o "Corporate spin-off) of [SRI International](https://en.wikipedia.org/wiki/SRI_International" \o "SRI International), which is a research institute financed by [DARPA](https://en.wikipedia.org/wiki/DARPA" \o "DARPA) and the [United States Department of Defense](https://en.wikipedia.org/wiki/United_States_Department_of_Defense" \o "United States Department of Defense).



**ABSTRACT**

Voice recognition technology is one from the fast-growing engineering technologies. It has several applications in different areas and provides potential benefits. Nearly 20% people of the world are suffering from various disabilities; many of them are blind or unable to use their hands effectively. The Voice recognition systems in those particular cases provide a significant help to them, so that they can share information with people by operating computer through voice input. This project is designed and developed keeping that factor into mind, and a little effort is made to achieve this aim. Our project is capable to recognize the speech and convert the input audio into text; it also enables a user to perform operations such as “save, open, exit” a file by providing voice input. It also helps the user to open different system software such as opening Ms-paint, notepad and calculator. At the initial level effort is made to provide help for basic operations as discussed above, but the software can further be updated and enhanced in order to cover more operation

**HOW IT WORKS?**

A virtual assistant is a ***[technology based on artificial intelligence](https://chatbotsmagazine.com/intelligent-assistants-i-a-85c21f9d3b8e" \t "_blank)***. The software uses a device’s microphone to receive voice requests while the voice output takes place at the speaker. But the most exciting thing happens between these two actions.

It is a combination of several different technologies: voice recognition, voice analysis and language processing.

When a user asks a personal assistant to perform a task, the natural language audio signal is converted into digital data that can be analyzed by the software. Then this data is compared with a database of the software using an innovative algorithm to find a suitable answer. This database is located on distributed servers in cloud networks. For this reason, most personal assistants cannot work without a reliable Internet connection.

With the increasing number of queries the software’s database gets expanded and optimized, which improves voice recognition and increases the response time of the system.

 The popularity of smartphones opened the opportunity to add technology into consumer pockets, while smart home devices, like Google Home and Amazon Echo, brought voice recognition technology into living rooms and kitchens.

**Various Libraries used**

* **Speech Recognition**: - Speech recognition is the process of converting spoken words to text. Python supports many speech recognition engines and APIs, including Google Speech Engine, **Google Cloud Speech API**, Microsoft Bing Voice Recognition and **IBM Speech to Text**.
* **pyttsx3**: - It is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline, and is compatible with both Python 2 and 3.
* **datetime**: - The datetime module supplies classes for manipulating dates and times. While date and time arithmetic is supported, the focus of the implementation is on efficient attribute extraction for output formatting and manipulation.
* **Wikipedia**: - It is a Python library that makes it easy to access and parse data from Wikipedia. Search Wikipedia, get article summaries, get data like links and images from a page, and more. Wikipedia wraps the MediaWiki API so you can focus on using Wikipedia data, not getting it.
* **web browser**: - The web browser module provides a high-level interface to allow displaying Web-based documents to users. Under most circumstances, simply calling the open () function from this module will do the right thing. Under Unix, graphical browsers are preferred under X11, but text-mode browsers will be used if graphical browsers are not available or an X11 display isn’t available. If text-mode browsers are used, the calling process will block until the user exits the browser. If the environment variable BROWSER exists, it is interpreted to override the platform default list of browsers, as an os. pathsep-separated list of browsers to try in order. When the value of a list part contains the string %s, then it is interpreted as a literal browser command line to be used with the argument URL substituted for %s; if the part does not contain %s, it is simply interpreted as the name of the browser to launch.
* **os**: - This module provides a portable way of using operating system dependent functionality. If you just want to read or write a file see open(), if you want to manipulate paths, see the os.path module, and if you want to read all the lines in all the files on the command line see the file input module. For creating temporary files and directories see the tempfile module, and for high-level file and directory handling see the shutil module.

**Components of Speech recognition System**

***Voice Input*** With the help of microphone audio is input to the system, the pc sound card produces the equivalent digital representation of received audio.

***Digitization*** The process of converting the analog signal into a digital form is known as digitization, it involves the both sampling and quantization processes. Sampling is converting a continuous signal into discrete signal, while the process of approximating a continuous range of values is known as quantization.

***Acoustic Model*** An acoustic model is created by taking audio recordings of speech, and their text transcriptions, and using software to create statistical representations of the sounds that make up each word. It is used by a speech recognition engine to recognize speech. The software acoustic model breaks the words into the phonemes.

***Language Model*** Language modelling is used in many natural language processing applications such as speech recognition tries to capture the properties of a language and to predict the next word in the speech sequence. The software language model compares the phonemes to words in its built-in dictionary.

***Speech engine*** the job of speech recognition engine is to convert the input audio into text; to accomplish this it uses all sorts of data, software algorithms and statistics. Its first operation is digitization as discussed earlier, that is to convert it into a suitable format for further processing. Once audio signal is in proper format it then searches the best match for it. It does this by considering the words it knows, once the signal is recognized it returns its corresponding text string.

**APPLICATIONS FOR VOICE RECOGNITION:**

***From medical perspective*** People with disabilities can benefit from speech recognition programs. Speech recognition is especially useful for people who have difficulty using their hands, in such cases speech recognition programs are much beneficial and they can use for operating computers. Speech recognition is used in deaf telephony, such as voicemail to text.

***From military perspective*** Speech recognition programs are important from military perspective; in Air Force speech recognition has definite potential for reducing pilot workload. Beside the Air force such Programs can also be trained to be used in helicopters, battle management and other applications.

***From educational perspective*** Individuals with learning disabilities who have problems with thought-to-paper communication (essentially, they think of an idea, but it is processed incorrectly causing it to end up differently on paper) can benefit from the software. Some other application areas of speech recognition technology are described as under:

***Command and Control*** ASR systems that are designed to perform functions and actions on the system are defined as Command and Control systems. Utterances like "Open Netscape" and "Start a new browser" will do just that.

***Telephony*** Some Voice Mail systems allow callers to speak commands instead of pressing buttons to send specific tones.

***Medical/Disabilities*** Many people have difficulty typing due to physical limitations such as repetitive strain injuries (RSI), muscular dystrophy, and many others. For example, people with difficulty hearing could use a system connected to their telephone to convert the caller's speech to text.

**CONCLUSION: -**

This project work of Voice Recognition started with a brief introduction of the technology and its applications in different sectors. The project part of the Report was based on software development for speech recognition. At the later stage we discussed different tools for bringing that idea into practical work. After the development of the software finally it was tested and results were discussed, few deficiencies factors were brought in front. After the testing work, advantages of the software were described and suggestions for further enhancement and improvement were discussed.

**REFERENCES: -**

1. <https://pypi.org/project/pyttsx3/>
2. [https://www.w3school.com/](https://w3school.com/)
3. [https://pypi.org/project/SpeechRecognition/](https://pypi.org/project/pyttsx3/)
4. <https://pypi.org/project/DateTime/>
5. <https://www.wikipedia.org>
6. <https://github.com/realpython/>