Voicella : System's Voice Assistant Using Python

K. Manikanta,M.Afreed,N.Gopal Krishna

Dept.of Information Technology, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, India

ABSTRACT: Voice assistants are programs on digital devices that listen and respond to verbal commands. A user can say, "What's the weather?" and the voice assistant will answer with the weather report for that day and location. They could say, "Tell me a story," and the assistant will jump into a tale. The user could even say, "Order my favourite pizza," and dinner will be on its way! Voice assistants are so easy to use that many people forget to stop and WONDER. How they work. How do voice assistants understand us? Is it magic? A complex system of codes? An actual person listening on the other end? The answer is less complicated than you might think. The application works like Siri, Google Assistant etc. The U.I of the application is self-explainable and very minimum. It takes voice as input. The system is being designed in such a way that all the services provided by the mobile devices are accessible by the end user on the user's voice commands.

Indexed Terms– voice to speech.

I. INTRODUCTION

The similar thought before making "Digital" Personal Assistant. Though it is not as capable and high as like Amazon's Alexa or Google Assistant, Home or Apple's Siri or JARVIS from Iron Man. Nowadays, People are troubled by typing commands into the computer. Be it procrastination or a busy schedule. Typing is a big obsolete process. The solution to this is that we switch over to an assistant which understands us and do the initial work for us. An assistant is the best replacement for typing commands.

It's named as Desktop Voice Assistant VOICELLA with Voice Recognition Intelligence, which takes the user input in form of user's voice and processes it and return the output in various ways like an action to be performed or the search result is spoken out to the end user.

In today's era almost all tasks are digitalized. We have Smartphone in hands and it is nothing less than having world at your fingertips. These days we aren't even using fingers. We just speak of the task and it is done. There exist systems where we can say Text Dad, "I'll be late today." And the text is sent. That is the task of a Virtual Assistant.

Virtual Assistants are software programs that help you ease your day today tasks, such as showing weather report, creating reminders, making shopping lists etc. This system is designed to be used efficiently on desktops. Personal assistant software improves user productivity by managing routine tasks of the user and by providing information from online sources to the user. VOICELLA is effortless to use. Call the wake word 'VOICELLA' followed by the command. And within seconds, it gets executed. Voice searches have dominated over text search. Web searches conducted via mobile devices have only just overtaken those carried out using a computer and the analysts a real ready predicting that 50% of searches will be via voice by 2020.Virtual assistants are turning out to be smarter than ever. Allow your intelligent assistant to make email work for you. Detect intent, pick out important information, automate processes, and deliver personalized responses. This project was started on the premise that there is sufficient amount of openly available data and information on the web that can be utilized to build a virtual assistant that has access to making intelligent.

II. MODULES

a) There are two modules in the phishing URL detection application. They are:
b) •   User Interface
c) •   System interface
d) User gives voice commands that a system need to perform and System greets the user and waits for the command.
e) System Interface has 4 steps in it.

f) Speech recognition software works by breaking down the audio of a speech recording into individual sounds, analyzing each sound, using algorithms to find the most probable word fit in that language, and transcribing those sounds into text.

g) Speech Recognition is an important feature in several applications used such as home automation, artificial intelligence, etc. This article aims to provide an introduction on how to make use of the SpeechRecognition library of Python. This is useful as it can be used on microcontrollers such as Raspberri Pis with the help of an external microphone.

h) PSutil (process and system utilities) is a cross-platform library for retrieving information on running processes and system utilization (CPU, memory, disks, network, sensors) in Python. It is useful mainly for system monitoring, profiling and limiting process resources and management of running processes. It implements many functionalities offered by classic UNIX command line tools such as PS, top, io top, lsof, netstat, ifconfig, free and others. PSutil currently supports the following platforms like

i) ☐ Windows

j) ☐ macOS FreeBSD

k) ☐ OpenBSD

l) ☐ NetBSD

m) ☐ Sun Solaris

n) ☐ AIX

o) ☐ Linux

p) The SMTP model is of two types:

q) End-to- End Method

r) The end-to-end model is used to communicate between different organizations whereas the store and forward method are used within an organization. A SMTP client who wants to send the mail will contact the destination's host SMTP directly in order to send the mail to the destination.

s) Store-and- Forward Method

t) The SMTP server will keep the mail to itself until it is successfully copied to the receiver's SMTP. The client SMTP is the one which initiates the session let us call it as the client- SMTP and the server SMTP is the one which responds to the session request and let us call it as receiver-SMTP. The client- SMTP will start the session and the receiver-SMTP will respond to the request.

u) Google Text-to-Speech Module

v) The text-to-speech (TTS) is the process of converting words into a vocal audio form. The program, tool, or software takes an input text from the user, and using methods of natural language processing understands the linguistics of the language being used, and performs logical inference on the text. This processed text is passed into the next block where digital signal processing is performed on the processed text. Using many algorithms and transformations this processed text is finally converted into a speech format. This entire process involves the synthesizing of speech. Below is a simple block diagram to understand the same.

w) The Text-to-Speech API enables developers to generate human-like speech. The API converts text into audio formats such as WAV, MP3, or Ogg Opus. It also supports Speech Synthesis Markup Language (SSML) inputs to specify pauses, numbers, date and time formatting, and other pronunciation instructions.

III. CONCLUSION

The Student portal Authentication system with threelevelpasswordsisdevelopedtofacilitate,student,administratoreasyprocessingforstudentsineducationalinstitutions.Manually,thisconsumesalotoftime,effortandpaperwork.Andalsoiftheconcernedauthorityisnotavailable,thetaskoflogginginbecomescomplicated.So,thiswebapplication overcomes all these limitations and offersa great deal of help at each and every stage in thewholeprocessofcommunicatingto students.

IV. FUTURESCOPE

This project Student portal Authentication system withthree level passwords has been developed in such amanner, that the future requirements of the user aremet.Theprojectisflexibletoadaptthechangesefficiently without affecting the present system. Infuture, there can be a provision to update attendance,results,andnotificationsthroughthewebapplication.

Wearealsoplanningtoimplementthewebapplication on various other platforms like Windowsand

ios.Thisisthefuturescopeofour project.

REFERENCES

1. Hirschberg, Julia, and Christopher D. Manning. "Advances in Natural Language Processing." Science 349 (6245): 261–266. doi:10.1126/science.aaa8685.
2. Moore, Clayton. "The Most Useful Skills for Google Home." Digital Trends (May 3, 2017). https://www.digitaltrends.com/home/google-home-most-useful-skills/.
3. Lee, Nicole. "Google Assistant on the iPhone Is Better than Siri, but Not Much." Engadget (May 17, 2017). https://www.engadget.com/2017/05/17/google-assistant-iphonehands-on/.
4. Hachman, Mark. "The Microsoft-Amazon Deal Leaves Cortana Speakers with One Advantage: Skype." PCWorld (August 30, 2017). https://www.pcworld.com/article/ 3221284/windows/the-microsoft-amazon-deal-leaves-cortana-speakers-with-one-advantageskype.html.
5. Spence, Ewan. "Windows Phone Is Dead, Long Live Microsoft's Smartphone Dream." Forbes (July 12, 2017). https://www.forbes.com/sites/ewanspence/2017/07/12/microsoftwindows-phone-windows10-mobile-strategy/.
6. Tilley, Aaron. "How A Few Words to Apple's Siri Unlocked a Man's Front Door." Forbes (September 21, 2016). https://www.forbes.com/sites/aarontilley/2016/09/21/applehomekit-siri-security/.
7. Liptak, Andrew. "Amazon's Alexa Started Ordering People Dollhouses after Hearing Its Name on TV." The Verge (January 7, 2017). https://www.theverge.com/2017/1/7/ 14200210/amazon-alexa-tech-news-anchor-order-dollhouse.
8. Zhang, Guoming, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. "DolphinAttack: Inaudible Voice Commands." In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, New York, NY, 103–117. doi:10.1145/3133956.3134052.
9. Tung, Liam. "Google Home Mini Flaw Left Smart Speaker Recording Everything." ZDNet (October 11, 2017). http://www.zdnet.com/article/google-home-mini-flaw-left-smartspeaker-recording-everything/.
10. Buhr, Sarah. "An Amazon Echo May Be the Key to Solving a Murder Case." TechCrunch (December 27, 2016). http://social.techcrunch.com/2016/12/27/an-amazon-echo-may-bethe-key-to-solving-a-murder-case/. MEDICAL REFERENCE SERVICES QUARTERLY 87
11. Vincent, James. "Mattel Cancels AI Babysitter after Privacy Complaints." The Verge (October 5, 2017). https://www.theverge.com/2017/10/5/16430822/mattel-aristotle-aichild-monitor-canceled.
12. Wolters, Maria Klara, Fiona Kelly, and Jonathan Kilgour. "Designing a Spoken Dialogue Interface to an Intelligent Cognitive Assistant for People with Dementia." Health Informatics Journal 22 no. 4 (December 1, 2016): 854–866. doi:10.1177/1460458215593329.
13. Palladino, Valentina. "Google Pixel Buds Are Wireless Earbuds That Translate Conversations in Real Time." Ars Technica (October 4, 2017). https://arstechnica.com/ gadgets/2017/10/google-pixel-buds-are-wireless-earbuds-that-translate-conversations-inreal-time/.