

**REPUBLIC OF TURKEY**

**ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY**

**FACULTY OF ENGINEERING**

**DEPARTMENT OF COMPUTER ENGINEERING**

**VOICE ASISTANT**

**MEHMET METO**

**BESTAMİ GÜNAY**

**ADANA 2022**

# Abstract

In this project, what is the voice assistant, its history, leading companies, timeline and application usage are mentioned. Python was used as the software language. Visual Studio Code was used as the compiler.

# What is a voice assistant?

A voice assistant is a digital assistant that uses voice recognition, language processing algorithms, and voice synthesis to listen to specific voice commands and return relevant information or perform specific functions as requested by the user. Based on specific commands, sometimes called intents, spoken by the user, voice assistants can return relevant information by listening for specific keywords and filtering out the ambient noise. While voice assistants can be completely software based and able to integrate into most devices, some assistants are designed specifically for single device applications, such as the Amazon Alexa Wall Clock. Today, voice assistants are integrated into many of the devices we use on a daily basis, such as cell phones, computers, and smart speakers. Because of their wide array of integrations, there are several voice assistants who offer a very specific feature set, while some choose to be open ended to help with almost any situation at hand.

# History of voice assistant

Voice assistants have a very long history that actually goes back over 100 years, which might seem surprising as apps such as Siri have only been released within the past ten years.

The very first voice activated product was released in 1922 as Radio Rex. This toy was very simple, wherein a toy dog would stay inside a dog house until the user exclaimed its name, “Rex” at which point it would jump out of the house. This was all done by an electromagnet tuned to the frequency similar to the vowel found in the word Rex, and predated modern computers by over 20 years.

At the 1952 World’s fair, Audrey was announced by Bell Labs. The Automatic Digit Recognizer was not a small simple device however, its casing stood six feet tall just to house all the materials required to recognize ten numbers!

IBM began their long history of voice assistants in 1962 at the World’s Fair in Seattle when IBM Shoebox was announced. This device was able to recognize digits 0-9 and six simple commands such as, “plus, minus” so the device could be used as a simple calculator. Its name referred to its size, similar to the average shoebox, and contained a microphone connected to three audio filters to match the electric frequencies of what was being said and matched it with already assigned values for each digit.

Darpa then funded five years of speech recognition R&D in 1971, known as the Speech Understanding Research (SUR) Program. One of the biggest innovations to come out if this was Carnegie Mellon’s Harpy, which was capable of understanding over 1,000 words.

The next decade led to amazing progress and research in the speech recognition field, leading most voice recognition devices from understanding a few hundred words to understanding thousands, and slowly making their way into consumers homes.

Then, in 1990, Dragon Dictate was introduced to consumers homes for the shocking price of $9,000! This was the first consumer oriented speech recognition program designed for home PC’s. The user could dictate to the computer one word at a time, pausing in between each word waiting for the computer to process before they could move on. Seven years later, Dragon NaturallySpeaking was released and it brought more natural conversation, able to understand continuous speech at a maximum of 100 words per minute and a much lower price tag of $695.

In 1994, Simon by IBM was the first smart voice assistant. Simon was a PDA, and really, the first smartphone in history, considering it predates HTC’s Droid by practically 25 years!

In 2008, when Android was first released, Google had slowly started rolling out voice search for its Google mobile apps on various platforms, with a dedicated Google Voice Search Application being released in 2011. This led to more and more advanced features, eventually leading to Google now and Google Voice Assistant.

Then, this was followed by Siri in 2010. Developed by SRI International with speech recognition provided by Nuance Communications, the original app was released in 2010 on the iOS App Store and was acquired two months later by Apple. Then, with the release of the iPhone 4s, Siri was officially released as an integrated voice assistant within iOS. Since then, Siri has made its way to every Apple device available and has linked all the devices together in a single ecosystem.

Shortly after Siri was first developed, IBM Watson is announced publicly in 2011. Watson was named after the founder of IBM, and was originally conceived in 2006 to beat humans at a game of Jeopardy. Now, Watson is one of the most intelligent, naturally speaking computer systems available.

Amazon Alexa is then announced in 2015. Its name being inspired by the Library of Alexandria and also the hard consonant “X” in the name, helping with more accurate voice recognition. With Alexa, the Echo line of smart devices are announced to bring smart integration to consumers homes for an inexpensive route.

Alan is finally publicly announced in 2017 to take the Enterprise Application world by storm. Being first born as “Synqq”, Alan is created by the minds behind “Qik”, the very first video messaging and conferencing mobile app. Alan is the first voice AI platform aimed at enterprise applications, so while it can be found in many consumer applications, it is designed for enterprises to be able to develop and integrate quickly and efficiently!

# Technology behind voice assistants

Voice assistants use Artificial Intelligence and Voice recognition to accurately and efficiently deliver the result that the user is looking for.

## Voice Recognition

**Voice recognition** works by taking an analog signal from a users voice and turning it into a digital signal. After doing this, the computer takes the digital signal and attempts to match it up to words and phrases to recognize the users intent. To do this, the computer requires a database of pre-existing words and syllables in a given language to be able to closely match the digital signal with. Checking the input signal with this database is known as **pattern recognition,** and is the primary force behind voice recognition.

## Artificial Intelligence

**Artificial intelligence** is using machines to simulate and replicate human intelligence.

In 1950, Alan Turing (The namesake of our company) published his paper “Computing Machinery and Intelligence” that first asked the question, can machines think? Alan Turing then went on to develop the Turing Test, a method of evaluating a computer to test its capability of thinking like a human. There were four approaches later developed that defined AI, Thinking humanly/rationally, and acting humanly/rationally. While the first two deal with reasoning, the second two deal with actual behavior. Modern AI is typically seen as a computer system designed to accomplish tasks that typically require human interaction. These systems can improve upon themselves using a process known as machine learning.

## Machine Learning

Machine learning refers to the subset of Artificial Intelligence where programs are created without the use of human coders manually creating the program. Instead of writing out the complete program on their own, programmers gives the AI “patterns” to recognize and learn from and then gives the AI large amounts of data to sift through and study. So instead of having specific rules to abide by, the AI searches for patterns within this data and uses it to improve its already existing functions. One way machine learning can be helpful for Voice AI, is by feeding the algorithm hours of speech from various accents and dialects.

While traditional programs requires an input and rules to develop an output, machine learning tools are given an input and an output and use that to create the program itself. There are two approaches to machine learning, supervised learning and unsupervised learning. In supervised learning, the model is given data that is already partly labeled, this means some of the data given will be already tagged with the correct answer. This helps guide the model into categorizing the rest of the data and developing a correct algorithm. In unsupervised learning, none of the data is labeled, so it is up to the model to find the pattern correctly. One of the reasons this is very useful is because it allows the model to find patterns that the creators might have never found on their own, but the data is much more unpredictable.

# Benefits of voice assistants

Some examples of what a Voice Assistant can do include:

* Check the weather
* Turn on/off connected smart devices
* Search databases

One of the main reasons of the growing popularity of Voice User Interfaces (VUI) is due to the growing complexity within mobile software without an increase in screen size, leading to a huge disadvantage by using a GUI (Graphical User Interface). As more iterations of phones come out, the screen sizes stay relatively the same, leading for very cramped interfaces and creating frustrating user experiences, which is why more and more developers are switching to Voice User Interfaces.

## Efficiency and Safety

While typing has become much faster as people have gotten used to using standard keyboards, using your voice will always be quicker, much more natural, and lead to less spelling errors. This leads to a much more efficient and natural intelligent workflow.

## Quick Learning Curve

One of the greatest benefits of voice assistants is a quick learning curve. Instead of having to learn how to use devices like mice and touch screens and get used to using specific physical devices, you can just use your natural conversation tendencies and use your voice.

## Wider Device Integration

Since a screen or keyboard isn’t necessary, it’s easy to place voice integration into a much wider array of devices. In the future, smart glasses, furniture, appliances, will all come with voice assistants already integrated into the device.

# Popular voice assistants

## Siri

Siri is the most popular voice assistant today. Created in 2010 by SRI Inc, and purchased in 2011 by Apple, Siri has quickly become an integral part of the Apple ecosystem in bringing all the Apple devices and applications together to use in tandem with one another.

## Alexa

Created by Amazon in 2014, Alexa was named due to its similarity to the Library of Alexandria. Alexa was originally inspired by the conversational voice system found on board the U.S.S. Enterprise in Star Trek. Alexa was released alongside The Amazon Echo, a smart speaker intended for consumers to dive into the world of home automation, uses the Alexa platform to allow users to interact with the Amazon ecosystem and allow for a plethora of smart devices to be connected.

## Google Assistant

Originally unveiled in 2016, Google Assistant was the spiritual successor of Google Now, with the main improvement being the addition of two-way conversations. Where Google now would return answers in the form of a search results page on Google, Google Assistant gives answers in the form of natural sentences and returns recommendations in the form of Feature cards.

## Cortana

Beginning in 2009, Cortana by Microsoft has had one of the longest visions of giving people access to voice assistants in their daily lives. Microsoft began shipping Cortana with all Windows 10 and Xbox devices, leading to a huge increase in the amount of registered Cortana users. In 2018 it was reported that Cortana had over 800 Million users.

## Alan

In 2017 Alan set out to take voice assistants to the next level, by enabling voice AI for all applications. Using domain specific language models and contextual understanding, Alan is focused on creating a new generation of Enterprise Voice AI applications. By using the Alan Platform, developers are able to take control of voice, and create an effective workflow that best fits their users with the help of vocal commands.

# Voice assistant timeline

* 1922 – First Voice activated consumer product hits store shelves as “Radio Rex”
* 1952 – Audrey, or the Automatic Digit Recognition Machine, is announced
* 1962 – IBM Shoebox is shown for the first time at the State Fair
* 1971 – Darpa funds five years of speech recognition research and development
* 1976 – Harpy is shown at Carnegie Mellon
* 1984 – IBM releases “Tangora” the first voice activated typewriter
* 1990 – Dragon Dictate is released
* 1994 – Simon by IBM is the first modern voice assistant released
* 2010 – Siri is released as an app on the iOS app store
* 2011 – IBM Watson is released
* 2012 – Google Now is released
* 2014 – Amazon Alexa and Echo are released
* 2015 – Microsoft Cortana is released
* 2017 – Alan is developed and released with the Alan Platform

# App usage

When our voice assistant app hears one of the words "hey, or hello", the assistant responds with "hey how can I help you".

When you ask one of the sentences "what is your name", "what's your name", "tell me your name", it will say its own name. It then asks for the user's name.( "what's your name ?")

For example:

User: Hey…

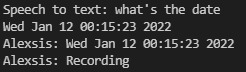
Asistant: Hey how can i help you ?

User: What’s your name ?

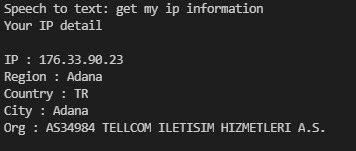
Asistant: My name is Alexis, what’s your name ?

The name change command says "your name should be ..... "

If you want to ask for the date, use the "whats the date" command



Just say "get my ip information" for your ip information.



You can use the expression "weather" for the weather forecast.

Say ("what is the " + search word) to search Wikipedia.

For simple mathematical operations you can use "plus, minus, divide, multiply".

For example:

User: Seven plus five

Asistant: twelve



You can use "capture, my screen or screenshot" for screenshot.

You can use the expressions "where am i or location" for your location information.

To search on google, you can use the phrases "search for + the word to dec searched for".

To search on youtube, you can use the phrases "youtube + the word to dec searched for".

For example:

User: youtube cats

exit commands are “exit”, “quit”, or “goodbye”

# References

<https://alan.app/blog/voiceassistant-2/#:~:text=A%20voice%20assistant%20is%20a,as%20requested%20by%20the%20user.>

<https://medium.com/yaz%C4%B1l%C4%B1m-ve-bili%C5%9Fim-kul%C3%BCb%C3%BC/python-ile-sesli-asistan-922b5e49b643>

<https://tr.linkedin.com/pulse/pythonda-sesli-asistan-olu%C5%9Fturmak-yunus-emre-g%C3%BCndo%C4%9Fmu%C5%9F>

<https://www.smartsheet.com/voice-assistants-artificial-intelligence>