EVALUATION OF INTERNSHIP REPORT

B.Tech: III Year

Department of Computer Science & Information Technology

Name of the Student	**:**	**Narayan Punase**
Branch & section	**:**	**CSIT-2**
Roll No	**:**	**0827CI201116**
Year	**:**	**2022-23**

Department of Computer Science & Information Technology
AITR, Indore,

ACROPOLIS INSTITUTE OF TECHNOLOGY & RESEARCH, INDORE

Department of Computer Science & Information Technology

Certificate

Certified that training work entitled "*ALEXA*" is a bonafied work carried out after sixth semester by "*Narayan Punase*" in partial fulfilment for the award of the degree of Bachelor of Technology in Computer Science and Information Technology from "*GARIMA KUMRAWAT (Assistant professor)*" Acropolis Institute of Technology and Research during the academic year 2022-23.

Name and Sign of Training Coordinator *Name & Sign of Internship Coordinator*

ACROPOLIS INSTITUTE OF TECHNOLOGY & RESEARCH, INDORE

Department of Computer Science & Information Technology

ACKNOWLEDGEMENT

I would like to acknowledge the contributions of the following people without whose help and guidance this report would not have been completed. I acknowledge the counsel and support of our training GARIMA KUMRAWAT (Assistant professor) , CSIT Department, with respect and gratitude, whose expertise, guidance, support, encouragement, and enthusiasm has made this report possible. Their feedback vastly improved the quality of this report and provided an enthralling experience. I am indeed proud and fortunate to be supported by him/her. I am also thankful to Dr. Shilpa Bhalerao, H.O.D of Computer Science Information Technology Department, for her constant encouragement, valuable suggestions and moral support and blessings. Although it is not possible to name individually, I shall ever remain indebted to the faculty members of CSIT Department, for their persistent support and cooperation extended during this work.

Student Name : Narayan Punase

Student Enrollment No. : 0827CI201116

ACROPOLIS INSTITUTE OF TECHNOLOGY & RESEARCH, INDORE

INDEX

Introduction to technology Undertaken (Alexa)

What is ALEXA?

We must first understand what the Alexa system is, and how it processes requests. For purposes of this discussion, the Alexa system is comprised of Echo devices listed above—the hardware and software that customers directly interact with and the cloud components—which have the majority of the “smarts”: Automatic Speech Recognition, Natural Language Understanding, and Response. Some responses are provided by third party services through “skills.” The 3rd-parties that write and publish those skills are responsible for their skill’s behavior. To start, we begin with a very simple tour through the entire system, demonstrated with the example request of “Alexa, what is the weather,” so we can see how the request is picked up by an Echo, sent through voice recognition, interpreted, acted upon, and then responded to

Echo Devices

Echo devices are the input and output devices for Alexa, and we designed and built them, from the beginning, with multiple layers of security and privacy protections and controls. Echo devices use on-device keyword spotting designed to detect when a customer says the wake word. This technology inspects acoustic patterns in the room to detect when the wake word has been spoken using a short, on-device buffer that is continuously overwritten. This on-device buffer exists in temporary memory (RAM); audio is not recorded to any on-device storage. The device does not stream audio to the cloud until the wake word is detected or the action button on the device is pressed. If it does not detect the wake word or if the action button is not pressed, no audio is sent to the cloud. The customer can turn the microphone off at any time by pressing the microphone button on the device. When the button is pressed to turn the microphones off, the microphones are electrically disconnected and a dedicated red LED on the microphone button is illuminated to indicate the microphones are off. As a safeguard, we designed the circuitry of Echo devices so that power can only be provided either to this dedicated red LED or to the device microphones, not to both at the same time. As a result, if the dedicated red LED is illuminated, the microphones are off and cannot stream audio to the cloud. In the usual case, with the microphone

on, when the user says, “Alexa, what is the weather in Menlo Park,” the Echo’s light ring turns blue to indicate that the device detected the wake word and is streaming audio to the cloud. The system is designed so that communication between the Echo device and the Alexa cloud is protected using TLS 1.2. All customers have the ability to delete the voice recordings associated with their account; enterprise users by using the Alexa for Business console, and personal users by using the Alexa app or the Amazon website. Customers using personal devices (enrolled users) can also use the Alexa app to view and play back the voice recordings associated with their account.

Automatic Speech Recognition (ASR)

The first Alexa system to receive data is Automatic Speech Recognition (ASR). It takes the audio stream and turns it into a text string (or set of possible text strings) that are sent to the Natural Language Understanding (NLU) system. In our case, possible strings could be:

- “what is the weather in Menlo park”
- “watt is the weather in Menlo park”
- “what is the weather in Menlo Park”
- “watt is the weather in men lo park”

These strings, and their related confidence scores, are used to improve speech recognition. The string with the highest score (i.e., the one that Alexa acts on) is also stored and is displayed to the user in the Alexa app.

Natural Language Understanding (NLU)

The “Natural Language Understanding” (NLU) interprets the recognition result and produces an intent. The NLU performs:

- intent classification (determining in this case that the user wishes to get the weather, and returning a Weather intent),
- entity recognition (determining that the user requested the location “Menlo Park”), and
- slot resolution (determining the location identifier for the location “Menlo Park” that

can later be used to retrieve the correct weather for that location).

The service now looks at the intent as “Weather” and routes the request to the proper application (Skill) with the slots filled in for location (94025) and time (today). The data about the chosen intent, and information related to entity recognition and slot resolution, is stored for machine learning purposes.

Skills

Skills extend what the Alexa system can do, and we’ve designed our skills program to share only limited information with the third-party developers of those skills. For example, voice recordings are not shared with skills.

In the above example, the skill takes the input (location (Menlo Park) and time (today)) and retrieves the appropriate information from the designated data source, which returns the needed data. The skill then formulates its response, taking the raw data (in this case, temperatures as well as the forecast) and constructing a textual response formatted with SSML (simple speech markup language) which tells the next step, TTS, how to respond. Once the response is generated, it is sent to the response system.

Customer’s personal information (e.g. name, address) are not released to the 3rd-party unless specifically requested to be shared by the customer. We also use a permission framework similar to the one used by mobile devices, which requires customers to grant permission to share certain data with skill developers–e.g., Lyft could request permission to access the address the customer has set for their Echo device so Lyft can send a ride to that location, and we would only share that address with Lyft after the customer granted permission. Even when a customer links their Amazon account to a 3rd-party skill account (e.g., when a customer links their account with Lyft, so that Lyft always knows which Lyft account to charge rides to), the 3rd-party doesn’t receive the customer’s Amazon account identifiers. Instead, they receive a token. Amazon can identify an Amazon customer from this token, but skill developers cannot. Each time a customer talks to a skill, the skill gets the same token for that user. However, different skills get different tokens when they talk to the same user, so even if skill developers share data, they cannot determine that they share common customers based on data that Amazon has sharedEach 3rd-party skill has their own policies concerning storage and retention of skill-specific data. Customers can read the 3rd-party skill developer privacy policies, and see what

customer information is being asked for, on the skill detail page on the Amazon website.

Responses (TTS)

The response system takes the SSML that is produced by the skill, uses text-to-speech (TTS) to generate the audio speech file, and streams the audio to the appropriate device. For many skills, this ends the interaction. Other skills are interactive and will ask follow-on questions that require answers. Echo devices are designed Alexa Privacy and Data Handling Overview (20180720) Page 4 of 10 so that the blue ring on the Echo device is lit when the device is waiting for a response to a question that Alexa has asked. The text of the response is stored by the Alexa system so that users of personal devices can review past answers using the Alexa app. Access to this data is not available to Alexa for Business users or administrators for devices managed by Alexa for Business. In addition, the response can be used by the Amazon team who built the specific skill to ensure that Alexa is providing relevant answers to queries and that the TTS system is properly translating the text to speech.

How Does an Echo Detect its Wake Word and Send Audio to the Alexa Cloud?

Amazon Echo devices are designed to use on-device keyword spotting to detect the wake word and only the wake word. Unless the microphone is turned off (discussed below), this technology inspects acoustic patterns in the room to detect when the wake word has been spoken using a short, on-device buffer that is continuously overwritten. There are multiple algorithms running on the Echo device looking for the specified wake word. At this point no audio is sent to the Alexa cloud.

If the algorithms do not detect the wake word, then the Echo device continues to wait for the wake word, continuously overwriting the contents of the small internal audio buffer. Importantly, Echo devices do not keep local records of audio; they keep only a small amount of audio to detect the wake word. This on-device buffer exists in temporary memory (RAM); audio is not

recorded to any on-device storage.
When the wake word is detected or the action button is pressed, a connection to the cloud is opened up. The Echo device turns on the blue ring and starts streaming the audio, starting with a fraction of a second of audio prior to the wake word and continuing until the Alexa system in the cloud turns off the audio stream. Echo devices use a signal processing technique called beam forming to emphasize the user's speech from the desired direction while suppressing audio interference (like conversations outside the room) from other directions. Customers see beamforming take place on an Echo device with a visual cue—the lightest blue color on the light indicator points towards the source of the audio that is being recorded.

If Alexa is activated using the wake word, the first step that occurs when the stream reaches the cloud is that the audio is reanalyzed using the more powerful processing capabilities of the cloud to verify the wake word was spoken. These additional algorithms are in the cloud, and not on the device, for reasons including requiring more processing power than the Echo device has available or using machine-learning derived models based on recent learnings. The on-device algorithms are automatically updated on a regular basis. If this cloud software verification is unable to confirm the wake word was spoken, the Alexa system stops processing the audio. If the wake word is verified (or if Alexa was activated using the action button), our ASR and NLU systems process the customer's request so Alexa can respond appropriately.
As our speech recognition system analyzes the audio stream, the system continually attempts to determine when the customer's request to Alexa has ended and then immediately ends the audio stream. The light ring then typically flashes blue/light blue until the response is ready for playback. It then sends the response (the blue ring pulses while Alexa is speaking), and the Echo device returns to monitoring for its wake word.

TO SUMMARIZE HOW AN ECHO DEVICE WORKS:

1. Unless the mic mute button is lit red, the Echo device is analyzing audio to detect its wake word and only its wake word. No audio is sent to the Alexa cloud while this is happening.
2. Wake word detection is done on the Echo device.
3. Only when the wake word is detected or the action button on the device is pressed does audio stream to Alexa in the cloud.

4. When audio is streaming to Alexa, the blue light is turned on.
5. Alexa is designed to stop processing the audio if it determines the speech is not intended for Alexa.
6. The audio stream closes immediately once our ASR system determines the customer has stopped speaking the request

Preventing an Echo Device from Responding

We make it very easy for anyone using the device to turn the microphones off. There may be certain times that users may not want the Echo device to respond to the wake word. All Echo devices have a microphone on/off button. When this button is pressed, the microphones are electrically disconnected and a dedicated red LED on the microphone button is illuminated to indicate the microphones are off. As an additional safeguard, we designed the circuitry of Echo devices so that power can only be provided either to this dedicated red LED or to the device microphones, not to both at the same time. As a result, if the dedicated red LED is illuminated, the microphones are off and cannot stream audio to the cloud. The system is designed so when in this state, the microphone cannot be turned back on by software or by voice (and if turned off prior to removing power from the Echo device, will still be off when the power is restored).

Alexa Calling

When using Alexa calling, the light ring (or bar) will glow green, to indicate that the microphone is on and audio is streaming. In this case, the audio is not being streamed to the Alexa ASR and NLU systems. Instead, it is being routed to either another Alexa-enabled product (for Alexa-to-Alexa calling) or to the phone system (if placing a telephone call). Calls are not recorded by Alexa. During the call, the Echo device is still monitoring for its wake word, so that you can say things like “Alexa, end call”. Even during a call, you will Alexa Privacy and Data Handling Overview (20180720) Page 6 of 10 note that the light ring on the Echo device turns blue when its wake word is detected, to show that the audio directly after the wake word is going to Alexa.

Data Storage

Different types of data are used and stored by the Alexa system to provide the Alexa service. Configuration parameters are set by the user either on the device or using the Alexa app. These parameters include such Alexa Privacy and Data Handling Overview (20180720) Page 7 of 10 things as the device location (set by the administrator or user), preferred time zone and unit measures, volume level, and other preferences. Audio and text inputs are the core piece of Alexa data. As described above, voice recordings are processed through speech-to-text algorithms and then through natural language processing algorithms to extract the user's intent and the parameters of the Alexa query. These systems use machine learning techniques to continuously improve themselves with each input.

Objectives

Nearly anything you can do with your device; Alexa can do through simple voice commands. This feature allows users to navigate their smart devices hands-free. Some of the popular tasks you can accomplish with Alexa include:

- Finding recipes and reading them out loud
- Controlling smart home tech
- Setting timers or alarms
- Performing internet searches
- Controlling music and podcasts
- Checking the local weather
- Reading news reports or sports updates
- Playing audiobooks
- Being a calculator or converting measurements
- Telling jokes or stories
- Finding nearby vaccines
- Helping shop on Amazon

Project undertaken

(Restaurant)

Introduction

Restaurant is a Alexa skill which is design with the help of voice flow platform. Voice flow is a platform of amazon Alexa skill where we can create skill of Alexa. It asks to order on restaurant. It ask for food which you want to order. The information of food is already added to the skill it shows the menu and we can order the food easily with the help of voice. After order it takes the feedback.

Working

Step 1:

You have to say “Alexa order the food”

Step 2:

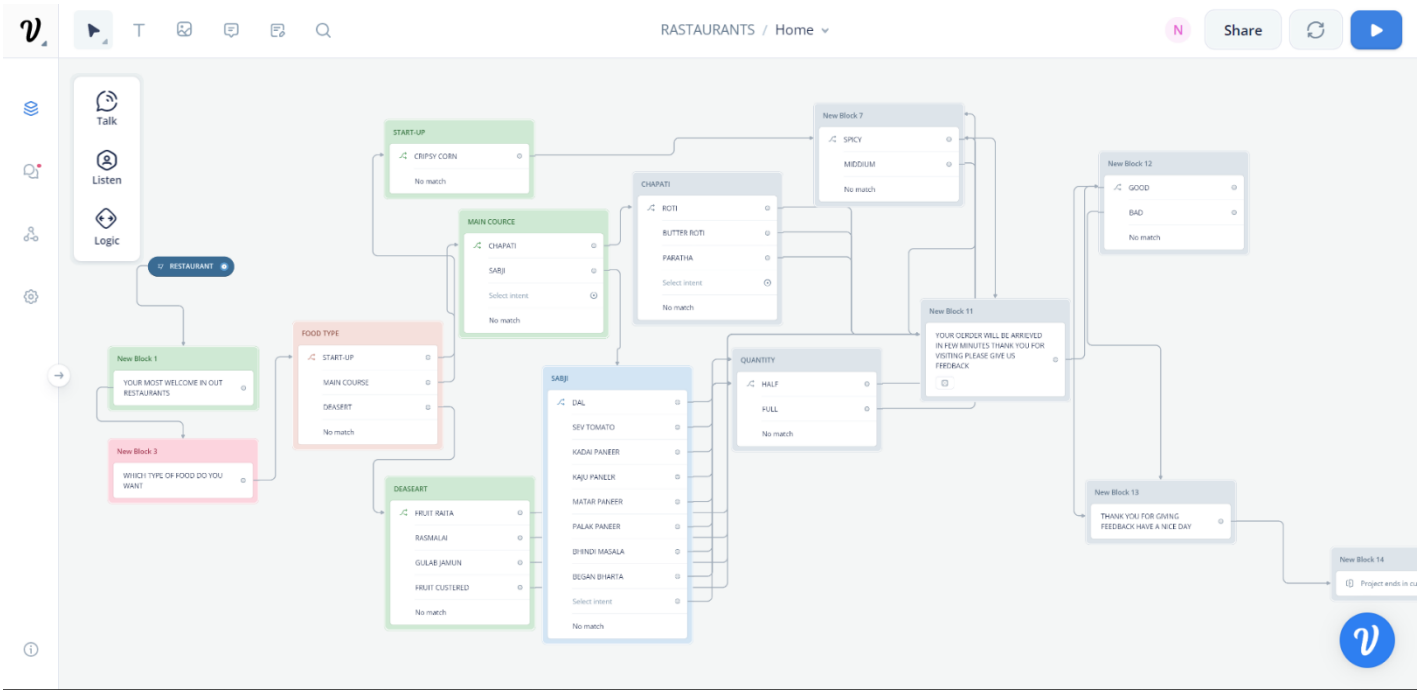
It shows menu and you must choose the food from menu.

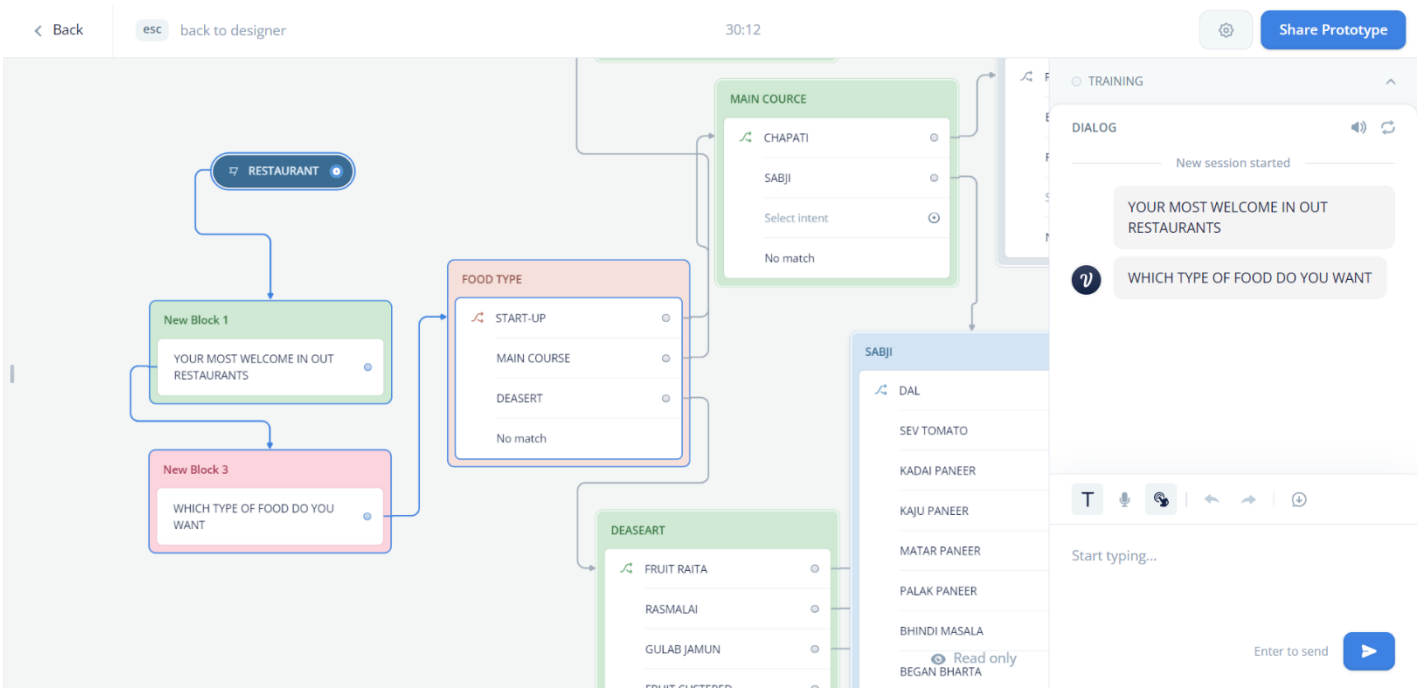
Step 3:

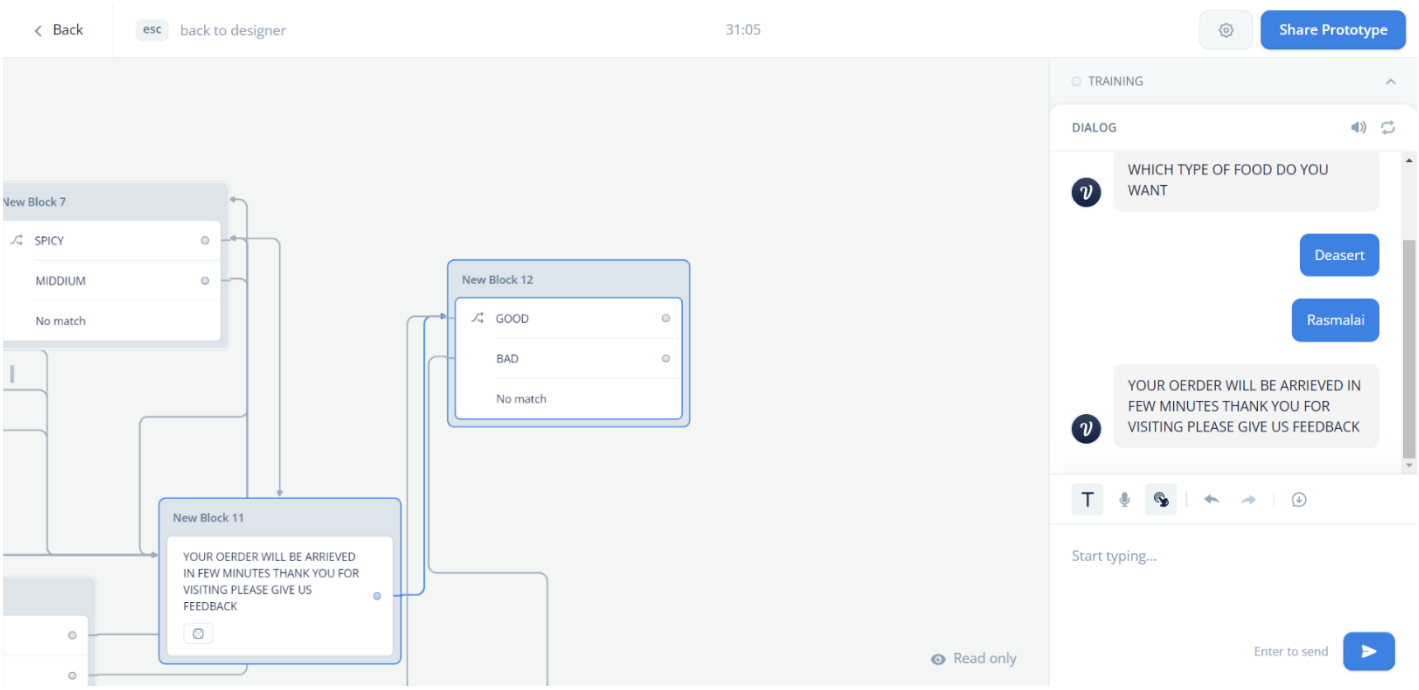
It send the information of food to the counter.

Step 4: The food will be delivered in some time.

Screenshots of Project




Back
esc
back to designer
30:12
Share Prototype
RESTAURANT
New Block 1
YOUR MOST WELCOME IN OUT RESTAURANTS
New Block 3
WHICH TYPE OF FOOD DO YOU WANT
FOOD TYPE
START-UP
MAIN COURSE
DEASERT
No match
MAIN COURCE
CHAPATI
SABJI
Select intent
No match
SABJI
DAL
SEV TOMATO
KADAI PANEER
KAJU PANEER
MATAR PANEER
PALAK PANEER
BHINDI MASALA
Read only
BEGAN BHARTA
DEASEART
FRUIT RAITA
RASMALAI
GULAB JAMUN
TRAINING
DIALOG
New session started
YOUR MOST WELCOME IN OUT RESTAURANTS
WHICH TYPE OF FOOD DO YOU WANT
Start typing...
Enter to send


Back
esc
back to designer
31:05
Share Prototype
New Block 7
SPICY
MIDDIUM
No match
New Block 12
GOOD
BAD
No match
New Block 11
YOUR OERDER WILL BE ARRIEVED IN FEW MINUTES THANK YOU FOR VISITING PLEASE GIVE US FEEDBACK
Read only
TRAINING
DIALOG
WHICH TYPE OF FOOD DO YOU WANT
Deasert
Rasmalai
YOUR OERDER WILL BE ARRIEVED IN FEW MINUTES THANK YOU FOR VISITING PLEASE GIVE US FEEDBACK
Start typing...
Enter to send

Conclusion

the latest trends in the market of systems for smart homes were studied. The work focused on developing a system that is not inferior to existing counterparts and, at the same time, will be an effective tool to meet user needs. The paper analysed the literature and other sources and explored the most popular voice assistants and systems that can be integrated. We also examined the applications and applications supported by the system and determined the best voice assistant to control this smart home system and connect all possible sensors. Because voice assistants by themselves do not allow you to make a smart home interface and all interaction with them takes place exclusively through external APIs.

References

https://developer.amazon.com/alexa
www.youtube.com
https://en.wikipedia.org/wiki/Amazon_Alexa