iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 Issue: II Month of publication: February

DOI: http://doi.org/10.22214/ijraset.2019.2137

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue II, Feb 2019- Available at www.ijraset.com

Voice Enabled Railway Assistant using Alexa Skills Kit for the Blind and Visually Impaired

Mrinal Bageshwari[1], Shreya Hegiste[2], Nitin Harane[3]
[1, 2, 3]*Computer Engineering, Thakur College of Engineering & Technology, Mumbai University*

Abstract: Blind and the visually impaired have a little or no access to the advancements in technology and the world of internet. This is because of the absence of affordable solutions and products (hardware and software). Existing technology that enables the blind and the visually impaired to make use of the internet or information in any digital form highly depends on Braille displays and keyboards which are scanty and expensive. Another limitation of the existing technology is that out of all the visually impaired population, not more than 2% know how to interpret Braille. Consequently, voice-controlled systems which transmit and receive information in the form of audio have already proven their usefulness to the blind and the visually impaired. Today's real-world voice-controlled systems are becoming smarter at an increasing pace. They tend to learn and understand the end user's speech patterns over the course of time to a degree that they are even capable of building their own vocabulary. With the speedily reducing cost of electronic hardware, way more pioneering approaches are practicable to consider for implementation in products. These products make use of algorithms that are linguistically sophisticated. The algorithms used are paired with extensive dictionaries. Hence, a new functionality providing dynamic information to the blind or the visually impaired to plan their daily commute using trains was designed using Alexa Skills Kit. The goal of this system is to build a voice assistant which is small in size and low at price to help the blind travel with a greater degree of psychological comfort and giving optimal schedules to reach the desired destination.

Keywords: AWS (Alexa Web Services), Amazon Alexa, Voice Recognition, ASK (Alexa Skills Kit), AWS Lambda, AVS (Alexa Voice Services), VUI (Voice User Interface), GUI (Graphical User Interface), NLP (Natural Language Processing).

I. INTRODUCTION

This system is a voice assistant for the Mumbai railway which is being developed by using the Alexa Skills Kit. The proposed system uses Alexa as its voice. Alexa is Amazon's cloud-based, virtual personal assistant which powers the family of Amazon devices and supports as the companion app on the Android and iOS smartphones as well. The tasks Alexa executes upon user request are called "Alexa Skills". In essence, an Alexa Skill is a voice-driven Alexa app. This system will help users plan their daily commute of the local trains of Mumbai. Voice being a rapidly emerging technology can be used to lend a hand to the visually impaired. Scaling the speech models to accommodate such a huge foreign dataset is a fascinating challenge. An estimated 7 million people living in Mumbai use the local trains daily to commute. Train stations can get crowded and using the assisting Braille systems at the stations can be a difficult task. Thus, making use of a VUI in such a situation can be of major use for the mentioned users. Fig. 1 shows the possible communication between a user and the device.

USER- Alexa, launch where's my local
USER- when can I go to Borivali from Churchgate around 4
ALEXA- The next train is at 4:08 pm coming at platform no. 3
USER- Thank you, Alexa!

Fig. 1. Structure of a sample User-Device Interaction

II. RELATED WORK

There are many different technologies to aid the visually impaired available such as voice stick, GPS based navigation system, the vOICe technology and the m-indicator application. The m-indicator application basically provides the timings of all public transport of the city. But this application is a self-help, screen- browse application. Developing an application which helps commuters plan their daily travel by using voice as a commodity is the basic principle of our system. It will help the visually impaired to know timings of the local trains, making accessibility one of the most important features of the proposed system. Another limitation found in this existing system is the use of a static database. The local train database is usually updated and modified according to the delays and changes in the platforms at which the train is going to arrive at.

III.DESIGN FOR USABILITY FOR BLIND AND VISUALLY-IMPAIRED USERS

Following are some basic concepts considered in order to design an application specifically for visually impaired user with an example.

A. Usability User Experience

Usability can be described as the overall experience of a person using a particular system or a product such as a computer application or a web application, especially in terms of how easy or pleasant it is. Usability can be one of the most efficient parameters to measure or understand the effectiveness of the system/ product. In touch screen-based devices, the description of guidelines and principles that help developers to design a formidable user interface and a pleasant experience for the user of any app is an example of this term. These guidelines are oriented to design apps for sighted users. [27]

B. Accessibility

This term can be understood as the degree to which a certain product, device, service, or environment is easy to reach to as many people as possible. Consider the example of Apple devices like iPhone and iPad that include a set of features specifically designed to provide accessibility to users with special needs. Some of these features are: Voice-based GPS system for blind users, Voice Stick, Color ID Free, Speech being converted to text, large fonts, hands-free microphone, audible, visible and vibrating alerts or assignable ringtones. [27]

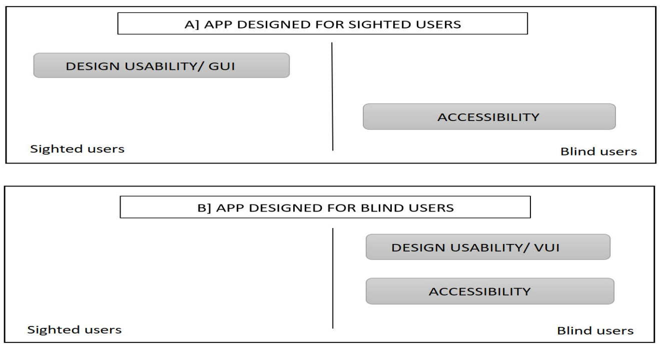


Fig. 2. App Design for A] Sighted Users and B] Blind Users

Figure 2. A] shows the design situation where an application is designed for sighted users with an extra layer including accessibility features. Even though blind users could use this application, the user interface is not essentially planned for blind users. In light of such an event, design for usability is only considered for sighted users and user experience for the blind is not entirely promised. [27]
Figure 2. B] shows a different design scenario where an application has been designed for blind users. Such usability design applies directly to visually impaired users. It improves the usability design of the UI, by keeping in mind the size, shape and position of the controls. Enhancing accessibility features by using voice as a commodity is not just an extra layerbut it connects directly to how the user feels that he/she is given the control to the system. [27]
Thus, the proposed system has considered of conceiving accessibility and usability features for the blind users from initial planning stages.

IV. TECHNOLOGY STACK

A number of technical components have been integrated in this project. These components include a combination of programming languages and software products. [13] A seamless functionality has been established amongst these components. The major building blocks of the technological stack is delineated as follows:

A. Physical Layer

The devices with which the user interaction takes place are present in this layer. They are as follows:

1) Amazon Echo and Amazon Echo Dot - It is a smart speaker that can be controlled using voice. On the basis of the request made, a response is returned to the user.
2) Amazon Fire Tablet- It is a tablet computer developed by Amazon.com. Fire tablet can be used as an alternative to Amazon Echo. It provides similar functionalities with additional rear and front facing cameras.

B. Application Layer

The application layer consists of the following components:

1) Alexa Skills Kit (ASK) [25] - It is a Software Development Kit (SDK) for developing custom skills for the various Amazon devices available in the market.
2) Amazon Web Services (AWS) Lambda [26] – It is an event-driven, serverless computing platform. It runs programs on invocation instead of hosting programs on a server.

C. Programming Layer

The source codes of our project are written in JSON and Python. [13]

V. METHODOLOGY

A. System Architecture

The system is shown as follows in Figure 3.

The entire system architecture comprises of following major phases:

1) User gives voice command
2) Intent Detection
3) Intent Processing
4) Device provides the user with a Voice Response

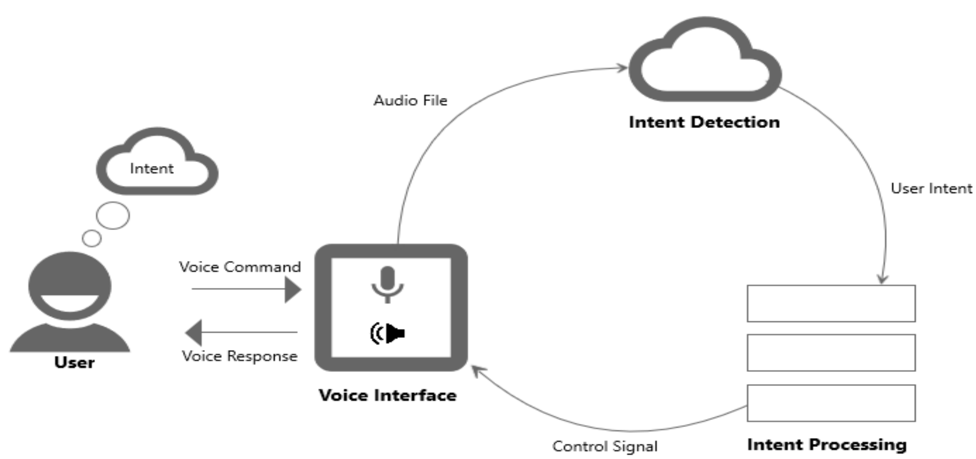


Fig. 3. Voice System Architecture

B. System Components

In this section, each component's role to make the system function is briefly explained.

1) *User:* A user in the proposed system is any person with an Echo device or a companion app device to Amazon who wants to plan their commute by the local trains of Mumbai city. Commuters can invoke the skill by voicing out the skill name which in this case is "Where's my local?" and following that he/she can request for the required train timing.
2) *Device:* This application makes the use of ASK and AWS to store the databases. Thus, this system can run on Alexa-enabled Amazon devices and also the supporting companion Android/ iOS devices. These devices will listen to the user's speech requests and return appropriate responses to the user.
3) *Skill Interface:* The skill interface basically processes the requests made by the user. This processing includes taking the user "requests/orders" and mapping it to the matching intent within the interaction model. The intents are actions that fulfil or carry out the spoken requests from the commuter. There are predefined utterances, words that the user might say in order to invoke the skill. Every intent is specified in one utterance. Detection of an intent results in the creation of a JSON encoded event by the Skill Interface. This event is passed further to the Skill Service.[27]

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue II, Feb 2019- Available at www.ijraset.com

4) *Skill Service:* The skill service deduces what actions to take as a reaction to the request of the user. It is a code that can be written in Python, Java, nodejs. This request is an encoded event in JSON with the skill interface. The interface sends this event to the Skill service. Once a decision is reached, the skill service returns response to the skill interface in the form of a JSON event for further processing. The speech response is sent back to the user through the Echo after the further processing is done. [27]

C. Algorithm for building the System /Modules for developing the System [23]

1) To design a voice user interface which explains how the user interacts with the skill.
2) Using the Developer Console to set up the skill.
3) Using the voice design to build your interaction model. Setting up the interaction model includes creation of utterances, intents, slots and slot types. JSON is used for this model.
4) Creating an AWS Lambda function for the skill.
5) Beta testing of the skill by making it available to the users to review the applications' functioning (optional).

VI. PROPOSED SYSTEM

The proposed system has the following features:

1) Unlike other voice assistants that work with Static data, this works with dynamic data providing way better functionality. Users get access to information about delayed trains, change in platform numbers, etc.
2) Auto synchronization used keeps the voice assistant updated with the data on its cloud server. [16]
3) Provides with an option to keep a "Wake me up" alarm to avoid missing your station.
4) Allows personalization of information and enables the user with opportunities to save favourite journeys, stations and trains. [21]

VII. WORKING OF THE SYSTEM

Alexa is built on the basis of Natural Language Processing (NLP) which is a procedure of converting speech into words, sounds, and even ideas. [22] When a user requests/orders the device, Amazon breaks the user's "orders" down into individual sounds. It further seeks information from a database which contains the pronunciations of words. It does this to pick the words that most closely correspond to the amalgam of each individual sounds. Furthermore, it recognises important words for identification of the tasks. Consequently, it carries out the corresponding functions. For instance, if Alexa observes words like "game" or "cricket", it would open the sports application. Amazon's servers revert the information to the user's device and Alexa speaks. In case Alexa needs to say anything in return, it goes through the same procedure in the reverse order.

VIII. SIMULATED RESULTS

A. Design of a VUI

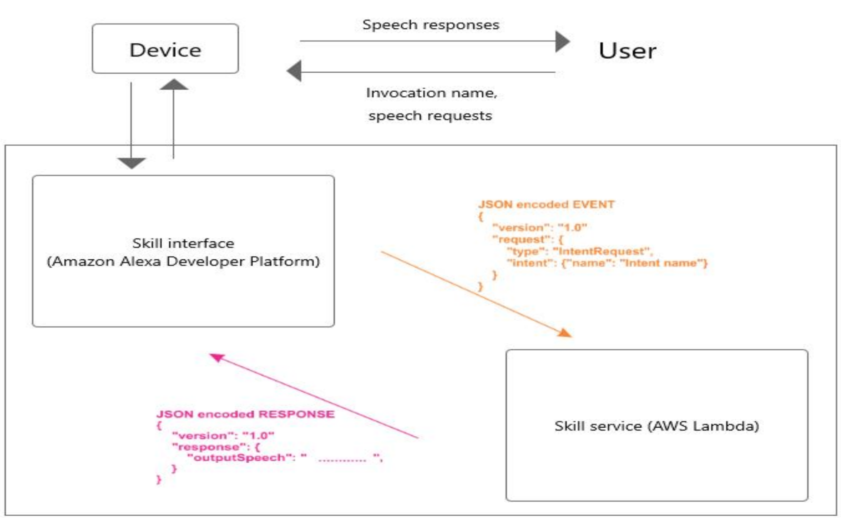


Fig. 4. Voice User Interface

B. Building an Interaction Model using the VUI

1) Beginning with setting up a skill in the Amazon Developer Console, the custom skill was named- Where's my Local? A study of the various interaction models available on the internet was undertaken prior to building the interaction model for this system. An interaction model refers to the collection of sample utterances, intents, etc. [23]

2) The requests/orders a skill can take care of are called as intents. An intent is an action that handles and fulfills a user's spoken request/order.

3) Intents can have arguments which are called slots. These slots are optional. Slots are a representative list of all possible values.

4) Sample utterances are used for mapping the intents to the words and phrases which the users can say in order to interact with a particular skill. This includes all possible representative phrases, making it an extensive list.

5) Example:

INTENT- songintent
SAMPLE UTTERANCES-
I want to listen to baby by Bieber
I want to listen to shallow by JZ
SLOTS- {title}, {artist}
SLOT VALUES- {title} baby, shallow, closer, etc.
{artist} Bieber, JZ, Gaga, etc.

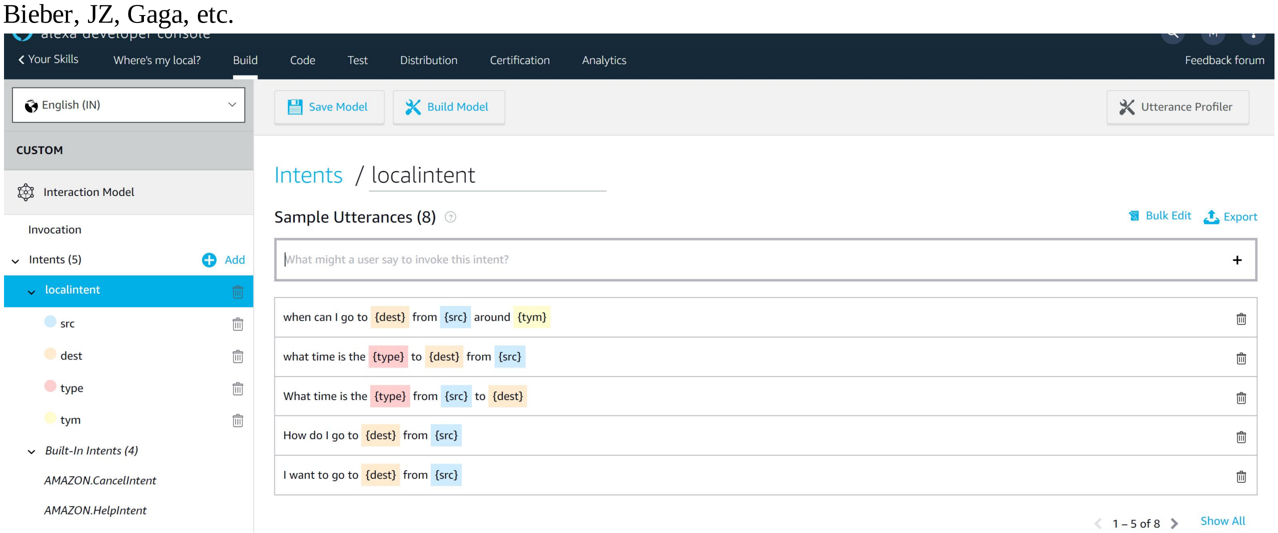


Fig. 5. Alexa Developer Console: Sample Utterances

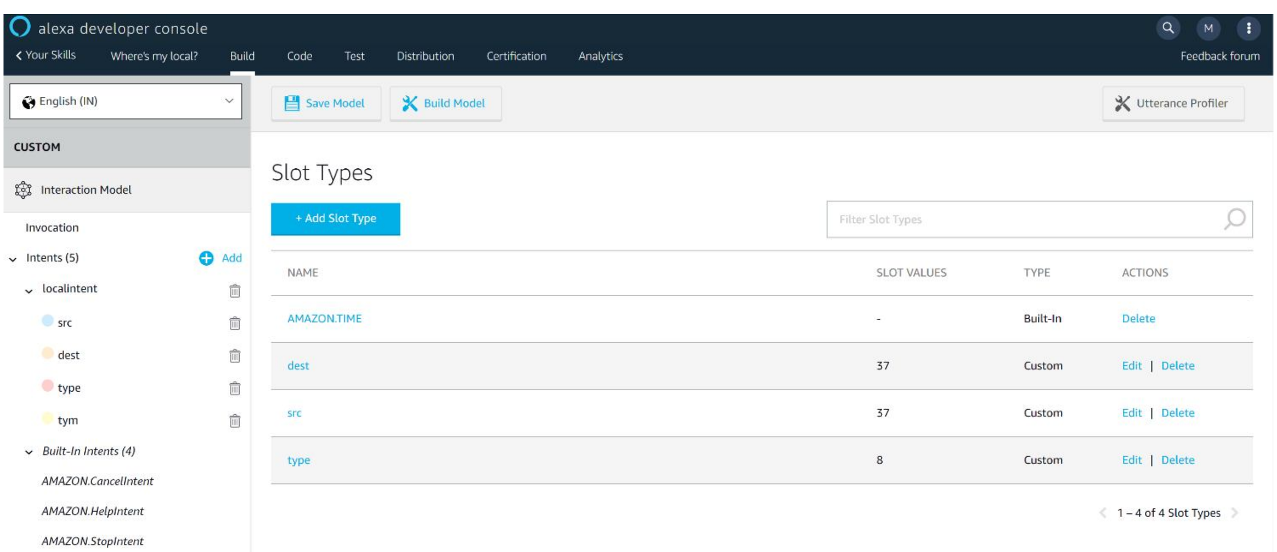


Fig. 6. Alexa Developer Console: Slot Types

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue II, Feb 2019- Available at www.ijraset.com

C. Analysis of a "Request/Order"

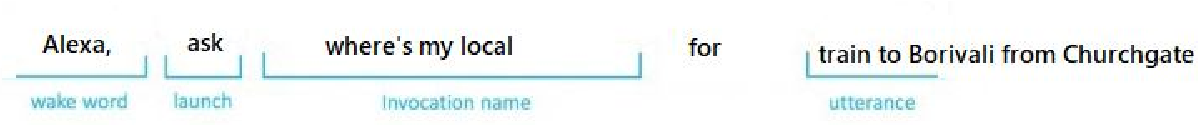


Fig. 7. Structure of a sample "request/order"

The above command has three important parts: Wake word, Invocation name and Utterance. [22]

1) *Wake word:* This term is analogous to a persons' name i.e we call a persons' name when we want to address them. Similarly, to address Alexa, we need to use the wake word. Once this word is registered by Alexa, she is all set to take in the user's requests.
2) *Invocation Name:* Just like the term suggests, this word is used to invoke the created skill. It is a keyword that can be put together with any command, action or doubt.
3) *Utterance:* 'train to Borivali from Churchgate' is utterance. These are the phrases that the user tends to use while making a request to Alexa. She recognizes the user's intent from the utterance and responds to the user accordingly.

D. Testing the "Request/Order"

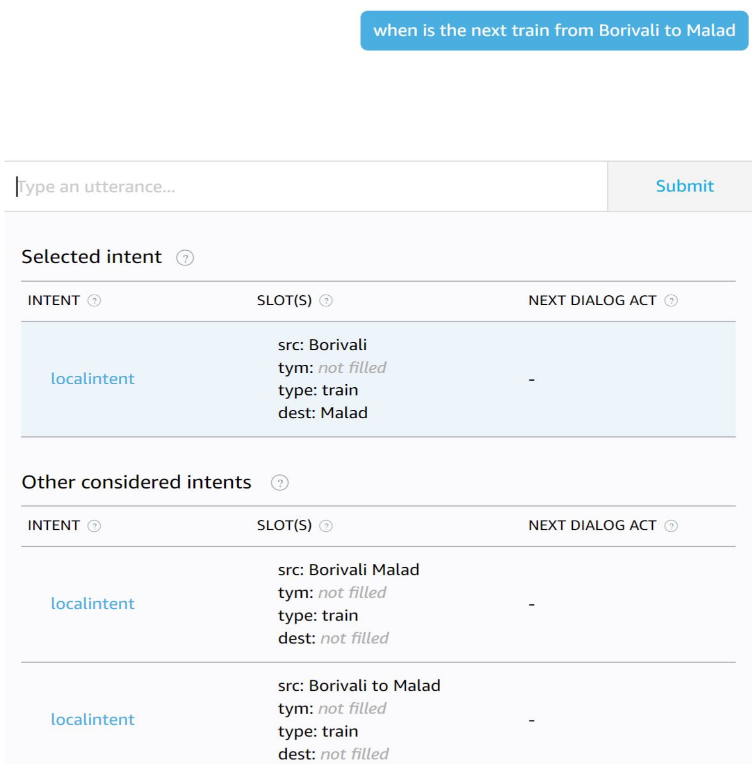


Fig. 8. Alexa Developer Console: Testing the structure of a sample "request/order"

IX. CONCLUSION

Based on the scientific breakthrough about speech synthesis and using efficient hardware evolutions along with powerful data processing, we have developed a railway voice assistant with unique functionalities, specially designed for the blind and the visually impaired focusing on creating a time-efficient interaction. Like GUIs try to limit the number of clicks a user has to make in order to complete an action, we have limited the load of voice interaction needed using our VUI. Unlike most systems that have been developed so far which lack the usage of dynamic data, our railway voice assistant uses it as a building block. We emphasize the functionality of this railway voice assistant to augment a blind and visually impaired commuter's experience with adequate and appropriate information to make him/her feel comfortable while traveling. This VUI also holds potential to be extended for devices with screens. Our motive was to find out how this VUI can create an impact on the people with special needs. With more work using the same VUI, our model possesses the capability of creating great and inclusive experiences for all users.

X. ACKNOWLEDGMENT

We would like to thank our Project Coordinator Mr. Anand Khandare for his expert advice and extraordinary support throughout this challenging project. We are grateful to the Head of Department, Dr. Sheetal Rathi and Dean Academic, Dr, R.R. Sedamkar for the encouragement and for creating a learning environment which has proved to be very valuable. We would also want to express our sincere gratitude towards Amazon.com for providing us with the Alexa Developer Console.

REFERENCES

[1] Aditi Bhalerao, Samira Bhilare, Anagha Bondade, Monal Shingade, “Smart Voice Assistant: a universal voice control solution for non-visual access to the Android operating system”, International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue: 01 | Jan -2017

[2] Granstrom, B., “Speech technology for the visually impaired - the Swedish perspective”, http://www.speech.kth.se/qpsr/

[3] Prof. Priya U. Thakare, Kote Shubham, Pawale Ankit, Rajguru Ajinkya, Shelke Om, “Smart Assistance System for the Visually Impaired”, International Journal of Scientific and Research Publications, Volume 7, Issue 12, December 2017 ISSN 2250-3153

Gurubaran, Gowrishankar Kasilingam, Mritha Ramalingam, “A Survey of Voice Aided Electronic Stick for Visually Impaired People”, International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163, Volume 1 Issue 8 (September 2014)

[4] Sanjana.B, J.Rejina Parvin, “Voice Assisted Text Reading System for Visually Impaired Persons Using TTS Method”

[5] Vincent Gaudissart, Silvio Ferreira, Céline Thillou, Bernard Gosselin, “SYPOLE: A MOBILE ASSISTANT FOR THE BLIND”, https://core.ac.uk/download/pdf/30467332.pdf

[6] GPS Based Voice Navigation System for Visually Impaired People, https://www.elprocus.com/gps-based-voice-alert-system-for-blind-people/

[7] Luke Scriven, “Five Tips for Teaching Speech Recognition to People with a Visual or Physical Impairment”,

[8] http://www.afb.org/info/programs-and-services/professional-development/technology/five-tips-for-teaching-speech-recognition-to-people-with-a-visual-or-physical-impairment/1235

[9] Mrs. S. S. Lavhate, Mr. Mandar Joshi, Mr. Akash Chhedha, “Secure Navigation System for visually impaired”, International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 4, Issue: 3

[10] Lisa Ran, Sumi Helal and Steve Moore, “Drishti: An Integrated Indoor/Outdoor Blind Navigation System and Service”, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.66.2419&rep=rep1&type=pdf

[11] Javier Sánchez Sierra, Joaquín Selva Roca de Togores, “Designing Mobile Apps for Visually Impaired and Blind Users”, ACHI 2012: The Fifth International Conference on Advances in Computer-Human Interactions

[12] Xinyu Lei , Guan-Hua Tu , Alex X. Liu , Kamran Ali , Chi-Yu Li , Tian Xie “The Insecurity of Home Digital Voice Assistants – Amazon Alexa as a Case Study”

[13] Achal S Kaundinya, Nikhil S P Atreyas,Smrithi Srinivas,Vidya Kehav , Naveen Kumar M R “Voice Enabled Home Automation Using Amazon Echo”

[14] How does the Amazon Echo help people with a vision impairment? https://www.rnib.org.uk/nb-online/how-does-amazon-echo-help-disabled-people

[15] Why Amazon’s Alexa is ‘Life Changing’ for the Blind, https://in.pcmag.com/news/118427/why-amazons-alexa-is-life-changing-for-the-blind

[16] Abhay Dekate, Chaitanya Kulkarni, Rohan Killedar, “Study of Voice Controlled Personal Assistant Device”

[17] Dr. Kshama V. Kulhalli, Dr.Kotrappa Sirbi, Mr. Abhijit J. Patankar,” Personal Assistant with Voice Recognition Intelligence”

[18] Prince Bose,Apurva Malpthak ,Utkarsh Bansal ,Ashish Harsola, “Digital assistant for the blind”

[19] Overcoming Challenges in Voice-Based Natural Language Processing (NLP) in Business, https://emerj.com/ai-podcast-interviews/spoken-voice-based-nlp-for-business/

[20] Trainline launches UK’s most advanced voice activated rail app, https://www.globalrailwayreview.com/news/64172/trainline-launches-uks-advanced-voice-activated-rail-app/

[21] National Rail Enquires, http://www.nationalrail.co.uk/times_fares/84874.aspx

[22] How Amazon Alexa works? Your guide to Natural Language Processing (AI), https://towardsdatascience.com/how-amazon-alexa-works-your-guide-to-natural-language-processing-ai-7506004709d3

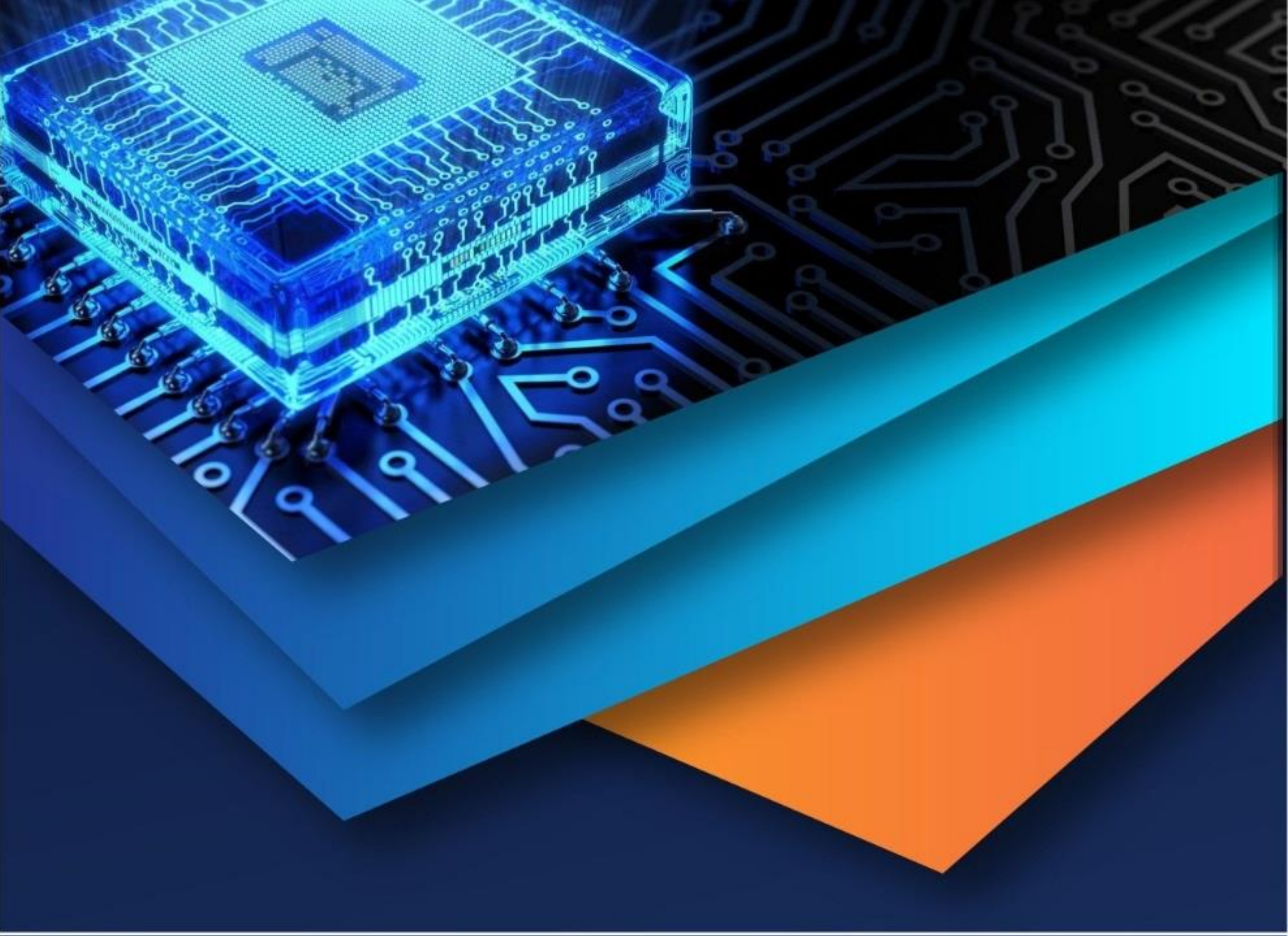
[23] Steps to Build a Custom Skill, https://developer.amazon.com/docs/custom-skills/steps-to-build-a-custom-skill.html

[24] Machine Learning in Practice, https://www.forbes.com/sites/bernardmarr/2018/10/05/how-does-amazons-alexa-really-work/#42413461937f

[25] Alexa Skills Kit, https://developer.amazon.com/alexa-skills-kit

[26] AWS Lambda, https://aws.amazon.com/lambda/?sc_channel=PS&sc_campaign=acquisition_IN&sc_publisher=google&sc_medium=ACQ-P%7CPS-GO%7CBrand%7CDesktop%7CSU%7CCompute%7CLambda%7CIN%7CEN%7CText&sc_content=lambda_bmm&sc_detail=%2Baws%20%2Blambda&sc_category=Compute&sc_segment=293611557117&sc_matchtype=b&sc_country=IN&s_kwcid=AL!4422!3!293611557117!b!!g!!%2Baws%20%2Blambda&ef_id=CjwKCAiAnsnjBRB6EiwATkM1Xmoo7fdi1aJtKuIR-GnQMAgnLQ2_uZ1gy7xwhbPxRWIrm_hcY1lfRhoC_qwQAvD_BwE:G:s

[27] Uros Ralevic, “How To Build A Custom Amazon Alexa Skill, Step-By-Step: My Favorite Chess Player”, https://medium.com/crowdbotics/how-to-build-a-custom-amazon-alexa-skill-step-by-step-my-favorite-chess-player-dcc0edae53fb



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY