

# Fall 2021 - CS 687 Capstone Project Report

## Developing a Voice assistant Chatbot with Alexa Skills

Sreelekha Sangitha

Advisor: Sion Yoon

MS in Computer Science

School of Technology & Computing (STC)

City University of Seattle (CityU)

[[sangithasreelekha@cityuniversity.edu](mailto:sangithasreelekha@cityuniversity.edu), [yunhee@cityu.edu](mailto:yunhee@cityu.edu)]

### Abstract

In recent years, many breakthrough advancements have been made in the field of intelligent voice assistants with the help of Machine learning technologies like Natural Language Processing (NLP), Natural Language Generation (NLG) and Deep Neural Networks (DNN). Several day-to-day business problems can be solved by developing skills for voice assistants that are backed by technologies based on Artificial intelligence and Machine learning. One of the essential components for any business is providing readily available support to their customers. Chatbots are essential and have become one of the essential features for any business, but with a limitation of being available to customers who visit their website or support page. To create new opportunities for a business, intelligent voice assistant skills can be developed by creating a conversational chatbot that can interact and respond to customer questions. This project researches on leveraging available tools and technologies like AWS, Lambda, Alexa skill kit (ASK) to build a simple and functional skill for voice assistant Alexa.

**Keywords:** Chatbots, Machine Learning, voice assistant, NLP, Technologies, DNN, Alexa Skill.

### 1. INTRODUCTION

Chatbots are dominantly being used by many websites to help customers answer the most commonly asked questions in an interactive way tailored to the questions a customer has. (Paikari & van der, 2018). However, customers must use a browser on a physical device to open the website and interact with the chatbot. With many hands-free devices in the market and homes, customers are increasingly relying on voice assistance to get information about a product or topic that they are interested in. This project explores the possibility of integrating the chatbot with the most convenient voice assistant so that the customer can get answers about our product using a hands-free voice assistant.

### Problem Statement

Any customer-facing corporation or business needs to provide customer support to address the questions and problems of its customers. As the scale of the company grows, it becomes more and more difficult to scale up the customer support team, and they turn into automated solutions like a chatbot. A chatbot can be programmed to serve many of the customers' needs, like answering the most common questions about the business and providing support and pointers with the problems a customer can face. The scope of many chatbots is currently limited to providing support when the customer visits their website on a mobile or desktop browser.

With evolving technology and the integration of hands-free voice assistants in compact devices, many consumers nowadays are relying on getting

answers from voice assistants rather than physically visiting the website on a browser. This project will focus on narrowing down the gap between customer expectation to get answers from a voice assistant and existing chatbot solutions that many companies use.

### **Motivation**

Current solutions of chatbots add extra overhead for customers to take additional time to look up the website of the product or organization and use a browser on a mobile or desktop to take the help of a chatbot. This project aims at evolving the chatbot solution into the next step of integrating with one of the voice assistants using Alexa skills. This will help customers get answers on any product or business just by invoking a voice assistant and start asking questions just like stating a conversation to an agent.

### **Value to the Student**

This topic helps me think from the customer's perspective and work backward from customer needs to develop a feature that would help any business to be able to help customers at their convenience. The Chatbot Alexa skill development would involve getting accustomed to project planning like requirement analysis, implementation details. It would also involve developing and testing an application using a programming language like Python and deploying the application in a serverless environment like Lambda, which would help to get familiarity with the software development lifecycle.

### **Alignment to Program Outcomes**

Development of chatbot and integrating with Alexa skill involves programming, developing application and testing code in Python. I can apply the skills I learned in Data Structures and Software Testing courses.

Once the chatbot application is ready, the application will be deployed to a Serverless Lambda to enable interaction with end-users. I can apply the skills I learned in courses like Introduction to Cloud Computing and Managing the Technology Project.

## **2. BACKGROUND**

Any customer-facing corporation or business needs to provide customer support to address the questions and problems of its customers. (Reddy Karri & Santhosh, 2020). As the scale of the company grows, it becomes more and more difficult to scale up the customer support team, and they turn into automated solutions like a chatbot. A chatbot can be programmed to serve many of the customers' needs, like answering the

most common questions about the business and providing support and pointers with the problems a customer can face.

Chatbots are dominantly being used by many websites to help customers answer the most commonly asked questions in an interactive way tailored to the questions a customer has. Current solutions of chatbot add extra overhead for customers to take additional time to look up the website of the product or organization and use a browser on a mobile or desktop to take the help of a chatbot. With many hands-free devices in the market and homes, customers are increasingly relying on voice assistance to get information about a product or topic that they are interested in (Singamaneni, Srinivas & P.S.L., 2021). This project explores the possibility of integrating the chatbot with the most convenient voice assistant so that the customer can get answers about our product using a hands-free voice assistant

This project aims at evolving the chatbot solution into the next step of integrating with one voice assistant using Alexa skills. This will help customers get answers on any product or business just by invoking voice assistant and start asking questions just like stating a conversation to an agent

## **3. RELATED WORK**

As virtual assistants, Chatbots are used to simulate human conversations with experts providing knowledge for users. A chatbot can be viewed as a question-answer system. A chatbot is a software developed to simulate a conversation with a human partner. The purpose of this survey paper is to provide an overview of a few existing approaches to implementing a chatbot system that responds to the Alexa skills and commands

### **Literature Review**

Human-machine communication through chatbots is accomplished by using machines like computers and mobile phones. Over the past couple of years, considerable progress has been made in the field of designing chatbots. (Mondal, 2018). Research on chatbots tends to be broadly distributed in the following areas: techniques for retrieving and generating the content, the length of a conversation, and the domain itself (e.g., open or closed).

Artificial intelligence branches such as voice recognition and voice assistants are highly

promising (Dickson, 2018). Due to the narrow scope of voice recognition, you can create a model that will recognize and translate voice commands in all circumstances, regardless of the background noises and accents. There have been several incidents in recent years that highlight significant security and privacy risks associated with intelligent virtual assistants (IVAs) (Chung, 2017). Better diagnostic testing of IVA ecosystems can reveal security and privacy vulnerabilities, making them more reliable.

### **Review on Serverless Application**

Serverless allows you to run code without provisioning or managing servers, creating workload-based scalability rules, managing event integrations, or managing runtimes. According to (Yan, 2016), the commoditization of natural language services has led millions of consumers worldwide to turn to chatbots as the latest platforms for communicating with companies, which provide developers with many building blocks to build chatbots inexpensively. Serverless technology is used by my project to implement Alexa skills to run chatbots.

### **Review on NLP**

In artificial intelligence, Natural language processing (NLP) is a branch that deals with understanding and processing the language that humans use. The ultimate objective of NLP is to decode, decipher, and understand human language in a worthwhile manner.

Unstructured text input cannot be understood by the system directly since it understands only structured formats. The user input received by the system is in an unstructured text format. With the use of NLU techniques, keywords and patterns are extracted from the unstructured text to create a structured format (Kłosowski, 2018). An NLU system needs to comprehend each word of a sentence to comprehend the complete sentence. Therefore, the first task is dividing a sentence into individual words. Then, the system must comprehend the grammar of each word. NLTK is a module in python which is ready to perform Natural Language Processing. (Karri, 2020)

### **Review on Chatbot**

According to the research (Hristidis, 2018), the author provided an overview of the technologies underlying chatbots, including Information Extraction and Deep Learning. The authors also examined the differences between conversational and transactional chatbots. There have been two different kinds of bots described:

manual chat logs based on free text and physical chatbots that are designed to achieve a specific objective, such as booking a flight. They also discussed commercial platforms and tools that can create and deploy chatbots.

We will build a chatbot that will process and understand the user's questions so that they can ask college-related questions without physically attending the college. The chatbot will be based on artificial algorithms that analyze the user's queries and understand their message. As a user, you can ask any question you want related to the college you are at. For example, you ask about the annual day, fees, faculty information, sports day, etc. It informs you about the events as they unfold in the college. Consequently, the Chatbot usually remembers the previous command to provide the functionality. Thus, no matter what the user asks, it will keep a record of it. Therefore, it helps the Chatbot and can analyze it for future use.

### **Review on AWS Managed Services (AMS)**

Amazon Web Services (AWS) offers 200 fully-featured services and data centers around the world, which makes it the most comprehensive and widely used cloud platform. It provides highly reliable, scalable, and cost-effective cloud infrastructure services(Bozic,Tazl, 2019).

With Amazon Lex, you can build conversational interfaces for applications using text and voice. With Amazon Lex, you'll receive the deep learning capabilities of natural language understanding (NLU), automatic speech recognition (ASR), and automatic speech translation (ASR). One of the applications created with LEX is Amazon Alexa. With LEX, chatbots for IoT can be created. A chatbot is a computer program that simulates human conversation by means of voice commands or text messages (Lee, 2020).

AWS Chatbot is an interactive agent for monitoring and interacting with AWS resources in Slack channels, Amazon Chime chat rooms. With AWS Chatbot you can receive alerts, execute commands to obtain diagnostic information, trigger AWS Lambda functions, create AWS support tickets. Amazon Lex and AWS Lambda make building a chatbot easy with real-time notifications, faster response, quick setup, and easy permission definitions (Sadavarte & Bodanese, 2019). AWS Chatbot is compatible with real-time notifications, smarter responses, and faster response times.

### 3. APPROACH

The business requirement is to automate the entire dataset of frequently asked questions (FAQs) related to Programs and services offered by the university. The scope of this project is not just to automate the questions and answers of the business by building a chatbot but also to integrate the chatbot with an intelligent voice assistant like Alexa so that users can interact with their smart devices at their convenience. The requirement of implementing this project is the access to and maintenance of AWS services for the development and deployment of the chatbot application.

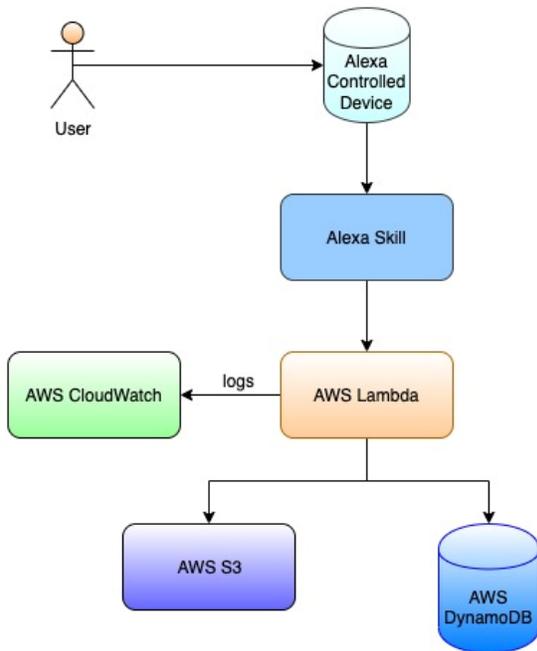


Fig1. Chatbot System Design

According to the system design shown in Fig1, a user interacts with our chatbot application with an Alexa Controlled device like an echo or Fire TV. The chatbot application can be launched by the user uttering the skill invocation name. After successfully launching the skill, the user can ask any question related to the business or product, and the voice assistant Alexa will be able to respond with an appropriate answer. Internally when the user asks a question, The question is processed by various AWS cloud services. The user's question will be processed by the advanced deep learning functionalities of automatic speech recognition (ASR) for converting speech to text

and natural language understanding (NLU) to recognize the intent of the user. Once the user intent is identified, Alexa skill will build a request with recognized intent, slots and send it to a Lambda function. The Lambda will be able to break down the request in the backend to get the values of Intents, slots and uses conditional statements to determine the appropriate response. The response will be built and returned in a JSON format, which will be routed back to the user's device through Alexa skill. The user's device can understand the JSON response and answer the question using the voice assistant. The complete interactions, request-response, and errors are logged in CloudWatch which can be analyzed to improve the scope and performance of the application

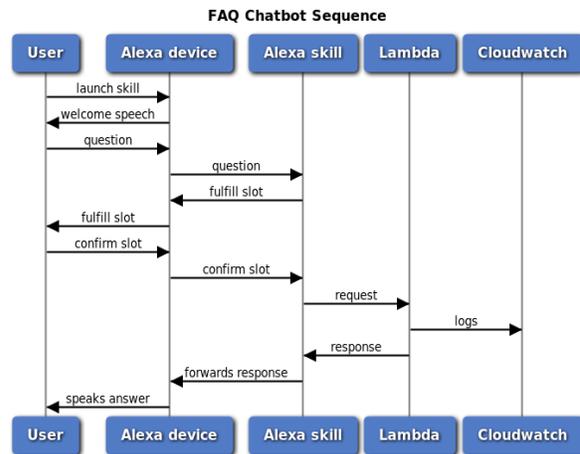


Fig2. FAQ Chatbot Sequence

### 4. DATA COLLECTION

The System Design of university chatbot consists of multiple technologies, which includes AWS, NLP, Python, Serverless, and Amazon Skill Kit

An AWS login is required for the configuration of an AWS Developer account and an AWS Management console Alexa skill. The chatbot will be implemented using the Dialogflow, which will be utilized as the NLU to perform artificial intelligence methods such as Natural Language Processing (NLP), POS tagging, and entity recognition to analyze text and carry out the appropriate actions. Dialogflow is a Natural Language Understanding Engine (NLU) used for the extraction of entities and intent from a user's message.

An Alexa skill is created by configuring and publishing it with the ASK Developer Console.

Alternatively, the AWS Management Console allows the programmer to leverage more than one AWS service, monitor cloud services, manage users, and control execution rights. The interaction logic for a custom dynamic skill is defined as part of the Alexa skill configuration. Custom skills can be developed by developers using the ASK custom interactions model. Using the Alexa Skills Kit (ASK) developer console, you can build the front-end of a custom Alexa skill. Several components must be configured to implement the interaction model.

As shown in Fig2, the Chatbot sequence follows Invocation: When users initiate an interaction with the Alexa skill, they use a specific keyword or phrase. Intent: There are different types of intents, which is to say, a structure for a particular set of functionalities where steps for voice interaction, user requests, corresponding actions are described. Intents can have optional arguments like slots.

Sample Utterance: ASK searches for possible matches of input requests in the lists of sample utterances to identify the intended intent. Sample utterances are probabilistic phrases indicative of users' propensity to express a particular purpose or react to specific questions. Dialog Model: It provides a structure to capture the mandatory data required for a particular intent by defining the steps involved in the conversation between Alexa and the user. The developer can also define and configure the intents, dialog delegations, slots, and utterances in the "JSON editor" in JSON format. Amazon Lambda: It imports all the resources necessary to configure ASK and NodeJS for the Lambda function, on top of a custom inbuilt template in the AWS Serverless App Repository.

Alexa skills can be tested on mobile devices via the Amazon Alexa app; it is available for Fire OS, Android, and iOS-based devices. Skills can be tested even before they are published, provided that the same account has been used to access the app.

**Data Source Selection:**

An authentic and reliable data source is needed to extract the backend data which is used for the chatbot design. Considering these features, the City University of Seattle website was selected as a dataset source. CityU is an educational university in the United States. The website

offers information content about topics such as scholarships, fees, guides, Student Services, Courses offered, recommendations, and suggestions about relevant miscellaneous topics. A valid AWS Lambda ARN endpoint is required for the skill to function. A developer must build and deploy the skill before testing it.

**5. EVALUATION**

A traditional chatbot requires a data collection of expected prompts from a user and the corresponding responses to the prompts that servers as the answers for any question asked by the chatbot user. This set of prompts that needed to be seeded or trained the model of chatbot can be categorized at Intents. Each intent can be trained with a series of questions, which the user of the chatbot can ask. The bot can understand any type of question that a user asks and interprets to one of the seeded questions using the Natural Language Processing and deep learning backed by the Alexa skills kit. Each Intent can be further classified into several optional sub-categories called slots which can drive a natural dialogue-based conversation so that the virtual assistant can further refine the granularity of the response to provide.

Fig3. Data Classification 1

Sample Response
The university is open from Monday to Friday, 9 am to 5 pm Pacific standard time
we are located at 521 Wall Street in Seattle washington
we are located at 789 West Pender Street in Vancouver British Columbia
We offer Master in Business Administration, Master in Program Management and Master in Computer Science
We offer Certificate in English Language Proficiency, Certificate in Marketing and Certificate in Full-stack development
We offer Bachelors in Criminal Justice and Bachelors in Information Systems
Please visit our website <a href="http://www.cityu.edu/programs-overview">www.cityu.edu/programs-overview</a> for more details related to programs we offer
Yes, City University of Seattle offers Cornish Commons, a 20-story residence hall building.
You can visit <a href="https://www.cityu.edu/housing/">https://www.cityu.edu/housing/</a> for further details

Fig4. Data Classification 2

For the voice assistant chatbot that is being developed, data has been collected from sources like the City University FAQ page and other detail pages like academics, scholarships, and admissions on the university website. As shown in Fig3, each data point collected has been classified into several intents. For example, any

details related to the programs offered by the university are classified as Programs Intent. As there are numerous programs offered by the university, this chatbot scope is defined to cover a specific set of programs to keep the responses concise and to have the opportunity to expand the program details in future work. With this defined scope, any question in the context of this intent is always followed up by a voice assistant to provide the program level that the user is interested to learn. This sub-classification of programs is defined by the feature of slots, and a slot named 'Levels' is defined. The user can respond with one of the defined slots like master's or bachelor's or Associate level. If the user does not know the level of programs available, a follow-up feature can help with the options available for the program level. Once the voice assistant chatbot has information related to the intent and the corresponding slot, it can provide the exact response defined for that slot.

As shown in Fig4, with the data that is available on the website of the university, the intents identified for the purpose of the university FAQ chatbot are Timings, Address, Programs, Housing, Tuition, Scholarship, Admissions, Commencement, Academic Calendar. Some of the intents like Housing, Timings do not have any optional slot, and the voice assistant can respond with the available response once it interprets that the request falls in one of these intents. For other intents like Address, a follow-up slot can be a location, and the user will respond with a location he is interested in for the voice assistant to respond with the exact address of that location.

## **6. FINDINGS**

The focus of this project is to take one of the features that have been the core for any business in recent times, a chatbot, into the next phase of evolution. As discussed in earlier sections about the recent advancements made in intelligent voice assistants and the advent of several smart devices, any emerging or existing business has an opportunity to expand its outreach by developing a smart skill that can answer any question related to business instead of the user visiting the FAQ page or the need to physically type questions using a chatbot on their website. With this use case and the requirements to serve as an FAQ chatbot, an Alexa skill has been developed with a limited scope for the City University, which can answer some of the topics related to admissions, location, scholarships, and housing. The data gathered for this project has been classified into sections called Intents and sub-sections called

Slots and based on user questions, the Natural Language Processing available in Alexa can map the request to one of the Intent. If needed, the voice assistant can request more details from the user to fine-tune the response using the defined slots. Once the Intent and Slot are identified as related to a request, then a python Lambda function is triggered, which would process the request based on conditional statements and return a response for the user.

With the configuration of all Intents and slots and defining the programming logic in Lambda, a functioning voice assistant smart skill has been delivered, which could very well replace the text-based chatbots with voice-based, self-evolving, and intelligent chatbots in a short period. This would also eliminate the need to visit the website for a customer who is interested to learn about a business and instead simply interact with the hands-free voice assistant to launch the skill and get as many details as possible at their convenience. Another major advantage to using a voice assistant skill, as discussed in this project, is how easy and quick a skill can be developed and deployed. AWS provides several managed services that run in the cloud like Lambda for writing and deploying the code, CloudWatch to analyze logs, a dashboard for data analytics, and many other machine learnings backed models like Natural Language Processing and Deep Learning in-built, without the need to provision and develop each feature from scratch. Using managed services, as a business can focus on delivering the products quicker and efficiently like our FAQ skill. The skill developed can be published to be available to all users that use an Alexa-supported device, and the skill can be activated by the user.

## **7. CHALLENGES**

The first major challenge encountered in this project was to export an AWS LEX chatbot to an Alexa skill. A functioning traditional text-based chatbot is developed in AWS LEX and based on the official documentation review. The chatbot can be easily exported to an Alexa skill. However, during the project implementation, the lack of documentation on the process of exporting the chatbot to an Alexa skill and the recent upgrade of AWS LEX console to v2 has posed a major risk to the project as the intents and Lambda are not working as expected after exporting. An alternate approach has been quickly researched and implemented the entire skill within the Alexa skills console without the need of AWS LEX due to the strict time constraints of this project

The second major challenge is to identify the correct intents and slots based on a user request and to troubleshoot several failures I encountered during the implementing phase. The AWS CloudWatch logs were helpful to quickly identify the issues and resolve them promptly, and ultimately deliver a functioning voice skill.

## 8. CONCLUSION

Chatbots have been around for a long time and have been a predominant feature for many growing businesses to provide automated support that answers the most common question or an issue that a customer has. The major limitation of the chatbot has always been it being available to a customer who can visit their website or reach their support page. This Project reviews all the recent advancements in the field of breakthrough hands-free voice assistants, Machine learning technologies like Natural Language Processing, and Serverless technologies like Lambda and CloudWatch in AWS. Intending to take the chatbots to next level, this project delivers a simple working voice assistant skill for Alexa that serves as a FAQ chatbot for the City University of Seattle website. The extra time required to provide your resources and configure your API endpoints are a thing in the past, and automatic provisioning in AWS managed services makes it simple and quick to deliver this functioning conversational chatbot that can be developed for any type of business.

## 9. FUTURE WORK

This project focused on delivering a functioning voice assistant skill that serves as a conversational chatbot for customers' FAQ related to the City University of Seattle. The scope is this project is built around the Time constraints to complete the literature review of emerging technologies related to Machine learning, Serverless Computing, and AWS managed services and deliver the skill that addresses the limitations with the existing chatbots that any business uses. As discussed in previous sections, we addressed the problem statement and built a voice assistant skill that responds to customer requests. This work can now be extended for any type of business and to address any type of used case related to serving a customer. The possibilities of business used cases are endless, like developing a smart skill that can troubleshoot the most common problems related to a hardware or software product of a company. The analytics provided by the Alexa developer console can be useful to analyze the features that might be

missing based on customer intents, and the scope of existing skills can be further enhanced. The responses provided by the skill can be dynamic by retrieving alternate responses for an intent, which can be achieved by using DynamoDB and extended for visual interaction by storing media in AWS S3 storage

## 10. REFERENCES

- 1) Bozic, J., Tazi, O. A., & Wotawa, F. (2019, April 1). Chatbot Testing Using AI Planning. <https://doi.org/10.1109/AITest.2019.00-10>
- 2) Chung, H., Iorga, M., Voas, J., & Lee, S. (2017). "Alexa, Can I Trust You?" *Computer*, 50(9), 100–104. <https://doi.org/10.1109/mc.2017.3571053>
- 3) Dickson, B. (2018, September 3). The fundamental problem with smart speakers and voice-based AI assistants. Retrieved from TechTalks website: <https://bdtechtalks.com/2018/09/03/challenges-of-smart-speakers-ai-assistants/>
- 4) Hristidis, V. (2018). Chatbot Technologies and Challenges. *2018 First International Conference on Artificial Intelligence for Industries (AI4I)*. <https://doi.org/10.1109/ai4i.2018.8665692>
- 5) Kłosowski, P. (2018, September 1). Deep Learning for Natural Language Processing and Language Modelling. <https://doi.org/10.23919/SPA.2018.8563389>
- 6) Mondal, A., Dey, M., Das, D., Nagpal, S., & Garda, K. (2018, November 1). Chatbot: An automated conversation system for the educational domain. <https://doi.org/10.1109/iSAI-NLP.2018.8692927>
- 7) Lee, L.-K., Fung, Y.-C., Pun, Y.-W., Wong, K.-K., Yu, M. T.-Y., & Wu, N.-I. (2020, August 1). Using a Multiplatform Chatbot as an Online Tutor in a University Course. <https://doi.org/10.1109/ISSET49818.2020.00021>
- 8) Paikari, E., & van der Hoek, A. (2018). A framework for understanding chatbots and

their future. Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering - CHASE '18.  
<https://doi.org/10.1145/3195836.3195859>

- 9) Reddy Karri, S. P., & Santhosh Kumar, B. (2020, January 1). Deep Learning Techniques for Implementation of Chatbots. <https://doi.org/10.1109/ICCCI48352.2020.9104143>
- 10) Singamaneni, S. K., Srinivas, L., & P, S. L. (2021, June). [Review of Deep Learning Chatbot using Python]. International Journal of Advanced Trends in Computer Science and Engineering, 10(3), 2019–2023. <https://doi.org/10.30534/ijatcse/2021/751032021>
- 11) Sadavarte, S. S., & Bodanese, E. (2019, December 1). Pregnancy Companion Chatbot Using Alexa and Amazon Web Services. <https://doi.org/10.1109/PuneCon46936.2019.9105762>
- 12) Yan, M., Castro, P., Cheng, P., & Ishakian, V. (2016). Building a Chatbot with Serverless Computing. *Proceedings of the 1st International Workshop on Mashups of Things and APIs*. <https://doi.org/10.1145/3007203.3007217>

## 11.APPENDIX

The link provided below is for the source code in my GitHub repository and the demo video uploaded to youtube for my project on university chatbot

GitHub Link for source code :  
<https://github.com/sreeleekhasangitha/cs687-capstone-chatbot-alexa-skill.git>

Youtube link for  
Demo : <https://youtu.be/ejs4nSJXNk>