ISS ChatBot Application

KOH SOOK BING A0195413E

LIM CHONG SENG HERMANN A0195392U

TAN YAO TAI TEERAPONG A0073460L

Table of Contents

[1](#_heading=h.2xcytpi) EXECUTIVE SUMMARY 3

[2](#_heading=h.1ci93xb) PROBLEM OVERVIEW 4

[2.1](#_heading=h.3whwml4) PROBLEM STATEMENT 4

[2.2](#_heading=h.2bn6wsx) PROJECT OBJECTIVES 4

[2.3](#_heading=h.qsh70q) PRODUCT OVERVIEW 4

[3](#_heading=h.3as4poj) TECHNICAL DISCUSSION 5

[3.1](#_heading=h.1pxezwc) INTRODUCTION 5

[3.2](#_heading=h.49x2ik5) TECHNICAL SOLUTIONS (GENERAL FAQ INTENTS) 5

[3.2.1](#_heading=h.2p2csry) DATA MINING AND CLEAN UP 5

[3.2.2](#_heading=h.147n2zr) DIALOGFLOW AND APP.PY REPRESENTATIONS 6

[3.3](#_heading=h.3o7alnk) TECHNICAL SOLUTIONS (RECOMMENDATION INTENTS) 8

[3.3.1](#_heading=h.23ckvvd) DATA MINING AND CLEAN UP 8

[3.3.2](#_heading=h.ihv636) DIALOGFLOW AND APP.PY REPRESENTATIONS 8

[4](#_heading=h.32hioqz) CONCLUSION & REFERENCES 10

[4.1](#_heading=h.z337ya) IMPROVEMENTS AND FUTURE ENHANCEMENTS 10

# EXECUTIVE SUMMARY

In recent years, due to fields such as artificial intelligence and data having been identified as upcoming strategic technological trends, and in line with Singapore’s vision to be a Smart nation of the future, the government has been pushing to train more professionals in the areas of artificial intelligence, data analytics and software engineering. Many people are looking for courses and training in those fields, which the Institute of System Science provides.

The Institute of System Science (ISS) is an institution that provides graduate and executive education for graduates and working professionals. It provides programmes and courses in Information Technology, Data Analytics, and Leadership in the Infocomm sector. ISS has also been appointed the National Continuing Education & Training (CET) Institute (NCI) for the National Infocomm Competency Framework (NICF) by SkillsFuture Singapore (SSG) and the Infocomm Media Development Authority of Singapore (IMDA).

With the multitude of courses offer by ISS, the visitors of the website might find it a hassle to navigate through all of the information provided on the website. The institute has therefore proposed that a chatbot system be implemented, so as to create a more user-friendly way for them to navigate and search for the information on the website.

Our project group of 3 members has developed a chatbot system where the user will be able to use natural language to interact with the chatbot. To build the chatbot system, we used Google Dialogflow to detect the intent of the queries, as well as Heroku to contain our solution files. For the user interface, we have implemented two alternatives. To interact with the chatbot, the user has the option of using the default web interface of Dialogflow, or through the program Slack interface.

Our team learned a lot in the process of working on this project. We managed to apply concepts such at Intents that we learned in our lectures, as well as technical skills learned from workshops like chatbot building and making use of webhooks in a business application scenario.

# PROBLEM OVERVIEW

## PROBLEM STATEMENT

With the multitude of courses offered by the Institute of System Science, it could be a hassle for users to try and find information on a specific course. Users who are looking for courses related to a specialization may also have trouble finding all the relevant information.

To help solve this problem, the institute has proposed that a chatbot system be implemented, where the users can use natural language to interact with the chatbot, to create a more user-friendly way for them to navigate and search for the information that they are looking for.

## PROJECT OBJECTIVES

The objective of this project is to build a chatbot system for the ISS website. The chatbot should be able to interact with users using natural language and will be able to answer enquiries related to ISS programs, courses, and related information.

# 

# TECHNICAL DISCUSSION

## INTRODUCTION

As the information on the current ISS site is very broad, the team has chosen to develop multiple intents to make the ChatBot as comprehensive as possible. The intents can be split into 2 broad categories as follows

**General FAQ Intents that**

* retrieve Course Details and Descriptions (GetCourseDetails)
* retrieve Course Fees (GetCourseFees)
* retrieve available Fundings/Sponsorships (GetFundingDetails)
* retrieve Teaching Staff details (GetTeachingStaff)

**Recommendation Intents that**

* recommends executive courses by areas of interest, e.g. data science, cyber security. (getCourseTopicintent)
* recommends courses to cater for users looking for courses which provides them professional credits required to maintain their professional certifications ( (getcoursePDUintent)

Both categories of Intents require significant effort in data mining and this will be elaborated in the respective sections below.

## TECHNICAL SOLUTIONS (GENERAL FAQ INTENTS)

### DATA MINING AND CLEAN UP

As the amount of information for this category of intents are extensive, python scripts were developed and used in the Data Mining process to develop both the responses and also tagging the custom entities in DialogFlow. The information required are listed as follows

* List of all Course Names – GetCourseDetails, GetCourseFees, @Course
* List of all Course Descriptions – GetCourseDetails, GetCourseFees,
* List of all Course URLS – GetCourseDetails, GetCourseFees
* List of all available fundings/sponsorships – GetFundingDetails, @Funding
* List of all Teaching Staff (56 Staff) – GetTeachingStaff, @TeachingStaff
* List of all Teaching Staff Biography – GetTeachingStaff
* List of all Teach Staff individual page URLs – GetTeachingStaff
* List of Executive programs with Professionsl credit hours - getCourseTopicintent @course@Topic
* List of Executive programs by topics - getcoursePDUintent @course@PDU

**Course Details and Descriptions (GetCourseDetails)**

Apart from Data Mining, a significant amount of effort was also put in to sieve and clean up the data. This is to ensure that when these data are eventually used to build the entities within DialogFlow, the chatbot can accurately detect the correct entity and respond accordingly.

For example, when data mining the list of courses from the CourseFinder webpages, many of the results actually exam certifications rather than courses such as the Project Management Professional (PMP® Exam Only). The results were eventually removed from the results.

<https://www.iss.nus.edu.sg/executive-education/course/detail/project-management-professional--(pmp-exam-only)/>

Also several of the teaching staff did not have any bio write up on their individual webpages. The writes ups were manually added to include the departments that they work in e.g. Sean Hoang.

<https://www.iss.nus.edu.sg/about-us/staff/detail/208/Sean%20HOANG>

Another significantly time-consuming task was in the cleaning up of the names of the teaching staff so that it can be accurately incorporated into the custom entity. For example “Barry Adrian Shephard”. In the custom enitity @TeachingStaff, the synonyms for “Barry Adrian Shephard” are “Barry” “Adrian” and also “Shephard”. Such that when a user searches for him using his surname “Shephard” or even his middle name “Adrian”. His profile will still show up as a response. This process was done deliberately but not blindly. In the entire list of teaching staff, none of the other staff had names which contained “Adrian” or “Shephard” and hence using these 2 synonyms will not have any conflicting response. However, surnames like “Tan” or “Lee” appeared in many of the teaching staff names and hence they were not included in the synonyms. For example, “Tan Peng Wei”, would be recognized in this entity as “Peng Wei” but not “Tan”. Last but not least, not forgetting a synonym for “Gu Zhan” to be “Sam”.

### DIALOGFLOW AND APP.PY REPRESENTATIONS

**Course Details and Descriptions (GetCourseDetails)**

This particular intent aims to address users question on the details of a particular course. The @Course custom entity which stores the course name of all 108 courses were created. When the user keys in a question and a particular course is identified by this custom entity, the intent GetCourseDetails will be selected which eventually drive the logic in the app.py file to respond accordingly with the Course details. The course name provided by the user is first converted to lowercase before the matching is done with the app.py file.

Sample Questions:

* Could you provide the course description for Master of Technology in Enterprise Business Analytics?
* What is the course PMI Agile Certified Practitioner about?

**Course Fees (GetCourseFees)**

This particular intent aims to address users question on the course fees of a particular course. Similar to the previous intent, the @Course custom entity which stores the course name of all 108 courses is used. The difference between GetCourseFee Intent and GetCourseDetails Intent is that GetCourseFee intent requires a mandatory second entity @Price. This is the differentiating factor between the 2 intents.

As the course fees can vary significantly (e.g. Singapore Citizen vs International, Part-Time vs Full-Time), a generic response of the course fee is provided together with the specific URL of the course webpage that the user has enquired. For example when a user enquires about the course fee for Visions Systems, DialogFlow identifies the GetCourseFee Intent, which eventually drive the logic within app.py to extract the relevant response together with the specific webpage URL for Vision Systems

Sample Questions:

* How much does Master of Technology in Digital Leadership cost? What is the course PMI Agile Certified Practitioner about?
* How much is the total school fees for Strategic Product Manager?

**Available Fundings/Sponsorships (GetFundingDetails)**

GetFundingDetails is a relatively simpler intent as compared to the rest. A custom entity @Funding is used which contains entities like “Sponsorship” and even names of Funding programs such as “SkillsFuture”. DialogFlow identifies this intent and responds with a URL that directs the user to the Funding webpage. Due to time and resource constraint, there is actually potential for this intent to be enhanced further similar to other intents. This is further elaborated in Section 5.1 of this report Improvements and Future Enhancements.

Sample Questions:

* Can I use skillsfuture credit?
* Is there any government grants available?

**Teaching Staff Details (GetTeachingStaff)**

In this particular intent, a custom entity @TeachingStaff is used. As elaborated earlier, a significant amount of effort was used to customize the synonyms for this particular Custom Entity such when users search for “Shephard” or “Adrian”, DialogFlow will still be able to identify the entity as “Barry Adrian Shephard”.

DialogFlow identifies this intent and subsequently passes the intent into the logic of app.py which then returns the details and specific URL webpage of the particular Teaching Staff.

Sample Questions:

* How much does Master of Technology in Digital Leadership cost? What is the course PMI Agile Certified Practitioner about?
* How much is the total school fees for Strategic Product Manager?

## TECHNICAL SOLUTIONS (RECOMMENDATION INTENTS)

### DATA MINING AND CLEAN UP

**Executive programs and the corresponding Professional Development Units (PDU) and Areas**

To tranform the data that provide the courses, topics and Professional development units in a usable format, we downloaded the PDF file (nus-iss-executive-education-planner-2019) from the ISS website. We then downloaded an internet tool to convert the PDF to an excel.

A time consuming task for the was to assign the headers which provides the area (e.g. data science, cyber security) that the course is categorised to the 100 over courses within excel.

We also used our domain knowldege understanding of the different words and lookup the internet resource to come up with a list of synonyms that could be used to represent PDU.

### DIALOGFLOW AND APP.PY REPRESENTATIONS

**List of courses related to an area of interest (getCourseTopicintent)**

This particular intent aims to provides the list of executive courses provided by ISS according to the area of interest. @ Topic custom entity which stores the area of interest all the 100 over executive courses in ISS. @ExecCourse provides the list of courses that meet the criteria for display.

ISS provides courses which are available to working professionals who seek to improve their professional knowledge. When the user keys in a question and the area of interest identified by this custom entity, the intent GetCourseDetails will be selected which eventually drive the logic in the app.py file to respond with the list of Executive courses accordingly with the area of interest.

Sample Questions:

* Cyber security related courses?
* recommend datascience courses
* What datascience courses are available?

**List of courses which qualify for PDU hours(getcoursePDUintent)**

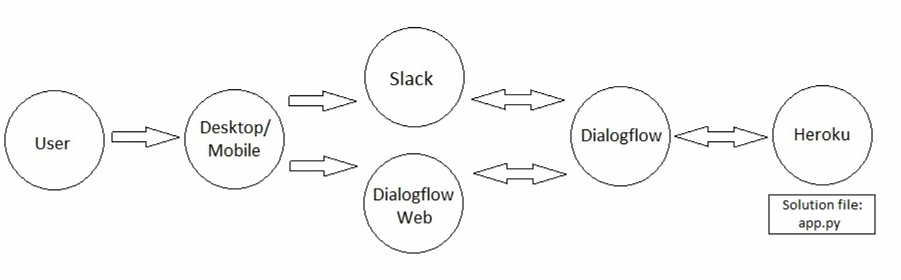
This particular intent aims to provides the list of executive courses provided by ISS which provides PDU required by working professionals for continuous professional development. @ PDU custom entity which stores the area of interest for all the 100 over executive courses in ISS. ExecCourse provides the list of courses and the PDU hours that meet the criteria for display.

Similar to the getCourseTopicintent**,** when the user keys in a question and PDU intent is triggered by this custom entity, the intent getcoursePDUintent will be selected which eventually drive the logic in the app.py file to respond with the list of Executive courses accordingly with the area of interest. Due to time and resource constraint, there is actually potential for this intent to be enhanced to cater for a variations in topics. This is further elaborated in Section 5.1 of this report Improvements and Future Enhancements.

Sample Questions:

* Which courses with Professional Development Unit?
* PDU courses?

## TECHNICAL SOLUTIONS (SYSTEM ARCHITECTURE)



The ChatBot application consist of several parts. A slack or Dialogflow web interface, the Dialogflow engine and finally Heroku. The user can choose to interface with the chatbot either through Slack Interface or DialogFlow Interface. When the user keys in a question, the query will be passed through to Dialogflow for intent detection. With the detected intent from DialogFlow, this is passed on to the Heroku Server through the app.py file where the logic within the app.py will run to determine the appropriate response. The response is then passed back to the user interface.

# CONCLUSION & REFERENCES

## IMPROVEMENTS AND FUTURE ENHANCEMENTS

Due to time and resource constraints, the chatbot and many of the intents do have opportunities for further improvements and enhancements. Below are a list of potential enhancements.

**Available Fundings/Sponsorships (GetFundingDetails)**

This current chatbot only redirects the user to the webpage that shows the list of fundings available. There is potential to further expand on this by data mining the individual URL/funding details/funding requirements of each type of funding and also redirecting the user to the specific individual fund webpage. e.g.

<https://www.iss.nus.edu.sg/executive-education/funding/self-sponsored-individuals/skills-future-funding>

**Teaching Staff Details (GetTeachingStaff)**

This current intent does not incorporate the list of courses taught by each teaching staff. This information could potentially be data mined and included in the response when the user enquires about the courses that this particular staff teaches.

**Potential Addition of Intents**

To make the chatbot all-encompassing, additional intents related to queries on ISS itself could also be added. This intent should also return external links such as links to the youtube promotional video of ISS

<https://www.youtube.com/watch?v=YCJk3PUNRB0>

or even respond with the university rankings of NUS when the user enquires about ISS.

**Courses (getCourseTopicintent)**

The current intent has only included limited area of interest to work on. An area of interest could include many variations (e.g. datascience could be machine learning, etc) We could incorporate including some of the knowledge maps to include related words.