UNIVERSITY OF CAPE TOWN

DEPARTMENT OF COMPUTER SCIENCE

# CS/IT Honours Final Paper 2021

Title: Virtual Student Advisor

Author: Michael Brough

Project Abbreviation: ADVICE

Supervisor(s): Aslam Safla

| **Category** | *Min* | *Max* | Chosen |
|---|---|---|---|
| Requirement Analysis and Design | *0* | *20* | 0 |
| Theoretical Analysis | *0* | *25* | 0 |
| Experiment Design and Execution | *0* | *20* | 10 |
| System Development and Implementation | *0* | *20* | 20 |
| Results, Findings and Conclusions | *10* | *20* | 20 |
| Aim Formulation and Background Work | *10* | *15* | 10 |
| Quality of Paper Writing and Presentation | *10* | | 10 |
| Quality of Deliverables | *10* | | 10 |
| Overall General Project Evaluation (*this section allowed only with motivation letter from supervisor*) | *0* | *10* | 0 |
| **Total marks** | | **80** | |

# ADVICE: Virtual Student Advisor

Michael Brough
The University of Cape Town
Cape Town, South Africa

## ABSTRACT

ADVICE is a prospective Virtual Student Advisor system for the University of Cape Town (UCT). ADVICE is a website for UCT students to get advice on their choice of study and learn more about their curriculum, it also hosts a chat bot that users can interact with and ask questions to in the place of a human student advisor. This paper investigates the best practices for designing chat bots and their potential implementation in the UCT digital ecosystem via a contextual inquiry. The study asks potential users about their experiences with chat bots and what they would like to see from a UCT-made chat bot. This paper discusses the most important features for building a chat bot such as this one. This paper also details the design and implementation of a UCT Student Handbook PDF reader application which was needed to populate the database that is used by the ADVICE website and chat bot.

## 1 INTRODUCTION

Having a Student Advisor to help guide students through their degree plays a critical role in their path to their future [5]. Advising students and answering student queries are some of the most important responsibilities a university must fulfill. The process of completing this task efficiently and effectively becomes extremely complicated when thousands of students need advice on their degree. The University of Cape Town (UCT) is faced with the challenge of enrolling over 25000 students annually, all of whom will have to come into contact with a Student Advisor at some point for various reasons.

The process of providing advice to each and every student is a momentous task that has been made more difficult due to the no-contact learning environment brought on by the COVID-19 pandemic. Successfully resolving queries via email is far more difficult compared to in-person meetings. Some universities in other countries have attempted to solve this problem by creating automated Student Advising systems that students can use for simple queries in place of speaking to a student advisor [8–10].

The aim of ADVICE is to create a website that hosts a chat bot which acts as a virtual student advisor. One advantage that this chat bot has over student advisors is the fact that it operates even outside of regular office hours. The goal for this tool is to offer guidance to students on their courses, electives, credit calculations, and general information about what they can expect from their degree.

This project was split amongst three people with three individual sections: The website and database, the chat bot, and the HCI component of chat bots as well as a tool to read the information from the UCT students handbooks (the documentation they would have to look through without the aid of the bot).

This paper covers the last section mentioned, this entails a contextual inquiry into UCT students' experience with chat bots in general, what they liked, disliked, and what features garner the most interest. The contextual inquiry seeks to gain insight into the average user such that the chat bot can be designed to best suit their needs.

This paper also details the design and implementation process of the Student handbook PDF reader application. This paper describes the technologies used to create the application and how the application achieves the desired outcomes.

## 2 AIM FORMULATION AND BACKGROUND WORK

### 2.1 Student Handbook Reader

For the website and the chat bot to operate, it needs access to a database of UCT Course and degree information. Since this project could not get permission to access UCT's database of Courses and Majors, a new tool needed to be used. The best alternative was to build a tool to upload this information to the website's own database, and the next best resource for accessing this information is the UCT student handbook. Due to this, an application that can read the course and degree information from the student handbooks and upload them directly to the website's database needed to be created.

Through a literature review, several tools and libraries were discovered to help build this tool. It was discovered that Apache PDFBox [2] is an open-source library that allows Java programs to create, extract data from, and manipulate PDF documents. Since the digital Student Handbooks are uploaded to UCT's official website as PDF files, this library was discovered to be most effective in converting the information from the handbook into a text format which a Java program can easily read and manipulate.

The Virtual Student Advisor website utilises a MongoDB database to read and store information. For a Java program to communicate with a MongoDB database, the MongoDB Java Drivers [7] library needs to be used. This library is most commonly added through Maven but the .jar file can be downloaded and manually added to a software project using an IDE.

### 2.2 Chat Bot Design Principles

Nowadays, people are familiar with conversational agents like Apple's Siri, Google's Google Assistant, or Amazon's Alexa. They use inputs such as the user's voice, vision (images), and contextual information to assist the user by answering questions in natural language, making recommendations, and performing actions. [6] Conversational agents are widely successful in the field of technology and have been adapted for many languages and uses across the world.

Historically, online support agents in commerce and marketing would be manned by real people who would give interactive advice and support to customers in a way that a static delivery of information (e.g. frequently asked questions) could not. Advances

in technology, however, have led to the development of unmanned chat bots which accurately mimic human speech patterns and give the user the impression that they are speaking to a real person. The paper written by Go and Sundar [4] addresses the “humanness” of a chat bot and why it is important for them to mimic human conversation. They mention that there are three main aspects of what makes conversational agents human-like: The use of human figures (visual cues), human-associated names or identity (identity cues), and the mimicking of human language (conversational cues). These three aspects influenced the questions asked during the contextual inquiry.

There has been little work surrounding Virtual Student Advisors but there have been studies such as one where a virtual advisor website linked students to Student Advisors over the internet. [3] The advisor was not automated however, it simply displayed 14 frequently asked questions, and if further assistance were required, it would connect them to an advisor when one was available.

The questions in the contextual inquiry aim to discover what aspects of a chat bot are the most important for user satisfaction and what the most valuable features of a virtual student advisor are. The inquiry also seeks to figure out how chat bots are perceived and if potential users would trust them with their personal information or if they would trust recommendations from the bot like they would from a human student advisor. The inquiry also seeks to uncover what features the user-base are most likely to enjoy and if there are any suggestions that could be employed in future ADVICE-related projects.

## 3 SYSTEM DEVELOPMENT AND IMPLEMENTATION

This section covers the details of the development and implementation of the student handbook PDF reader application. This section shows what the application looks like, how it is used, and how it works (code breakdown).

### 3.1 Student Handbook PDF Reader

The application launches a simple GUI (graphical user-interface) containing a toolbar with the "File", "Edit", and "Help" options. (See Figure 1 below)


**Figure 1: An image of the PDF Reader GUI**

Upon selecting "Open" from the "File" tab, the user is prompted with a file chooser (See Figure 2). From here, the user can select the Student Handbook pdf file stored locally on their machine. Once they have selected the file and clicked on "Open", the program processes the text in the PDF file and asks the user if they would like to upload the data to the database. If the user selects "Yes", the program uploads the data to the ADVICE website’s database.

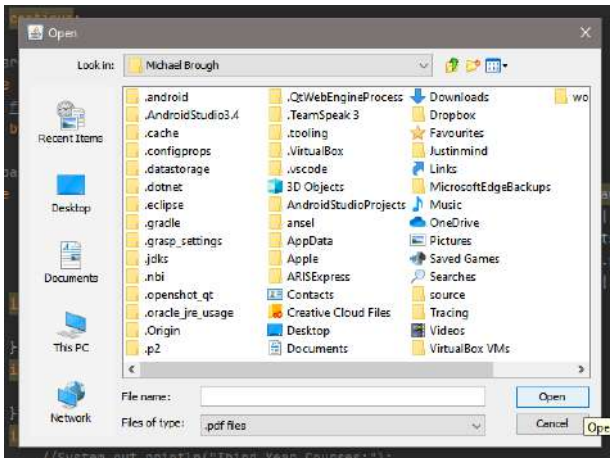



**Figure 2: An image of the PDF Reader File Selection**

### 3.2 Technology Used

This application was built using the Java programming language with the use of the following open-source libraries:

- Apache PDFBox [2]
- Apache FontBox
- Apache Commons [1]
- MongoDB Java Driver [7]

Apache PDFBox, FontBox, and Commons were used to access and subsequently process the text in the Student Handbooks given as input while the MongDB Java Driver library was used to communicate with the web server and upload the processed information to the MongoDB database.

### 3.3 Classes and Code Structure

The application consists of 6 Java classes:

- Main.java
- GUI.java
- SettingsGUI.java
- Program.java
- Major.java
- Course.java

A full overview of the classes and methods can be found in the class diagram (See Appendix A).

Program.java, Major.java, and Course.java are custom objects designed to store the information about each Program (e.g. SB001), Major (e.g. Computer Science, Applied Statistics), and Course (e.g. CSC1015F, MAM1000W) that the program finds in the student handbook. These classes each have their own accessor and mutator methods which are used by the GUI class.

Main.java is the class that launches the GUI and GUI.java is responsible for managing user inputs and the processing of the PDF

files. GUI.java is the largest class of the application and contains methods for reading the PDF document, searching for specific instances of strings, and manipulating the data to store them in an array of custom objects (e.g. Programs, Majors, Courses). These arrays are then used to add the data to the database for the use of the chat bot and website.

### 3.4 Methods used to Extract Data

This section details the low-level process of how the program accesses the student handbook, extracts the relevant information, and uploads it to the mongoDB database.

Once a file has been selected by the user, the program uses a pdf text stripper (belongs to PDFBox library) to check the first page of the document to determine what handbook it is. Depending on the version (i.e. Science or Commerce handbook), the program looks for a specific key phrase in the PDF's text to start the search for information about Majors in the faculty.

The program creates a text file and stores the PDF document's text in it. The program then iterates through the text file one line at a time, looking for specific keywords and symbols in order to save the relevant information about each major (i.e. name, code, first year courses, second year courses, etc) and stores every major in an ArrayList of Major objects.

The programs then looks for the start of the courses section using another keyword (also depends on handbook faculty) where the information about each course is located. The program then iterates line-by-line through the text document and saves the Course Code, course name, credits, NQF level, course entry requirements, and course outline each to a Course object in an ArrayList of Course objects.

Once both ArrayLists have been populated, the text file is then deleted and the information uploaded to the database. The program prompts the user with the option to upload the extracted data to the database. (This Yes/No option exists to prevent accidentally uploading items and for testing and demonstration purposes) If the user selects "Yes", the program iterates through the Course ArrayList and adds the entries to the database and then starts the processing the majors list. For each course that belongs to a major, the program queries the database using the course code as reference and fetches the ObjectId field. The program then uses the ObjectId as a reference to the course when inserting each Major into the database. This is done so that the website can look up the information about each course from the Majors collection using the ObjectId as reference.

### 3.5 Testing

The testing of the handbook reader involved repeatedly giving the Faculty of Science and Commerce handbooks to the application and checking the output for any artefacts or errors. Multiple editions (i.e. 2021, 2020, 2019, etc) of the student handbook were tested with the application. The testing was done until the number of errors or artifacts in the application's output were acceptable.

The program was also tested on files that were not UCT Student handbooks, the application recognises when it is reading a student handbook or not and will indicate to the user to submit a valid file as input.

## 4 EXPERIMENT DESIGN AND EXECUTION

In order to gain an understanding of the possible user base of the website and chat bot, a contextual inquiry was conducted using Google Forms as a platform to record responses to questions about chat bots and the feelings towards the potential implementation of a UCT chat bot.

### 4.1 Contextual Inquiry

To gather data for the contextual inquiry, a survey was conducted using Google Forms. The beginning of the survey included the Informed consent form which the participants would have to agree to before they could proceed to the questions.

The informed consent form and survey were made possible with the help of the UCT Ethics Committee who offered advice on the form and ultimately gave ethics clearance to conduct the survey. The form consisted of 23 short questions where one could choose between multiple options.

With the assistance from the project supervisor, Aslam Safla, a link to the survey was shared across multiple Vula tabs at UCT to gather responses. The link was also shared with close friends and peers who are either studying at UCT or have recently graduated from the institution. All respondents were familiar with the UCT electronic ecosystem and registration process, this was important as the survey contained questions regarding the potential implementation of a chat bot for UCT and how it would be perceived.

### 4.2 Results

This section outlines each question asked in the survey and the results gathered from a total of twenty (n=20) responses.

The first question simply asked if, after reading the informed consent form, they agreed to participate in the research. All of the respondents (n=20) agreed to take part.

The second question asked the participant which of the options bets described them. (See Figure 3 Below)

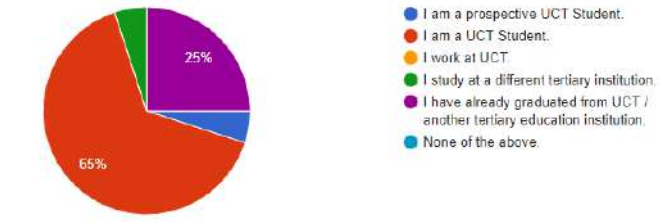



**Figure 3: A Pie Chart showing the distribution of responses for question 2**

Although not all participants were UCT students, they were familiar with the UCT environment as they had already graduated or moved to another institution.

The next question asked the participant to specify their gender. This was to potentially draw a relationship between gender and opinions on chat bots (e.g. if gender played a role on if the participants wanted the chat bot to be friendlier or wanted it to have a name, etc.)

Figure 4 below shows the percentage of male to female and if they preferred not to share. Only (n=1) participant elected to not share their gender information.

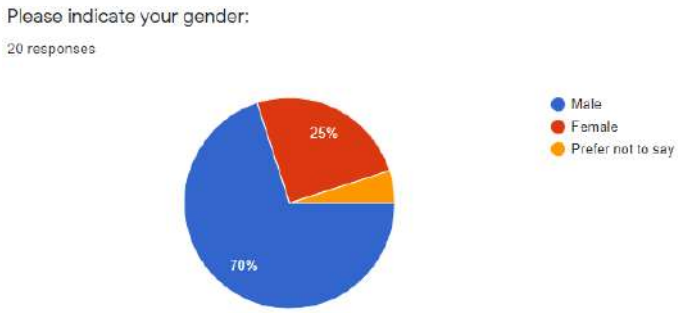



**Figure 4: A Pie Chart showing the distribution of responses for question 3**

Question 4 presented the participant with an image of a simple chat bot and asked if they were familiar with the concept of chat bots in general. Only one (n=1) participant was unfamiliar with the concept. (See Figure 5 below).

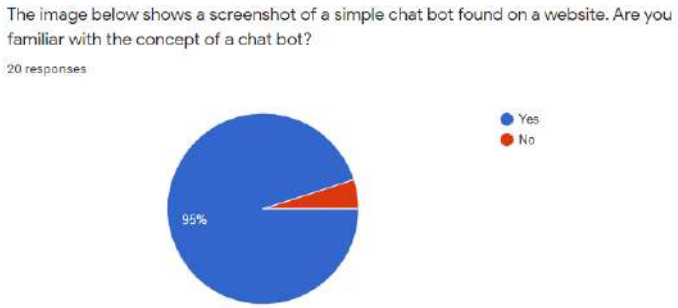



**Figure 5: A Pie Chart showing the distribution of responses for question 4**

The form then asks questions about the participant's experience with chat bots in the past. The question asks, "Have you ever chatted with a chat bot, if yes, how did it help you?" The participants could choose between 4 options, the responses to this question can be found in Figure 6 below.

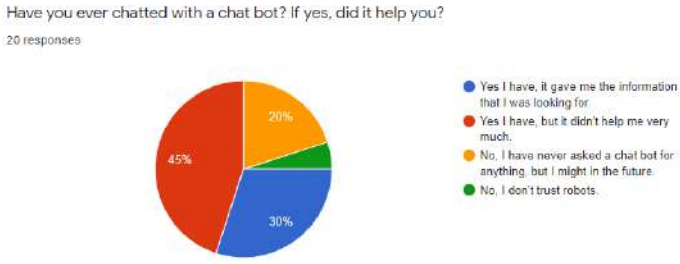



**Figure 6: A Pie Chart showing the distribution of responses for question 5**

75% of participants (n=15) indicated that they have chatted with a chat bot, but only 45% (n=9) of participants found benefit in using it. 20% (n=4) of participants elected that they have not chatted to a bot, but they might do so in the future. One participant (n=1) indicated that they would not chat to a chat bot because they didn't trust the technology.

The next question asked the participant if they would like a chat bot to pretend to be a human or if it should tell the user beforehand. (See Figure 7 below). 90% of participants (n=18) indicated that they would like to know if they are talking to a bot whereas 10% (n=2) of participants preferred if the bot pretended to be a human talking to them.

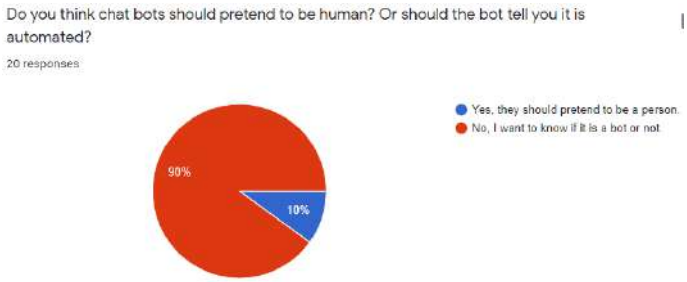



**Figure 7: A Pie Chart showing the distribution of responses for question 6**

The next question posed to participants was whether or not a chat bot should have a name. The responses for this question were divisive with a near 50% split between Yes and No. (See Figure 8 Below)

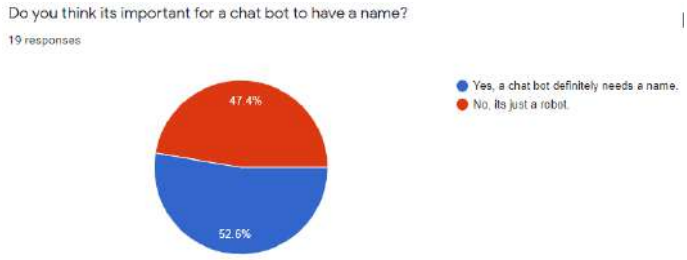



**Figure 8: A Pie Chart showing the distribution of responses for question 7**

The subsequent questions asked the participants what the most positive and negative aspects of a chat bot were. (See Figures 9 and 10). The options they could select from were as follows:

**Positive Aspects**

- The chat bot answered quickly.
- The chat bot successfully helped me.
- The chat bot was friendly.
- The chat bot helped me outside of office hours.
- The chat bot transferred me to a human who then helped me.

- None of the above
- Other: (Custom user input)

**Negative Aspects**

- The chat bot couldn't help me because I needed to speak to a real person.
- The chat bot did not solve my issue.
- The chat bot could not understand me.
- The chat bot took too long to help me.
- I dislike talking to a bot.
- The chat bot was not friendly enough
- None of the above
- Other: (Custom user input)

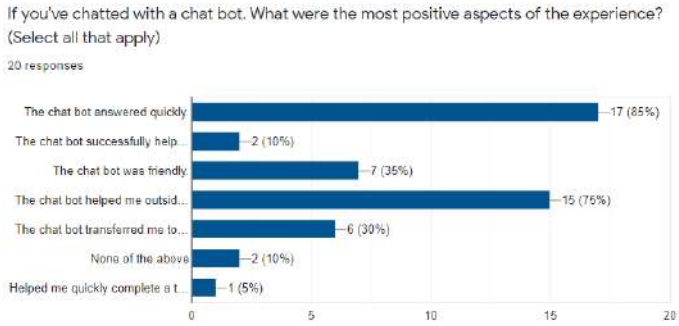



**Figure 9: A Bar Graph showing the distribution of responses for question 8**

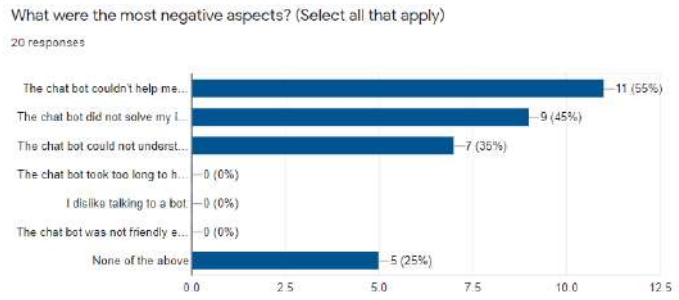



**Figure 10: A Bar Graph showing the distribution of responses for question 9**

Two of the most positive aspects chosen by participants were, "The chat bot answered quickly", (n=17) and "The chat bot helped me outside office hours". (n=15) Two other notable, but less popular choices were, "The chat bot was friendly" (n=7) and "The chat bot transferred me to a human who then helped me" (n=6).

The three most negative aspects of using a chat bot were, "The chat bot couldn't help me because I needed to speak to real person" (n=11), "The chat bot did not solve my issue" (n=9), and "The chat bot could not understand me" (n=7).

No participants selected options 4, 5, and 6, and (n=5) participants selected "None of the above".

Participants were then asked to select the most important aspects that a chat bot should get right. Figure 11 below shows that 80% of participants elected for the feature of connecting them to a real person if they wanted to. The second and third most important aspects were that the chat bot needed to be friendly and that it should always let the user know that it is automated before they interact with it.

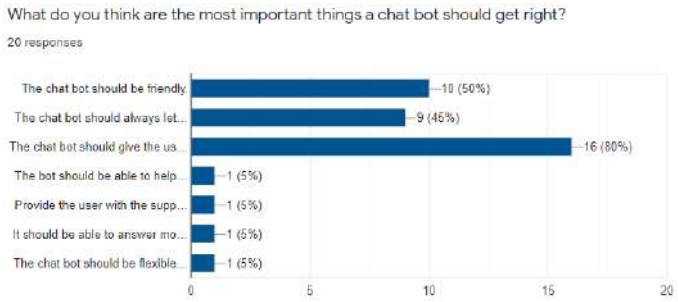



**Figure 11: A Bar Graph showing the distribution of responses for question 10**

The next section of the survey focused on UCT and the possible implementation of a chat bot for UCT students.

The first question in this section was a Likert scale where participants were asked to rate the UCT registration process on a scale of 1-5 (1 = Incredibly Inconvenient, 5 = Very Easy). As can be seen from Figure 10 below, the opinion towards the UCT registration process was low with two modes of a scale rating 1 (Incredibly Inconvenient) and 3 (middle of the scale). Over all entries, the mean rating for the UCT registration process was 2.45.

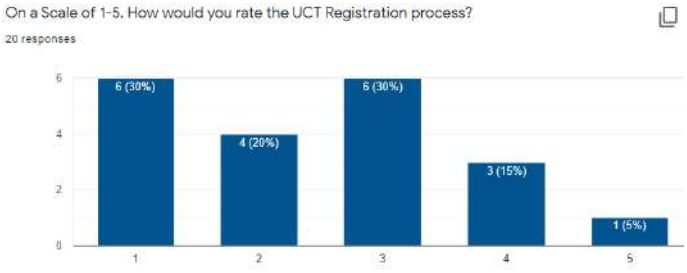



**Figure 12: A Histogram showing the distribution of responses for question 11**

The next question asked participants if they would find benefit in a course credit calculator (See Figure 13 below). 95% of participants indicated that they would find benefit from using a tool like this.

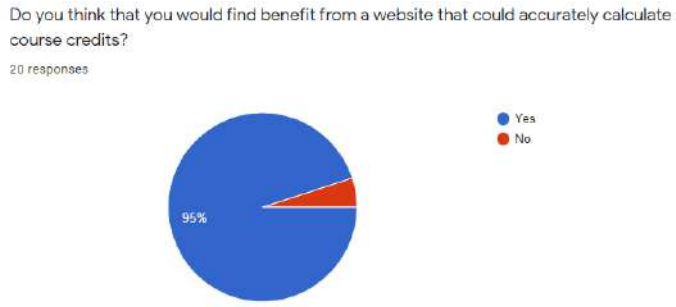



**Figure 13: A Pie Chart showing the distribution of responses for question 12**

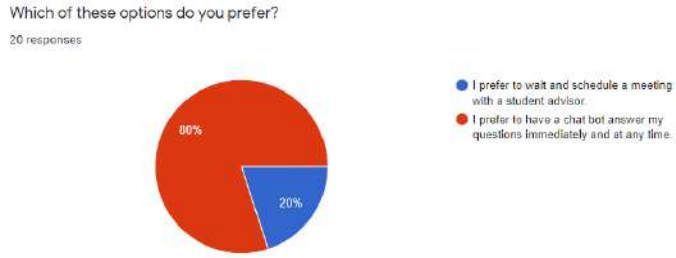



**Figure 14: A Pie Chart showing the distribution of responses for question 13**

Participants were then asked to pick between two options. (See Figure 14 above). It shows that 80% of participants preferred to have their questions answered immediately and at any time over scheduling a meeting with a student advisor.

The next question posed to participants was if they would find benefit from using a UCT chat bot for advice. Figure 15 below shows that 90% of participants (n=18) would find this tool to be beneficial.

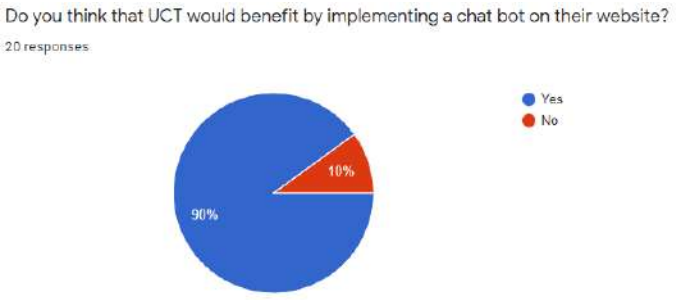



**Figure 15: A Pie Chart showing the distribution of responses for question 14**

Participants were then asked to elaborate on their choice to the previous question in the form of a short text answer. Some notable comments were, "A chat bot that could accurately answer frequently asked questions would be a godsend." and "It would take some of the burden off of lecturers and UCT staff."

They were then asked if they would be comfortable with giving their name, student number, degree, and completed courses to a UCT-made chat bot (The results for which can be found below).

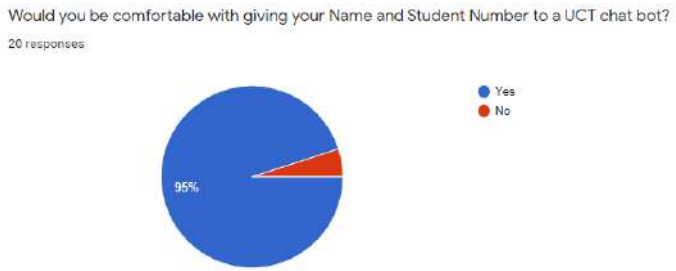



**Figure 16: A Pie Chart showing the distribution of responses for question 15**

Ninety-five percent of Participants indicated that they would be comfortable with sharing their name and student number, but only 75% of participants were comfortable with sharing their degree and completed courses. Twenty percent of participants however, selected "Maybe", instead of "No" on sharing their degree and courses.

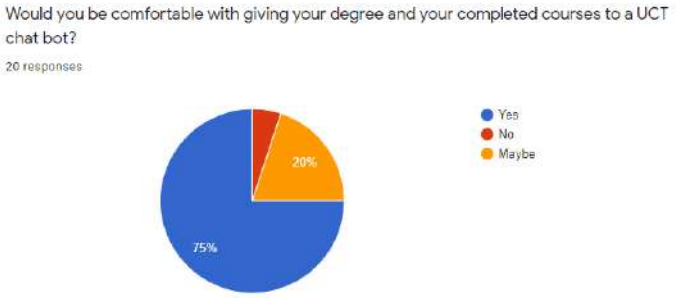



**Figure 17: A Pie Chart showing the distribution of responses for question 16**

The next two questions were, "On a scale of 1-5. How willing are you to talk to a chat bot who could help you instead of a student advisor?", and "On a scale of 1-5. How willing are you to talk to a chat bot first so that you can meet with a student advisor?" (1 = "Absolutely no chance", 5 = "I definitely would").

As can be seen from figures 18 and 19 below, these questions both had a modal class of 5. The first question had a mean of 4 whereas the second had a higher mean of 4.6.

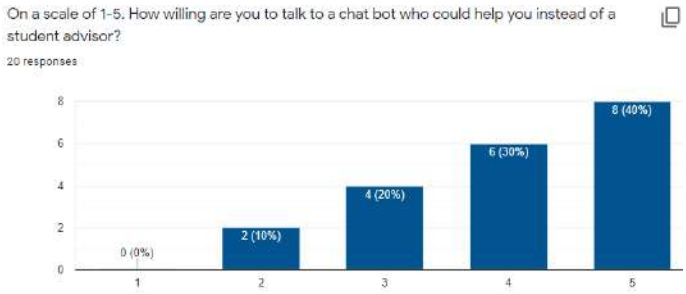



**Figure 18: A Histogram showing the distribution of responses for question 17**

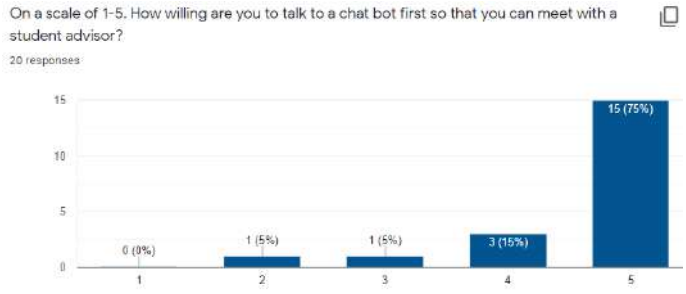



**Figure 19: A Histogram showing the distribution of responses for question 18**

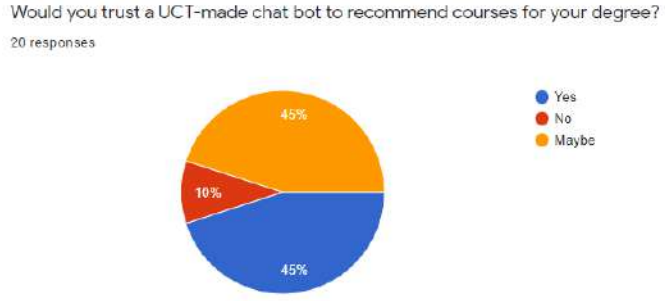



**Figure 20: A Pie Chart showing the distribution of responses for question 19**

The participants were probed further and asked if they would trust a UCT-made chat bot to recommend courses for them to take. Forty-five percent (n=9) of participants answered "Yes" and another 45% (n=9) answered "Maybe". Only 10% (n=2) of participants selected that they would not trust the bot (See figure 20).

Question 20 involved asking participants to select from a list of questions which they would ask a UCT chat bot if they could. The options were:

- When are the UCT Holidays?
- When does registration start?
- What degrees does UCT offer?
- What jobs can I expect from my degree?
- What is the weather like?

All of the options were very popular with no votes for "None of the above" and the next lowest being "What's the weather like" with 5 votes (See figure 21 below).

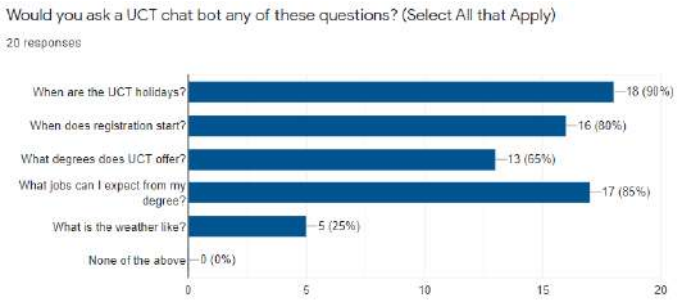



**Figure 21: A Bar Chart showing the distribution of responses for question 20**

Participants were also asked which features they would most like to see in a UCT chat bot. They were given the following list to select from:

- Course registration recommendations
- Answer general questions about UCT
- Course Credit Calculator
- Potential jobs for a certain degree
- Schedule a meeting with a student advisor
- Directing me to the UCT website where I should be able to find what I'm looking for

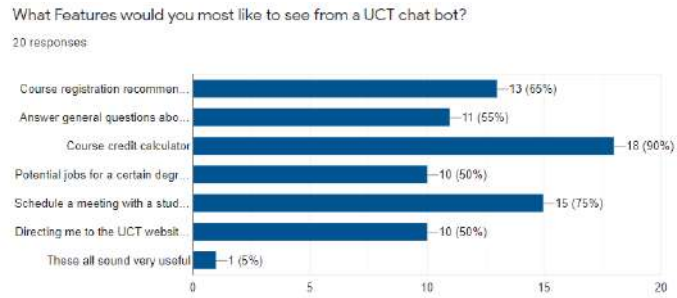



**Figure 22: A Bar Chart showing the distribution of responses for question 21**

All of these options were very popular, the three most popular features were the course credit calculator, scheduling a meeting

with a student advisor, and course recommendations (See figure 22 for full breakdown).

The final question was an optional short text answer which asked the participants if they had any suggestions for features that were not already listed. Some notable suggestions were:

- A feature to get exam/test venues and times for a particular course
- I think that the chat bot should be able to tell you what courses you need to register for as part of your degree.
- The ability to see alternate course timelines, i.e. can I take this course at a later semester and still graduate as expected. Instead of having a very intense first year course load and tapering off can it not be balanced better?

These could be interesting features for the further improvement of the ADVICE website and chat bot.

## 4.3 Discussion

This section discusses the meaning behind and possible outcomes that can be surmised from the analysis of the data in the results section above.

#### *4.3.1 Perception of Chat Bots.*

From the results gathered, it can be seen that the majority of participants had first-hand experience with online chat bots. A vast majority of participants (90%) indicated that a chat bot definitely needs to let the user know that its automated. This feature should be included in the UCT chat bot before it goes live for the public to use. It is important to be transparent with users, a simple, "Hi, I am an automated chat bot, how can I help you?" or something similar will enhance the user experience because the user will know how to react appropriately. This would also help with the chat bot's overall success, if the users knows it is a bot, they will most likely structure their language in such a way that a bot could understand (not including colloquial language, abbreviations, etc).

The participants' opinion on whether or not a chat bot should have a name was nearly evenly split between "Yes" and "No". There was also no correlation between this choice and any other factors (e.g. gender). This could indicate that a chat bot doesn't need a name, but giving it a name would appease the majority of users. It is unlikely that users who selected "No" would be upset that it did have name, so it seems that a name is held with high regard and a name should potentially be given to the UCT chat bot before it is open to the public.

According to the data in figure 9, it appears that the best aspect of a chat bot is that it answers very quickly compared to talking to a person online. Eighty-five percent (n=17) of participants selected this answer and 75% (n=15) of participants selected "The chat bot helped me outside of office hours". This means that the two most important aspects of the designing and hosting of a chat bot are efficiency and accessibility. The chat bot should be accessible to users at any time of the day and it should be quick to reply to users.

The most negative aspects of a chat bot (as can be seen in figure 10) are because of the chat bot not being able to solve nuanced or complex issues. Fifty-five percent (n=11) of participants indicated that in the past, chat bots could not help them because they needed to speak to a real person and 45% (n=9) indicated that the bot could not solve the issue at all, this is perhaps closely related to the third most popular choice: the chat bot not understanding the user.

When asked what the most important aspects of a chat bot were, the most selected option was the ability to schedule a meeting with a real person. (See figure 11) A UCT-made chat bot should definitely include this feature to maximise the user experience. This may be difficult to implement on a large scale, but perhaps the bot could include the ability to send an email on your behalf to an email address that student advisors can check and follow up on.

The next most important aspect of a chat bot is that it should be friendly, therefore a UCT chat bot should be designed in such a way that it comes across as polite and friendly, this will maximise user engagement and improve the overall user experience.

#### *4.3.2 Chat Bot Implementation at UCT.*

The general opinion on UCT's registration process is quite low with a mean rating of = 2.45 (see figure 12), this is mainly due to waiting in long lines to meet student advisors, paging through hundreds of pages in the UCT handbook, and administrative errors. Perhaps this website could remedy that by being a tool that allows them to understand the UCT registration system and how their degrees are structured.

The manual calculation of course credits seems to be a tedious process so much so that 95% (n=19) of participants indicated that they would find benefit from using a course calculator tool. This feature is a must-have for when a UCT chat bot goes live.

A majority of participants indicated that they would rather have their issues solved at any time rather than scheduling a meeting with a student advisor. This indicates that students would be willing to use a UCT chat bot to ask questions. Ninety-five percent (n=19) of participants indicated that they would be comfortable with giving their name and student number to a UCT chat bot if it meant that it could assist them.

Participants are slightly more concerned for their privacy when asked if they would share their degree and the courses they have completed at UCT, although only 75% of participants were willing to share, 20% selected "Maybe" instead of "Yes". This could mean that users would only want to share this information if it were necessary. If the chat bot works well, looks reputable, and can guarantee assistance, users would probably be more likely to share this information with the bot.

Figures 18 and 19 display that users are willing to use the chat bot for advice. It is interesting to note that although participants are willing to communicate with a chat bot to solve their issues, they are much more willing to talk to a chat bot so that they could meet a student advisor. There is some skepticism towards a chat bot being able to solve issues in the place of a student advisor, this is a natural response as there is no tool like it yet. If the chat bot is able to solve simple queries and void the need to talk to a student advisor, the public perception would sway more in the favor of talking to the chat bot instead of a student advisor.

There is also skepticism towards a chat bot recommending courses for users to register for, only 45% (n=9) of participants said that they

would trust the chat bot for this task while another 45% chose the "Maybe" option. If users could see and interact with a functioning chat bot that has access to UCT major and course information, the public perception would likely shift towards trusting the chat bot.

Users also indicated that they would like the chat bot to answer general questions about UCT; such as when the holidays are, when registration starts, what degrees UCT has to offer, and what jobs can be expected from a particular degree. When implementing the chat bot, one should take care to include answers for these general questions too.

## 5 CHALLENGES AND LIMITATIONS

This project was not without its own set of challenges and limitations. The initial plan for the PDF reader was for it to be able to read all student handbooks from the Science, Commerce, and Humanities faculties, but due to time and scope constraints, only the data from the Science and Commerce faculty handbooks can be successfully extracted.

Since there was no incentive provided for responses, it was difficult to convince a large number of people to answer the questionnaire. In order to obtain results that best represent the potential user base of the website and chat bot, a much larger number of responses is required. The results of this study only reflect the views of a small group of people and may not accurately reflect the opinions of the majority of potential users (prospective and current UCT students).

As suggested by one participant, another potential feature to consider adding to the chat bot is the functionality to ask for test/exam venues for particular courses. This would only be achievable if the chat bot had access to UCT's database(s) and could query it. If this was possible, it would also be able to query course and degree information from multiple different faculties and provide more functionality to more students across UCT.

## 6 FUTURE WORK AND IMPROVEMENTS

Work on this project may continue in the future, possible avenues for improvement would be extending the functionality of the pdf reader and conducting a more thorough contextual inquiry.

The PDF reader could be improved to an extent such that it can successfully extract all major and course information from the handbooks of all the faculties at UCT. This way, the chat bot could recommend courses and give information about degrees for all UCT students. Realistically however, it would be a better option to access UCT's already-existing course and degree database instead of scanning the handbooks if ADVICE is being implemented into the UCT digital ecosystem.

Another contextual inquiry could be conducted in the future, perhaps it could include in-person interviews where participants use a prototype of the chat bot and answer questions about the usability and features they like or dislike. The information gathered from this contextual inquiry could then be used to further influence the design of the chat bot and also serves as valuable user-testing and software training for the final product.

## 7 CONCLUSIONS

### 7.1 Student Handbook Reader

The PDF Reader can successfully read the format of the Faculty of Science and Faculty of Commerce UCT student handbooks and subsequently upload the extracted data to the mongoDB database. Working with the students handbooks has highlighted the inconsistencies between them and some issues regarding course information. Sometimes courses only appear in the handbook under the majors and no additional information can be found for them in the courses section. The handbook is not a complete source of information as not all of the courses that you can take appear in one handbook. If this website is to be pursued further it may be better to use UCT's database (or at least a copy of part of it) this is because the student handbooks from each faculty are structured slightly differently from each other and taking all of these small inconsistencies into account is a large scope of work.

### 7.2 Chat Bots Features and Perceptions

The contextual inquiry highlighted some key aspects about chat bots and potential features that could be considered in the future. The participants have voiced that chat bots should be transparent with the user and indicate that they are an automated chat service. The two most positive aspects of a chat bot are its speed and availability. When maintaining the ADVICE website, one must make sure that the bot is replying quickly and is responsive no matter what the time or network traffic. Participants identified their least liked chat bots experiences were when the bot could not help them, it is important to train the bot such that it can respond to user input and provide them with alternative sources of information in the case that it cannot help.

Although participants indicated that they would speak to a chat bot instead of a student advisor, they would still like the option to get into contact with a human student advisor. The usability of the ADVICE website would increase dramatically if this option was made available, but focus should be placed on the chat bot helping the user instead.

Participants had an overall low opinion on the UCT registration system, so one of the goals of this platform should be to help students along the process and make it easier for them to understand the UCT system and how the degrees are structured.

The majority of participants indicated that they would find benefit from using a platform like ADVICE and many others have expressed interest in the idea, they were slightly distrusting of giving their degree information out to a bot but this opinion may change when users can see and interact with a responsive and good-looking interface.

## REFERENCES


[1] Apache. Apache commons. https://commons.apache.org/, 2021.

[2] Apache. Apache pdfbox | a java pdf library. https://pdfbox.apache.org/, 2021.

[3] Coates, W. C., Ankel, F., Birnbaum, A., Kosiak, D., Broderick, K. B., Thomas, S., Leschke, R., and Collings, J. The virtual advisor program: linking students to mentors via the world wide web. *Academic Emergency Medicine 11*, 3 (2004), 253–255.

[4] Go, E., and Sundar, S. S. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior 97* (2019), 304–316.

[5] Gurantz, O., Pender, M., Mabel, Z., Larson, C., and Bettinger, E. Virtual advising for high-achieving high school students. *Economics of Education Review 75* (2020), 101974.

[6] Hauswald, J., Laurenzano, M. A., Zhang, Y., Li, C., Rovinski, A., Khurana, A., Dreslinski, R. G., Mudge, T., Petrucci, V., Tang, L., et al. Sirius: An open end-to-end voice and vision personal assistant and its implications for future warehouse scale computers. In *Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems* (2015), pp. 223–238.

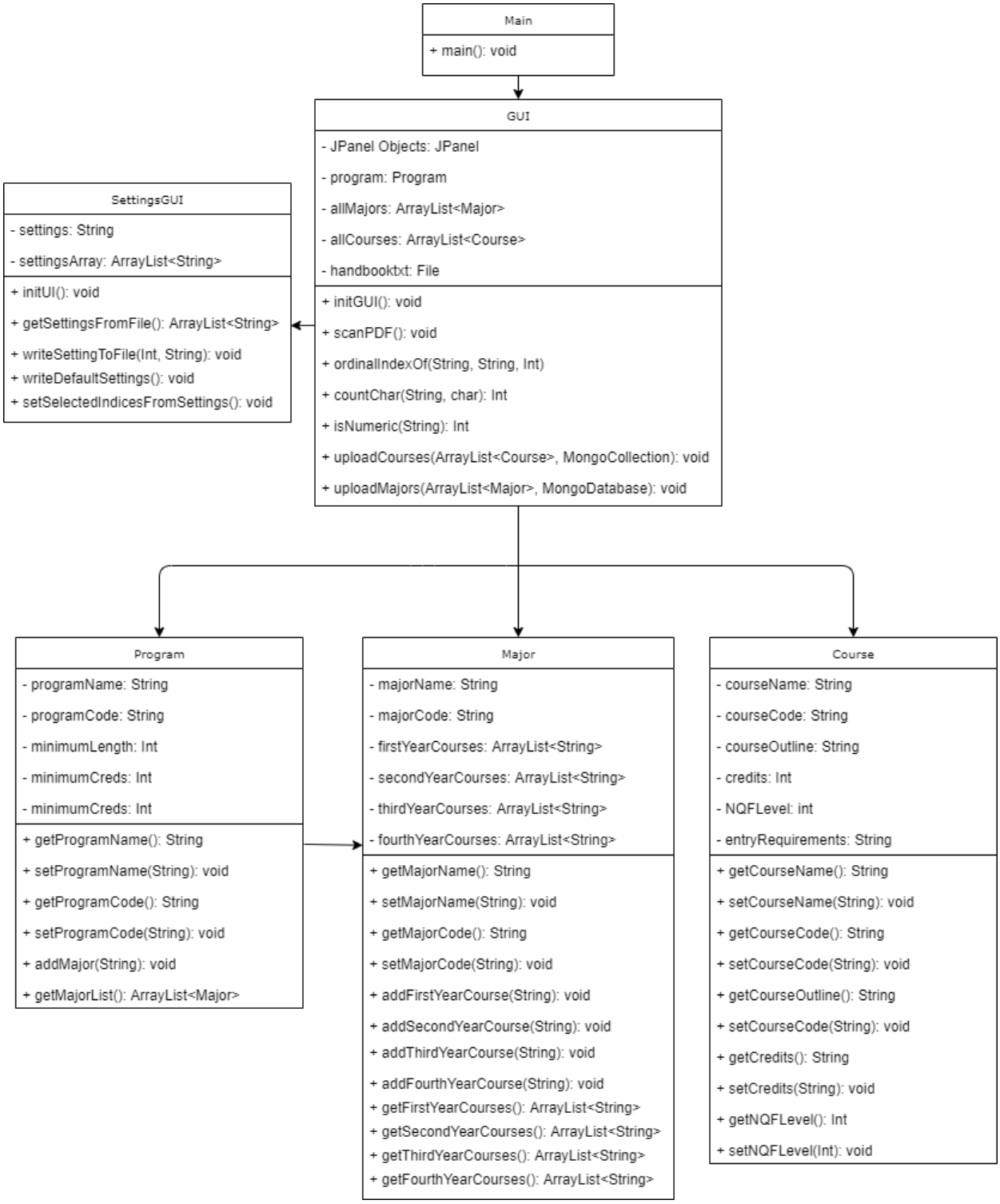
[7] MongoDB. The next generation java drivers for mongodb. https://mongodb.github.io/mongo-java-driver/, 2021.

[8] Ranoliya, B. R., Raghuwanshi, N., and Singh, S. Chatbot for university related faqs. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (2017), pp. 1525–1530.

[9] Suvethan, N., Huzaim, K. A. M., Mathusagar, R., Gamage, M., and Imbulpitiya, A. Virtual student advisor using nlp and automatic appointment scheduler and feedback analyser.

[10] Wijaya, H. D., Gunawan, W., Avrizal, R., and Arif, S. M. Designing chatbot for college information management. *IJISCS (International Journal of Information System and Computer Science) 4*, 1 (2020), 8–13.

## APPENDIX A




**UML Class Diagram for the Student Handbook PDF Reader**