An Innovative VUI system for IT Support at Middlesex University Mauritius.

Moosafur M. Hishaam
Middlesex University
MM3313@live.mdx.ac.uk

ABSTRACT
The IT Support Department of Middlesex University Mauritius has a role to act as helpdesk for all the staffs of the campus while also having to cater for students' issues such as access to the University website (Unihub). As a result, the IT support team is often overloaded with work. Consequently, in order to lighten their workload, a Voice-User Interface could possibly be used to assist struggling students. However, limited investigation has been done towards the use this innovative approach to help students. Due to this limitation, the usability of VUIs has not been well studied though it is a crucial aspect so as to improve adoption by end users. This study delves into the usability of a VUI system for assisting students having difficulties to install a software or accessing their University accounts. Using Html, CSS and JavaScript, a prototype was developed. Evaluation of the prototype was done using the System Usability scale method.

Keywords
Voice assistant; Speech Recognition; Web-Speech-Api, System Usability Scale.

INTRODUCTION
The IT Support department of Middlesex University Mauritius is one of the most important department of the campus, its absence impacts direct on the smooth running of other departments such as the academic, administrative, financial department amongst others. This is because the IT support team has as task to act as helpdesk for all the staffs of the campus while also catering for students having technical difficulties with their access to University website or installing university given software. In order to address the above issues and revolutionise the whole department, Voice user interface (VUI) could potentially be incorporated in the department. Voice recognition is the ability a system has to receive and interpret spoken commands. By definition, voice recognition refers to "the technology by which sounds are transformed into coding patterns" [1]. In recent years, voice recognition has gained prominence especially thanks to the booming of intelligent assistants such as Apple's Siri, Microsoft's Cortana or even Amazon's Alexa. According to a Statista's report around 3.25 billion digital voice assistants were used around the world in 2019, and the number is even forecasted to reach 8 billion users by 2023 [2]. However, there is limited study related to the application of VUI in the helpdesk sector. This study delves into the usability of a VUI system for assisting students having difficulties to install a software or accessing their University accounts. This paper is expected to be insightful to researchers in the area of VUI.

DESIGN CONCEPT
The objectives behind the designing of such a system is to:

- **Improve the reach** of the IT support department. The VUI system will be available to students round the clock and that too on a 7x7 basis.
- **Save money**. With the Voice user interface, staff will have a lightened workload, hence working less overtimes. This will prove to be cost effective for the University who will have lesser wage to pay.
- **Cater for students with visual impairments**.
- **Increase student interest and satisfaction**. A well-designed speech system can provide a user experience that is engaging and enjoyable [3]. Moreover, it will provide ubiquity since students will be able to access the system anywhere using their Laptops.
- **Increase in efficiency and productivity** of the department. Again, with less work to do, the staffs will be able to give more time to work that they are doing and deliver a better quality of work especially in their helpdesk duties. Figure 2.0 shows how workload affects performance in the Helpdesk department according to The Industry Standard in Helpdesk reporting [4]. Refer to **appendix-C** for illustration.

The services that the system will provide to the user is as follows:

- Giving the user instructions about how to login on Unihub and then giving him the option to open the page directly.
- Giving the user instructions about how to register on Unihub.
- Allowing the user to contact the IT Support team via email by opening the email of the user with the receiver email already save.
- Teaching the user how to install Microsoft 365 by opening a picture with instruction about the installation.

At first, the design considered was an auditory interface whereby communication between the user and the system would be purely through sound. That is, typical speech input from students, speech and nonspeech output from the system. However, the nature of such a system would be ephemeral as is only uses transient message therefore placing an important cognitive demand on users. Consequently, the design concept was settled on a multimodal interface, that is one which uses a combination of speech with other modalities such as a small screen to display the conversation between the user and the system. The addition of such a modality mitigates the whole issue of cognitive demand placed on users [5].

The interface will consist of the title of the page at the very top. Followed by 3 buttons, the help button, the Record button and the mute button respectively. The interface was designed in respect to Nielsen's heuristics and Amazon's Voice Design Best Practices (Legacy) [6]. For example, the 2nd rule of Nielsen's heuristics, that is match between system and the real world is followed as the design speaks the user's language. In addition, there is a close and back button in pop-up windows on the application allowing user control and freedom as per the third heuristic rule. Besides, the 8th heuristic rule is also respected as the design is aesthetic and minimalist, that is there is no irrelevant information on the system. **Figure 3 in appendix-D** is an illustration of the design mock-up.

The help button will open a command list and instructions to guide the user in using the system. The mute button will simply remove sound from the system. The record button for its part, whenever activated, will start listening/recording from the device's microphone to detect the user's input. Below is an illustration of how the user's input is processed and the corresponding system response is returned.

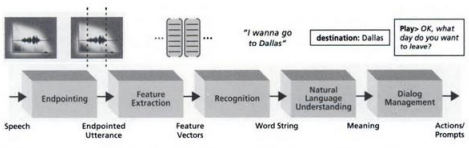


Figure 1- The process of handling spoken input from the user [5].

A typical scenario of a student interacting with the system will be as follows: The student opens the help area, read the instructions and the list of commands available. Following which he will proceed to give commands to the user. For example, the student can ask: "How to install Microsoft 365 on my laptop?". To which the system will reply via audio output: "Here is how you can install Microsoft 365" and proceed to display a document of with the installation guide.
Another typical dialogue could be the student wanting to learn how to login on unihub. Here, the student will say a command such as: "how can I login on Unihub" or "I want to login on Unihub". To which the system will reply: "to login on unihub you will need to have your student Id and your Unihub password, do you want me to open the page for your". Student will reply with a simple "yes" or "no" and the system will open the unihub login page in a new tab.

From the scenarios above, we can see that the system VUI will only accept speech utterances as input, however, the output utterances will vary. The output can be speech from the VUI, text that will appear in the text area, pictures/documents that it can display, open links such as the login link, or even open the user's email with the desired receiver's email pre-set.

The system caters for both conversational design and one off-tasks, that is it can engage in a dialogue with the user as illustrated in the conversational flow diagram (refer to **Figure 4 appendix-E**) or do tasks upon one command. Moreover, the VUI can handle alternatives, for example saying "how can I sign up" instead of "how can I register" or even all sorts of greeting messages, be it formal or informal such as "hello", "hi", "hey", "good morning" amongst others. According to Mikko Kuriomo, in agglutinative language the vocabulary for the conversational speech should include alternatives to cover spelling variations due to colloquial pronunciations, in addition to the word compounding and inflections [7].

PROTOTYPE

The above-described design concept was then transformed into a running prototype. Firstly, the front-end of the application was created using html, CSS and Bootstrap (Refer to a**ppendix-F** for screenshots of the interface**)**. The backend was built on JavaScript. Furthermore, in order to incorporate speech recognition and synthesis into the system, the following 3 methods were considered:

1st-Method	2nd- Method	3rd-Method
Integrate existing voice technologies into the application by using special APIs and other development tools.	Build an intelligent assistant with the help of open-source services and APIs	Create from scratch with further integration into the application.

Table 1-Speech Recognition Incorporation Method

After reviewing, each method, the first method was regarded as best fitted one for this project. Therefore, Web Speech API was implemented, more specifically the web kit speech recognition API. Speech recognition was used both as an input for continuous dictation and control and as scripting to generate text-to speech output. Finally, the web speech API was also used to control activation and timing and to handle results and alternatives.

While developing the prototype, a lot of challenges were faced. The following are some examples:

1. Detecting long sentences.

The system has a hard time detecting long sentences uttered by the user. One of the main reasons behind this problem is that accent of people may vary therefore some words sound different. To cater for this problem, the system was designed to accomplish a task in the shortest possible dialogue flow as recommended in Amazon's best VUI design practices [6].

2. Conveying what the assistant cannot do.

It is crucial to convey to users what the system cannot do [8]. For example, when the user asks "How to install Microsoft 365?", the system displays a list of instructions about the installation. If the user tries to say something such as "download the software". The system will reply "I didn't catch what you said, please try again" because this feature doesn't exist yet. A possible solution would be: Since the above is a *likely utterance,* the system can let the user know by saying "sorry, downloading the software is not yet available, we are still on it".

3. Error correction.

A classic problem about VUI was encountered. When the voice assistant hears wrong, there is no way to correct it. We tried to address this problem by providing different error messages encouraging the user to try again Another possible solution could have been the inclusion of a text area for the user to type in the command all over again.

EVALUATION

In order to gather crucial information related to user experience, it is imperative to conduct an evaluation of the proposed prototype. To achieve this, the System Usability Scale (SUS) was used [9]. The latter is the mostly widely standardized usability assessment questionnaire [10]. SUS is simple and based on a ten-item scale providing a general perspective of subjective usability tests. In addition, SUS is reliable and provides a clear view of user on the system [11]. Furthermore, in some previous studies related to VUI, SUS was effectively applied [12]. **Table 4** in **appendix-G**, shows the 10 items part of SUS, adapted from previous works which involved this scale [12], [13]. The Likert-5 scale where (1 is strongly disagree and 5 is strongly agree) was used to assess each point.

Evaluation was conducted remotely using skype due to covid-19 pandemic situation. A total of 3 third year, computer science students were involved in the exercise, the latter were targeted as participants since the application aims at providing technical support to computer science students. Moreover, for such a study, the minimum number of required participants were met [14]. Besides, as procedures, participants were firstly briefed about the purpose of the research and the application to assess. Following this, the consent form in **appendix-H** was presented to them and informed consent was sought. The participant was then provided with a link to the application and let free to interact with it. During the interaction, the researcher observed the and recorded any difficulties the participant encountered and also the reactions of the participant to those problems. Upon completing the interaction, the SUS questionnaire in **appendix-I** was given to the user to fill-in.

FINDINGS AND RESULTS

Following the evaluation, data related to usability was gathered using the SUS questionnaire. The results were processed and stored in a tabular form as illustrated in **Table 2.** From this result set, **S2** got a very good score as most participants did not agree that the application was unnecessarily complex to use. The students participating found the VUI interaction entertaining as it changes from mainstream text input and output. They even suggested adding

fun response from the VUI to make it even more enjoyable. Participants also strongly disagree that system required learning a lot of things prior to using it as the mean score of **S10** is the second highest one. This is because, a simple click is needed to activate the system and the rest is done by reading the commands out loud. Parallelly, participants found it pleasing and relieving to use an interface with a minimalist design as navigating through the functionalities was quick and easy learn, this explains the high mean scores of **S3** and **S7**. Besides, the score of **S6** is high as the participants did not agree that the system had many inconsistencies. The students unanimously claimed that the VUI was consistent in its way of responding for all functionalities. All of their voice inputs were displayed in the text area for each functionality. Also, the button sizes were consistent along with the font weight and sizes inside the text area. Furthermore, **S4** also obtained a good score. All of the students strongly agreed that the support of a technical person would not be needed as they did not run into any bug or got stuck on any functionality. Additionally, as computer science they know that with such a system, in case of any contingency, just refreshing the browser might solve most bugs such as connection problems. The evaluation results were encouraging, hence showing that such a system could possibly be incorporated in Middlesex Mauritius's IT Support Team to support all Computer Science students ranging from year 1 to final year as the technical proficiency required to use the system is low.

On the other hand, **S1** has a low mean score as participants argued that they will not use the application frequently as the functionalities such as learning how to login, register on unihub and installing Microsoft 365 are tasks they will do only once in a while. Also, low mean scores were obtained from the remaining SUS items, more specifically between 2 and 4. Some issues related to usability may explain these scores. For example, in the case of **S5**, 2 of the 3 participants were neutral about whether the functions of the system were well integrated while the 3rd one strongly disagreed. They claimed that for the 'learning how to install Microsoft' function, the pdf of instructions could have been included instead of a screen of the pdf. Moving on, among the 3 participants, 1 felt confident using the system while the 2 others did not, resulting in the low score of **S9**. Those 2 students argued that the system was having issues to recognize their voice. For, both of them were unfamiliar with VUI and did not have a natural British accent, unlike their 3rd counterpart. They had to repeat the same command a couple of times until the system recognized it. In addition, for the similar reasons, 66.7% of the participants strongly agreed that the system was cumbersome to use, resulting in a poor mean score for **S8**.

ID	SUS Point	Mean Score
S1	I think that I would like to use this system frequently.	2.0
S2	I found the system unnecessarily complex.	1.33
S3	I thought the system was easy to use.	4.33
S4	I think that I would need support of a technical person to be able to use this system.	4.0
S5	I found the various functions in this system were well integrated	2.33
S6	I thought there was too much inconsistency in this system	1.67
S7	I would imagine that most people would learn to use this system very quickly.	4.33
S8	I found the system very cumbersome to use.	3.67
S9	I felt very confident using the system	2.67
S10	I needed to learn a lot of things before I could get going with this system.	1.67

Table 2- SUS Results

The overall SUS score obtained was exactly 70.0 which is deemed as an acceptable system according to the scale [15]. The score suggests a positive feedback of the systems but also room for improvement. To summarise the positive points, participants felt that the system: had a good interface design, is user friendly, is intuitive and easy to use. However, drawbacks of the system were also pointed out as part of this paper. The following are possible improvements that could be implemented:

1. **Integration of pdf files**

It was mentioned during the evaluation that instructions for installation of Microsoft 365 was displayed as a picture. Those instructions can be given as a pdf file, hence more information could be given in a single screen.

2. **Text Display**

On a few occasions, missed whatever the VUI was saying. Therefore, as an improvement, whatever response said by the VUI can be also displayed in text format in the text area.

3. **Instruct users about how to speak to the VUI.**

As for the inability of the system to recognize different accents quickly, a note can be included in the command list notifying users to speak in a slow pace in case the system is not recognizing their voice input.

CONCLUSION

This research delved into the usability of a Voice User Interface system for the IT Support Department of Middlesex University. The system was built using Html, CSS and JavaScript including technologies such as web-speech-api. The system's usability was evaluated using the System Usability Scale method and the results were deemed acceptable as per the standards. As future works, the improvements proposed could be implemented in the system.

References

[1] J. Baumann, "Voice recognition," *Human Interface Technology Laboratory, fall,* 1993.

[2] Y. C. Y. L. S. M. Q. a. Y. M. Han, "Voice-Indistinguishability--Protecting Voiceprint with Differential Privacy under an Untrusted Server," 2020.

[3] J. Larson, "Voice user interface design for novice and experienced users," *Practical Spoken Dialog Systems,* 2004.

[4] M. Olson, "Conversation Performance," *The Industry Standard in Help Desk Reporting,* 2018.

[5] M. C. M. G. J. a. B. J. Cohen, Voice user interface design, Addison-Wesley Professional., 2004.

[6] H. Lee, Voice User Interface Projects: Build voice-enabled applications using Dialogflow for Google Home and Alexa Skills Kit for Amazon Echo, Packt Publishing Ltd, 2018.

[7] P. S. S. V. a. M. K. S. Enarvi, Automatic Speech Recognition With Very Large Conversational Finnish and Estonian Vocabularies, vol. 25, no. 11, pp. 2085-2097 ed., IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017.

[8] C. Liljestam, "Exploring a voice user interface to convey information in an e-commerce website.," 2019.

[9] J. Sauro, "Interpreting single items from the SUS," 2018.

[10] J. Lewis, "The system usability scale: past, present, and future," *International Journal of Human–Computer Interaction,* 2018.

[11] A. G. T. a. B. G. Nathoo, "Exploring the Use of Tangible User Interfaces for Teaching Basic Java Programming Concepts: A Usability Study," in *In 2019 Conference on Next Generation*

Computing Applications, 2019.

[12] D. F. P. Z. S. a. Z. S. Ghosh, "Assessing the utility of the system usability scale for evaluating voice-based user interfaces," in *In Proceedings of the Sixth International Symposium of Chinese CHI*, 2018.

[13] A. a. J. T. Pyae, "Investigating the usability and user experiences of voice user interface: a case of Google home smart speaker.," in *In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 2018.

[14] J. D. J. a. R. J. Dumas, A practical guide to usability testing, Intellect books, 1999.

[15] P. K. a. J. M. A. Bangor, "An empirical evaluation of the system usability scale," *International Journal of Human–Computer Interaction*, 2008.

Appendix

Appendix- A: Link to Webpage: https://hishaam18.github.io/Voice_Assistant/

Appendix- B: Link to Video: https://youtu.be/92RaaquZdJ4

Appendix- C

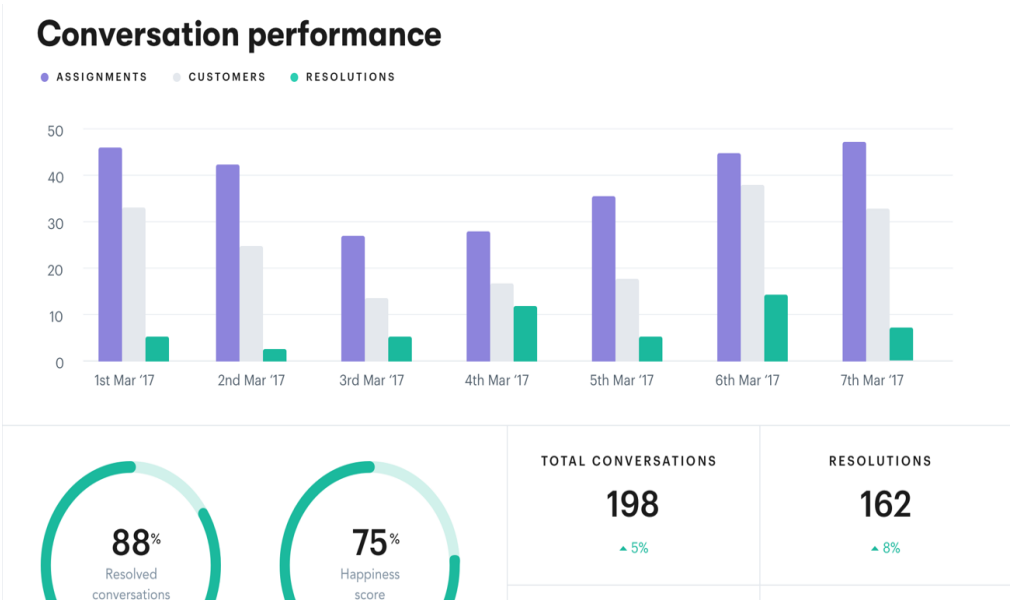


Figure 2-Performance x workload chart - The Industry Standard in Help Desk Reporting

Appendix- D

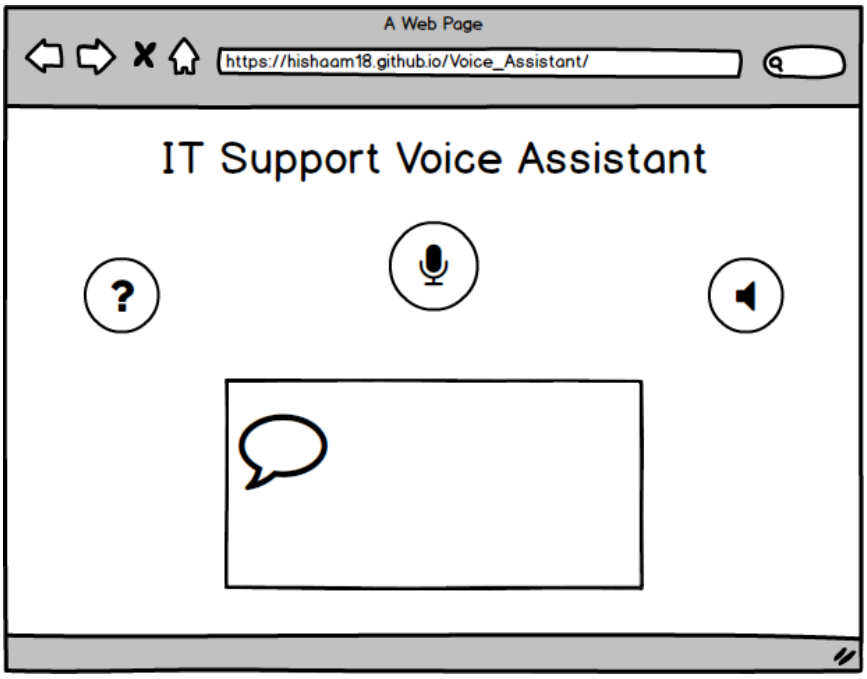


Figure 3-Interface Mock-up

Appendix – E (Conversation Flow)

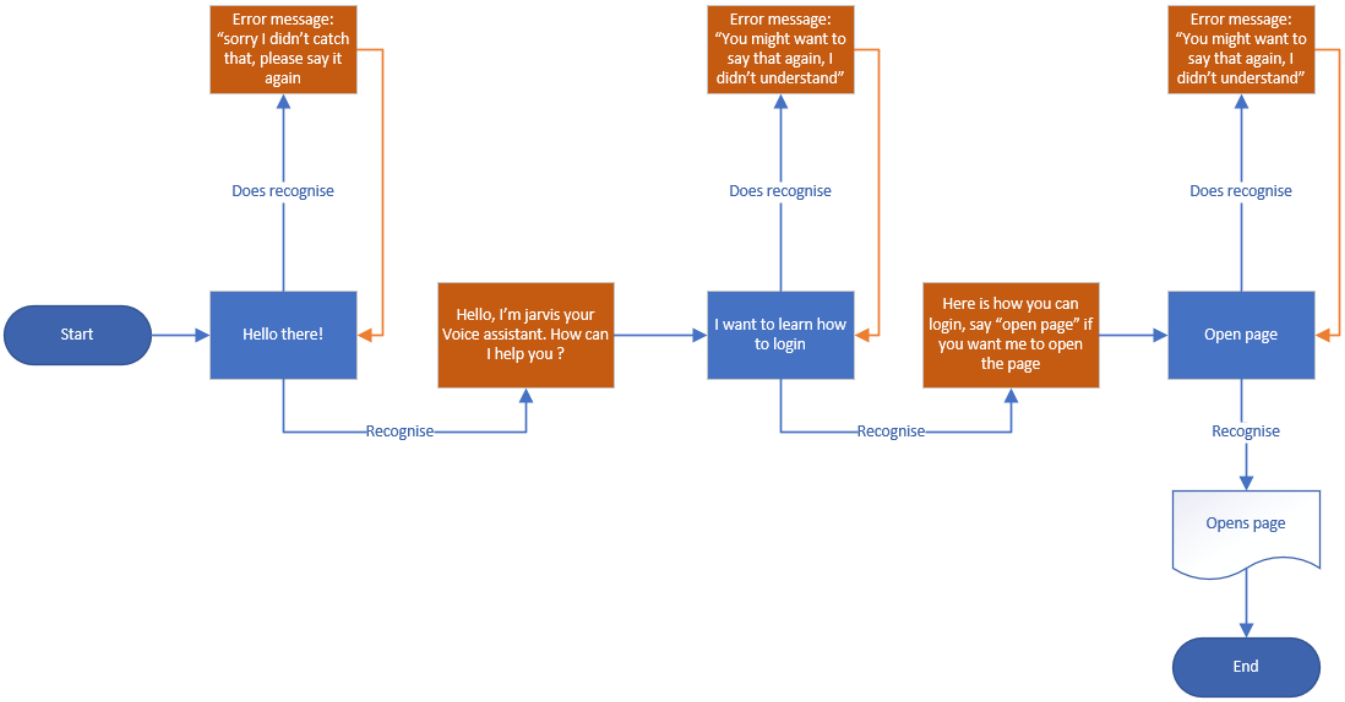
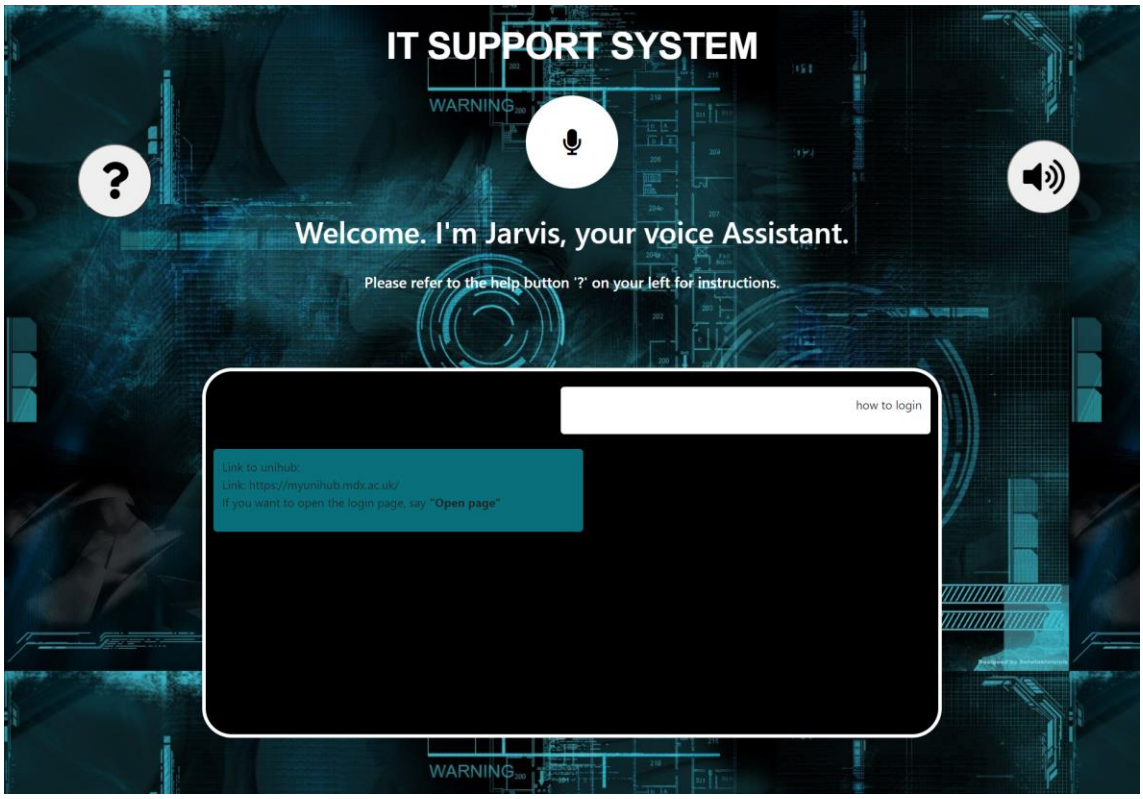
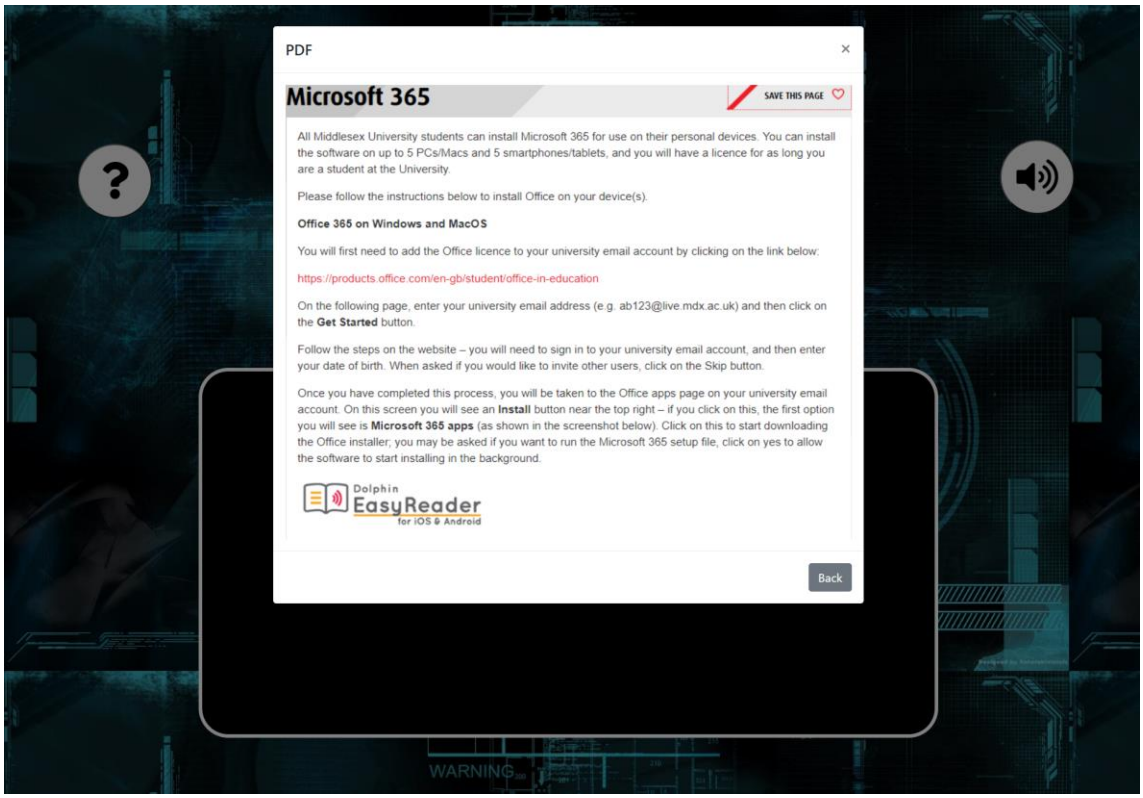


Figure 4- Conversation Flow

Appendix- F (screenshots of the webpage)



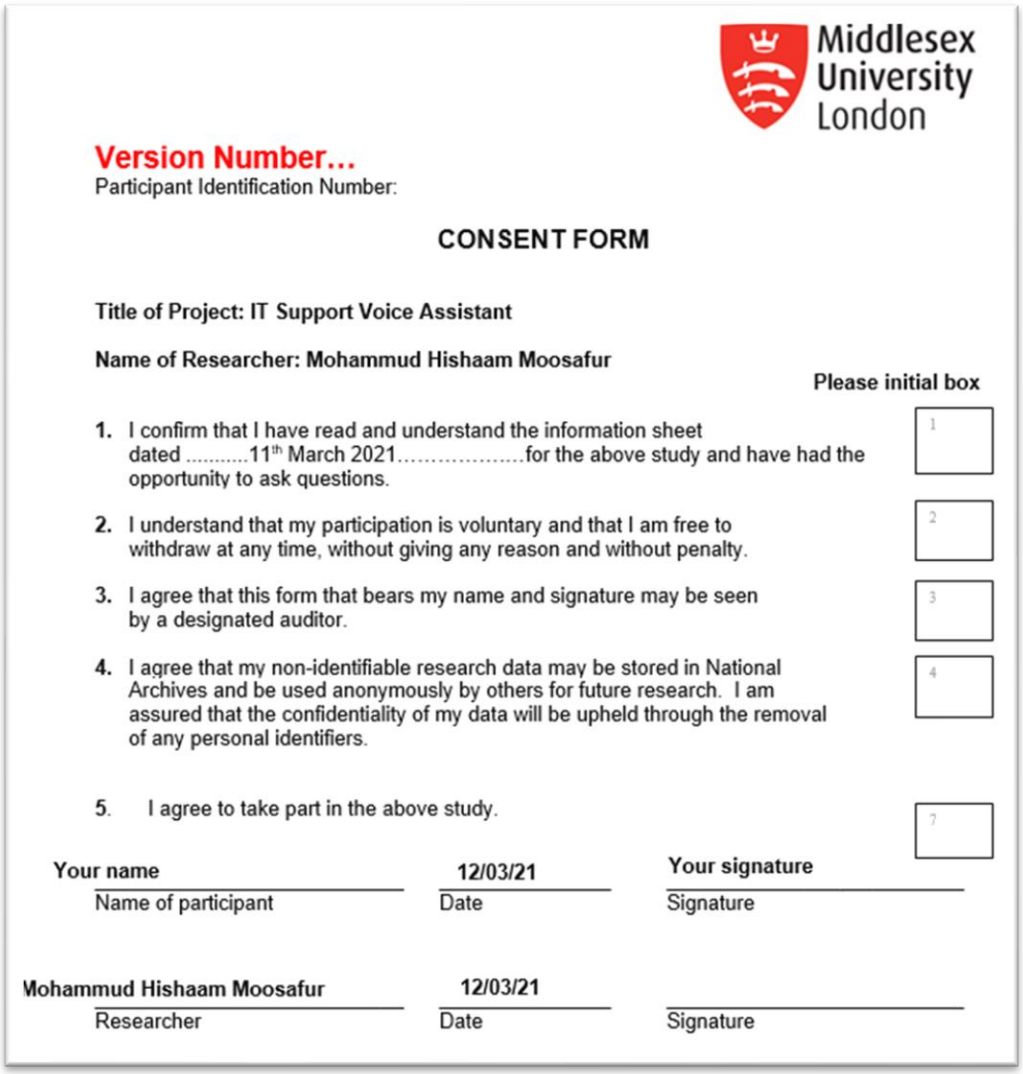


Appendix-G

ID	SUS Point
S1	I think that I would like to use this system frequently.
S2	I found the system unnecessarily complex.
S3	I thought the system was easy to use.
S4	I think that I would need support of a technical person to be able to use this system.
S5	I found the various functions in this system were well integrated
S6	I thought there was too much inconsistency in this system
S7	I would imagine that most people would learn to use this system very quickly.
S8	I found the system very cumbersome to use.
S9	I felt very confident using the system
S10	I needed to learn a lot of things before I could get going with this system.

Table 3- SUS points

Appendix-H



Middlesex University London

Version Number…

Participant Identification Number:

CONSENT FORM

Title of Project: IT Support Voice Assistant

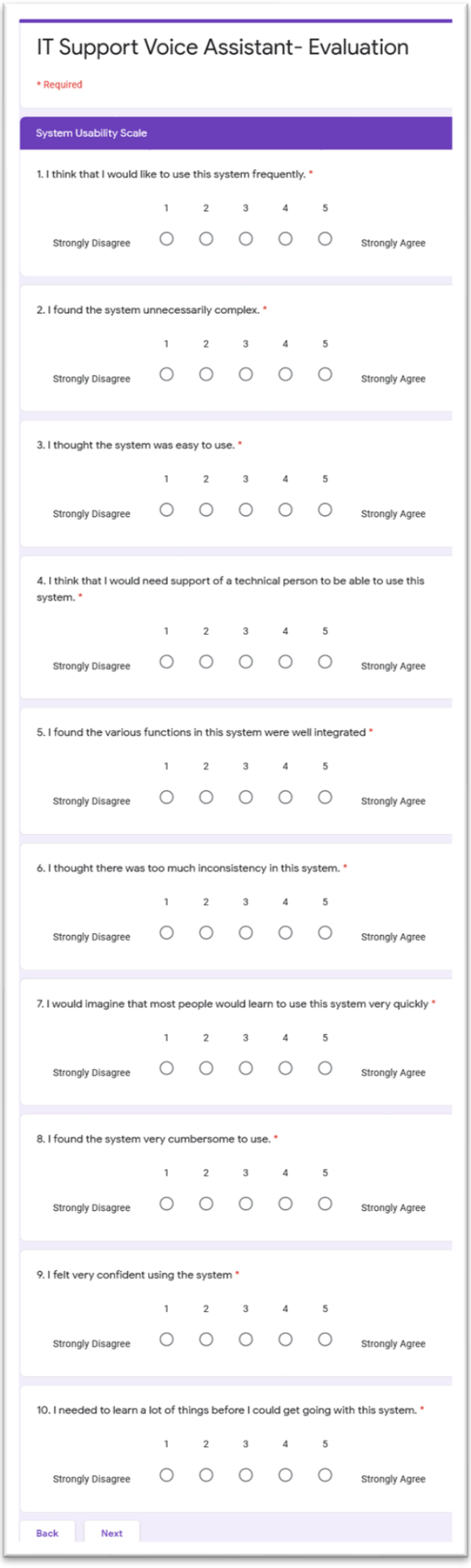
Name of Researcher: Mohammud Hishaam Moosafur

Please initial box

1. I confirm that I have read and understand the information sheet dated .....11th March 2021.....for the above study and have had the opportunity to ask questions. [1]
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without penalty. [2]
3. I agree that this form that bears my name and signature may be seen by a designated auditor. [3]
4. I agree that my non-identifiable research data may be stored in National Archives and be used anonymously by others for future research. I am assured that the confidentiality of my data will be upheld through the removal of any personal identifiers. [4]
5. I agree to take part in the above study. [7]

Your name	**12/03/21**	**Your signature**
Name of participant	Date	Signature
Mohammud Hishaam Moosafur	**12/03/21**	
Researcher	Date	Signature

Appendix-I



IT Support Voice Assistant- Evaluation

* Required

System Usability Scale

1. I think that I would like to use this system frequently. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

2. I found the system unnecessarily complex. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

3. I thought the system was easy to use. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

4. I think that I would need support of a technical person to be able to use this system. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

5. I found the various functions in this system were well integrated *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

6. I thought there was too much inconsistency in this system. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

7. I would imagine that most people would learn to use this system very quickly *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

8. I found the system very cumbersome to use. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

9. I felt very confident using the system *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

10. I needed to learn a lot of things before I could get going with this system. *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Back Next