Faculty of Technology and Society
Department of Computer Science and Media Technology

Master Thesis Project 15p, Spring 2019

Interaction with Smart Assistant

By
Prakriti Dhang
Manaswini Kolluru

Supervisors
Per Linde
Bahtijar Vogel

Examiner
Johan Holmgren
Contact Information

Authors:
Prakriti Dhang
E-mail: prakritidhang@gmail.com

Manaswini Kolluru
E-mail: manaswinikolluru@gmail.com

Supervisors:
Per Linde
E-mail: per.linde@mau.se
Malmö University, Department of Art, Culture and Communication

Bahtijar Vogel
E-mail: bahtijar.vogel@mau.se
Malmö University, Department of Computer Science and Media Technology

Examiner:
Johan Holmgren
E-mail: johan.holmgren@mau.se
Malmö University, Department of Computer Science and Media Technology
Abstract

Newer technologies are being developed all the time in order to help us in our daily tasks. Smart Assistant (SA) is one such broad category of devices which, when instructed through speech or command, perform certain tasks.

Most SAs have two primary tasks: identify the user, and perform the task communicated by the user. The task of identification can be achieved in several ways and one such method involves face recognition interface integrated within SA. Further interaction with the user is achieved via either a touch screen, or a speech-based interface integrated within SA. Since the primary objective of an SA involves interaction with an individual user, it is necessary to identify users' concerns related to the usability of SA.

In this thesis, the goal is to perform a usability study of Smart Assistant(SA) in the context of providing assistance in scheduling meetings at the workplace. Adopting a qualitative approach and we performed the usability study by designing a questionnaire to document public perception of new technologies like a smart assistant. We performed an online survey and collected responses from a section of population. We also performed a similar survey with employees at Cybercom who have experience in interacting with the SA being developed in their company, and which is meant for providing assistance in scheduling meetings. In the process, we also provide relevant background literature on human-computer interaction to contextualize the results of our survey. The main outcome of this work is a list of guidelines to be followed for developing a socially interactive speech-based SA.

The results of our survey indicate that there are several privacy and security related concerns that users have when using a face recognition interface and speech-based interaction. Although some of the concerns require technological advancements, a few of the concerns can be addressed by adopting certain suggested strategies.

Keywords: Smart Assistant, User Experience, Face Recognition, Speech Based Interaction, Human Interaction, Challenges, Internet of Things.
Popular Science Summary

Our reliance on digital devices has been increasing over the past decade or so. With the continuous evolvement of technology, new products are launched rapidly into the market aimed at further improving the lives of common people. Now we live in an era where we can control the air conditioner with the flick of our phone, and turn off the oven with another flick of the phone. All such products involve intricate design integrating hardware and software platforms. The next generation of such products see the seamless integration of several devices (IoT) and seek to further improve the user experience.

The new age products are often categorized as “Smart Appliances, which not only receive instructions but also are capable of intelligent interaction.” Possibly some of the most commonly known smart devices available in the market are Alexa, Google Assistant, Siri, and Cortana.

In this thesis, we perform a qualitative user study of the SA app built by Cybercom which is designed for smart scheduling. From the theoretical side, we focus on the foundational aspects of conceiving and designing a smart device, which broadly is categorized as user-centered design (UCD). On the practical side, we aim to test the usability of the product by designing a questionnaire to gather information on the perceived challenges faced by users when interacting with a smart device.
Acknowledgement

I am profoundly grateful to my supervisors Per Linde and Bahtijar Vogel for their guidance and continuous encouragement throughout the thesis work. I would like to thank Mahmoud Passikhani and Dennis Zikovic from Cybercom.

Secondly, I would like to thank my fellow classmates for their support and also to all my professors.

Finally, I express my sincere heartfelt gratitude to my parents and my beloved sister for supporting and encouraging me with their best wishes, during the course of the work.

Prakriti Dhang

I am profoundly grateful to my supervisors Per Linde and Bahtijar Vogel for their guidance and continuous encouragement throughout to the project. I would like to thank them for the interest they took in the project idea. I would like to thank Mahmoud Passikhani and Dennis Zikovic from Cybercom.

Secondly, I would like to thank my fellow classmates for their support and also to all my professors.

Finally, I express my sincere heartfelt gratitude to my husband and my beloved son for supporting and encouraging me with their best wishes, during the course of the work.

Manaswini Kolluru
# Contents

Abstract ................................................................. ii  
Popular Science Summary .............................................. iii 
Acknowledgement ......................................................... iv  
List of Figures ............................................................ vii  
List of Tables .............................................................. viii  
List of Acronyms ........................................................... viii  

## 1 Introduction  
1.1 Motivation .......................................................... 2  
1.2 Project Idea .......................................................... 3  
1.3 Research Goals ...................................................... 5  
1.4 Research Questions .................................................. 5  
1.5 Outline ................................................................. 5  

## 2 Literature Review  
2.1 Smart Assistant ....................................................... 8  
2.2 Interaction and Identification ....................................... 12  
2.3 User-Centered Design (UCD) ......................................... 17  
2.4 Issues and Challenges ................................................ 19  
2.4.1 Face Recognition .................................................. 19  
2.4.2 Speech-based Interaction ......................................... 20  
2.4.3 IoT based Application ............................................. 21  
2.5 Summary ............................................................... 22
# Research Methodology

3.1 Research Methodology and Approach ........................................ 26
  3.1.1 Pilot Study ................................................................. 28
  3.1.2 Survey ................................................................. 29
  3.1.3 Co-design on different contexts of SA ................................. 30

3.2 Data Collection Techniques .................................................... 31

3.3 Ethical Considerations ............................................................ 32

# Results

4.1 Pilot Study ................................................................. 33
  4.1.1 Improvements in the Questionnaire .................................. 36

4.2 Surveys ................................................................. 37
  4.2.1 Online Survey ............................................................. 37
  4.2.2 Web-based Questionnaire with Cybercom Employees ............ 42
  4.2.3 Summary of the Surveys ............................................... 46

4.3 Co-Design on different contexts of SA .................................... 47

# Discussion

5.1 Design Guidelines ............................................................. 52
  5.1.1 Design: Conceptual and Physical .................................... 52
  5.1.2 Identification ............................................................... 53
  5.1.3 Interaction ................................................................. 53

5.2 Limitations and Validity Aspects ........................................... 55

# Conclusion and Future Work

6.1 Conclusion ................................................................. 56

6.2 Future Work ................................................................. 58

Appendix A - Questionnaires for pilot study and online survey ............................... 59

Appendix B - Survey Questions for Cybercom Employees ........................................ 61

Appendix C - Co-Design ........................................................ 63

References ................................................................. 65
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Applications of Internet of Things [4]</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Flow diagram of the thesis</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Summary of Literature Review</td>
<td>24</td>
</tr>
<tr>
<td>3.1</td>
<td>Applied Research Methodology</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Co-Design: Interaction between hosts and participants</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>Preference to identify themselves (Pilot study)</td>
<td>34</td>
</tr>
<tr>
<td>4.2</td>
<td>Preference of platforms (Pilot study)</td>
<td>34</td>
</tr>
<tr>
<td>4.3</td>
<td>Medium of interaction (Pilot study)</td>
<td>35</td>
</tr>
<tr>
<td>4.4</td>
<td>Have used smart assistant (Online survey)</td>
<td>37</td>
</tr>
<tr>
<td>4.5</td>
<td>Preference to identify themselves (Online survey)</td>
<td>38</td>
</tr>
<tr>
<td>4.6</td>
<td>Identified via face (Online Survey)</td>
<td>39</td>
</tr>
<tr>
<td>4.7</td>
<td>Privacy Concern with face recognition (Online Survey)</td>
<td>39</td>
</tr>
<tr>
<td>4.8</td>
<td>Security Concern with face recognition (Online Survey)</td>
<td>40</td>
</tr>
<tr>
<td>4.9</td>
<td>Preference of platforms (Online survey)</td>
<td>41</td>
</tr>
<tr>
<td>4.10</td>
<td>Mode of interaction (Online survey)</td>
<td>41</td>
</tr>
<tr>
<td>4.11</td>
<td>Have used Smart Assistant (Online survey with Cybercom)</td>
<td>43</td>
</tr>
<tr>
<td>4.12</td>
<td>Comfort Vs Privacy (Online survey with Cybercom)</td>
<td>43</td>
</tr>
<tr>
<td>4.13</td>
<td>Trust Vs Experience with Speech-based interaction (Online survey with Cybercom)</td>
<td>44</td>
</tr>
<tr>
<td>4.14</td>
<td>Experience of Frustrating Moment (Online survey with Cybercom)</td>
<td>45</td>
</tr>
<tr>
<td>4.15</td>
<td>Medium of Interaction (Online survey with Cybercom)</td>
<td>45</td>
</tr>
<tr>
<td>4.16</td>
<td>Trust Vs Experience Vs Comfortability on Speech-based (Online survey with Cybercom)</td>
<td>46</td>
</tr>
</tbody>
</table>
List of Tables

2.1 IPAs [8] ................................................................. 11
2.2 Learned from the Literatures ................................... 25
3.1 Data Collection Technique ....................................... 32
4.1 Preference to end the conversation (Pilot study) ............ 35
4.2 Preference to end the conversation (Online survey) ........... 42
4.3 Summary of the Surveys ........................................... 46
4.4 Summary of the Co-design session ............................... 49
List of Acronyms

**IoT** Internet of Things

**SA** Smart Assistant

**IoTaP** Internet of Things and People

**IA** Inteligent Assistant

**PA** Personal Assistant

**IPA** Inteligent Personal Agent

**ISO** International Standards Organization

**UCD** User-Centered Design

**UML** Unified Modeling Language

**SUI** Speech User Interface

**GUI** Graphical User Interface

**RFID** Radio Frequency Identification

**IVRS** Interactive Voice Response System

**WSN** Wireless Sensor Network

**TAM** Technology Acceptance Model
Chapter 1

Introduction

The history of communication [1] is broadly divided into three ages. The first age is wireless, which is given by Marconi [2] who gave us station to station wireless telegraphy using spark equipment. The second age is a station to people where unforeseen expansion in the wireless is made. The third age of wireless is people to people connectivity where the internet has played a major role in the day to day life. Apart from people to people connectivity, we can connect people to things and things to people, and things to things. This led the way for the Internet of Things (IoT) [3] which is currently in great demand. Figure 1.1 shows the different applications of the internet of things.

![Figure 1.1: Applications of Internet of Things [4]](image)

Networking various sensors, actuators, circuits, and many other physical devices in our day to day lives is the key concept of IoT [3], paving way for smart lives.
Master Thesis Project: Interaction with Smart Assistant

through smart homes, smartphones, smart cities with sustainable transportation and energy solutions [5]. We now know that the next generation of smart devices will also incorporate context-based decision making, which involves state of the art machine learning and artificial intelligence technologies [6]. Clearly, research in these directions will lead to improve industry efficiency, energy consumption, agriculture, business, and our health, leading to a sustainable society [7].

SA has already penetrated a large number of households. Alexa is one such smart home assistant. There are many SAs which can be termed with different names such as “Intelligent Personal Assistant(IPA)”[8], “Digital Assistant(DA)” [9], “Autonomous Robot(AR)” [10]. We know there have been several ideas regarding SA, already incorporated in a smart home. However, there is still enough room to introduce new elements to the existing application of the smart house, cities, etc as we see new devices getting introduced almost every fortnight. With these new devices, we are getting swamped with more and more data, which could potentially be utilized to help us in leading a sustainable and healthy lifestyle. For instance, data from fitness tracker [11] devices can be used by smart assistant to suggest daily diet, and amount of rest to keep the body healthy and stress-free. In addition, keeping in line with modern green technologies and sustainable living, we could integrate electricity, heating, and other energy-intensive equipments with a smart assistant to minimize our carbon footprint.

Our broad aim is to perform a usability study for a SA being used and developed at Cybercom, and thereby to help them in the process of further developing their SA. Our primary focus areas in relation to a smart assistant are identification(of the user) and interaction (between SA and the user). Through the usability study, our goal is to identify different issues faced by potential users.

1.1 Motivation

Smart Assistant(SA) as an application of IoT [8] is an intelligent agent that helps users in their day-to-day work. The SA can help improve the assistance which is offered to users by collecting information autonomously from the environment. Siri, Alexa, Google Assistant are such examples of SA which we use in our daily life.

In our perspective, a SA, will help to book a meeting room at the desired place. The SAs’ task is to book a room for a recognized user. The user is recognized by face, which requires that the user’s information must be in the database. In general,
we can think of a SA in Universities, Schools, and homes as well. For example, users interact with the smart assistant in ATM machine [12] via speech-based interaction. The identification of the individual is achieved via face recognition [13], which is one of the many ways to biometrically recognize the individual. Indeed due to the complex algorithm involved in face recognition, there are several practical issues to be considered for widespread implementation of such software in potentially sensitive situations.

For the purpose of the study, we define SA developed at Cybercom as a device which has the ability:

- to identify the user by face
- communicates via speech, and
- helps booking meeting rooms.

Treating this as a functioning prototype our motivation is to explore the extent of usability of the SA in the general public domain which includes offices, homes, healthcare, public spaces.

Specifically, our objective in this thesis is to reach out to the current and future users of a SA and identify potential concerns that users may have faced regarding identification and interaction with SA. Also, in the process, explore users’ the perspective of the expansive reach of SAs etc. By performing a usability study on the prototype and unearth potential issues to be addressed in utilizing SA in the aforementioned domains.

The novelty of our work lies in investigating potential usages of SA and also unraveling potential concerns/issues to be faced by future users of SA with integrated face and speech-based interaction. Additionally, the novelty of our work also lies in bridging the gap between theory and reality. The cutting edge research in smart technologies is often claimed to be driven by demand and potential applicability of the technologies. Through this study, we shall gain significant insight into the extent of acceptability of new technology by users at large, and by juxtaposing the results of our survey with the available literature we shall be able to provide validation to future users’ concerns.

### 1.2 Project Idea

The theme of this work is in the broad category of Smart Living, which is a part of research projects in the Internet of Things and People (IoTaP) research
center at Malmö University in collaboration with Cybercom. The basic premise of Smart Living centers around the application of IoT in building, offices, institutions, universities, and homes to not only ease our lives, but also enhance the efficiency of our work, and hence add value to our lives.

There are many smart assistant systems available in the market. Cybercom is one of the companies that has developed the smart assistant for meeting room planner. In short, the overall system works in the following way: A registered person is first identified via face recognition; the identified person can then book a room for a meeting by instructing the smart assistant to do so. Instruction to a smart assistant is conveyed via speech. Several issues may occur during such an interaction due to the inclusion of two very complex components: interaction through speech and face recognition.

We explore several practical and technical aspects of such a smart assistant which may be relevant in large scale implementation of such technology. We adopt a qualitative approach to explore the aforementioned aspects. We began with a pilot study with a few questions. After analyzing the result, we reformulated the questions and initiated two online surveys: one comprising of a heterogeneous mix of people from a section of population, and second with Cybercom employees. The results are presented to infer people’s perception about smart assistants. Next, we used a co-design approach to explore the SA in different contexts such as homes, offices, health care, and public spaces. The figure 1.2 shows the overall processes done in the thesis.

Figure 1.2: Flow diagram of the thesis
1.3 Research Goals

Predominantly, the primary objective of this study is to perform a usability study of a smart assistant focusing only on identification and interaction aspects of the smart assistant.

We performed a questionnaire-based survey to contact people from a wide array of professional backgrounds and collect information about their experience with similar smart assistants. Specifically, we were interested in gathering information about various negative and positive perceptions related to using a smart assistant. Additionally, through this survey, we strived to understand perceptible changes in which people are looking forward to the basic design of smart assistants. Additionally, we performed co-design sessions with a small group of individuals to further, gain an extensive understanding of users experience and perspective of SA.

Using the results of our survey and co-design sessions, we formulate guidelines to be followed for designing a SA which is widely accepted by the users in different domains such as offices, home, health care, and public spaces.

1.4 Research Questions

In this thesis, the following Research Questions are identified:

RQ 1. What are the potential issues in using face recognition and speech-based interaction with a smart assistant?

RQ 2. What degree of comfortability one can achieve when SA is integrated in different contexts?

The above research questions will be discussed in chapter 6. To answer the two research questions we used an online survey to gather and sense (potential) users’ concerns and their level of comfort when interacting with a SA. We specifically, performed two surveys, one of which was targeted to gather public opinion at large, and the other was aimed to gather information from the users of SA at Cybercom. We also, performed co-design sessions to address RQ2.

1.5 Outline

The master thesis is divided into several chapters. Having introduced the basic motivation, project idea, research goals and research questions in this chapter, we
in chapter 2 present the relevant literature review of smart assistants and human-robot interaction. Next, in Chapter 3 we describe the research methods applied during the research process of the thesis. Subsequently, in Chapter 4 we discuss the results of our surveys and co-design sessions, leading to the discussion of our work in Chapter 5, where we present the design guidelines, limitations and validity aspects of this work. Finally, in Chapter 6, we present the conclusions and possible future directions to explore.
Chapter 2

Literature Review

A literature review is conducted with the aim to give insight into the current state-of-the-art research related to various aspects of smart assistants. We begin Section 2.1 with a basic discussion of smart assistants presenting a historical overview of the developments in socially interactive robots/devices, and the desiderata to be implemented in any socially interactive device/robots, subsequently, discussing different avatars of smart assistants: intelligent assistants and personal assistants. We end this section with a discussion of few examples of smart assistants in a range of challenging environments pushing the boundaries of applicability of smart assistants and also documenting the new challenges.

Since our focus is on the interaction and identification aspects of a smart assistant, we then move onto discuss various aspects of interaction and identification in Section 2.2 by beginning with examining the theoretical foundations of any verbal and non-verbal communication channel, and the evolutionary aspects of communication which involve learning from imitation. We then move our attention to the practical implementation of interaction technologies in smart assistants by providing a discussion of various issues to be handled when integrating a speech-based interaction within a smart assistant in different contexts. We finally conclude the section with a discussion about identification in smart assistants. Thereafter, in Section 2.3 we discuss how to quantify if the basic objective of a smart assistant, which is to help users. This is achieved by going over relevant international certifications in what is called, *user-centred design (UCD)*. Next, in Section 2.4 we discuss few very recent and relevant papers detailing the issues and challenges involved in identification and interaction with regards to smart assistant. Finally, we summarise the literature review in Section 2.5.
2.1 Smart Assistant

Heuristically, SA is an advanced robot with a certain specific set of tasks, which often involve delicate verbal/non-verbal communication with the user. We must not here that the history of robots dates back to the industrial revolution. Any new analysis of various technological and social aspects of usage and development of robots, therefore, needs to dig through reams of papers of research. The authors in [10] present an exceptional survey of various issues and developments of socially interactive robots. The authors begin with documenting the origins of the development of robots, specifically focusing on certain aspects like stigmergy to simulate human-like collective behavior, which is central to any socially interactive system. Broadly, Fong et al. [10] focus on robots which are defined as socially interactive, and list specific aspects of design methodology when conceiving a socially interactive robot:

- First is the design approach which is classified as: biologically inspired ([10], pp: 147), which have a sound basis in science, and functionally designed [14], which appear intelligent, even though their internal design is not very well founded in science.

- Next, having chosen one of the approaches, one must address the issues associated with the specific design ([10], pp: 148), which include correctly perceiving human activity; expecting believable human behavior from robot thus ensuring effective peer-peer communication; etc.

- Having resolved the design issues, one must focus on embodiment [15] of the robot, which broadly relates to how the physical presence of robot effects its physical neighborhood, and how the environment, in turn, effects the robot.

- Embodiment also includes morphological aspects like the appearance of the robot, since appearance effects how the robot is treated. For instance, a cleaning robot like Roomba is treated differently as compared to putting a life like character CERO as a robot “representative” ([10], pp: 150).

- Another critical aspect of human-robot interaction is the emotional aspect ([10], pp: 151). Since it is one of the most challenging qualitative aspect, the first steps involve clear categorisation, and dimensionalization of emotions. Speech, facial expressions, and body language, are just a few individual aspects of emotions.
For an interacting robot, the medium of communication is via dialogue, which involves several aspects, and they are discussed in Section 2.2.

Thereafter, when conceiving a socially interactive robot, it is important to associate a personality to the robot, which are clear, distinctive qualities of the robot, which make it “unique” ([10], pp: 146,149). The authors note the well established “Big Five Inventory” ([10], pp: 155) when describing a personality: extroversion; agreeableness; conscientiousness; neuroticism and openness.

Next is a highly abstract but equally relevant aspect that of perception [16]. For a coherent communication between robot and human, the robot must be equipped to understand and appreciate the human perception of various objects/subjects. Perceptions are usually encoded/decoded via sensors. The authors discuss literature related to human visual perception for the robot to have similar “visual motor control” ([10], pp: 155).

Closely related to the above is the ability to perceive human behavior ([10], pp: 157), specifically the ability to interpret and react to a certain behavior. Mostly, such ability can be incorporated by specifying information about potential users of the robot.

Another aspect, which is discussed [17, 18] in detail in Section 2.2 corresponds to learning from the environment by way of social learning and/or imitation.

Finally, for a robot to be able to establish a clear channel of communication, it must also understand and predict behavior ([10], pp: 158).

Having established the basic functional platform of any socially interactive machine like a smart assistant, we focus now on specific design aspects of an intelligent assistant, which is designed to perform cross-domain tasks spanning an array of resources. Specifically, Sun et al. in [19] focus primarily on an intelligent assistant capable of performing the following three tasks:

- “discovering meaningful intentions from users’ past interactions;” ([19], pp: 169)
- “leveraging surface intentions from group of apps;” ([19], pp: 169)
- “talking about intentions via natural language.” ([19], pp: 169)
In order to design such an assistant, the authors in [19] propose to learn the flow of actions to be performed by way of collecting usage data from the user’s smart phone, for example. For simple tasks, maintaining a database of past interactions together with the associated sequence of dialogues suffices to work. However, believably the space of past intentions and sequence of dialogues is going to be large, and thus one can utilise standard clustering algorithms using training datasets to perform actual clustering, or nearest neighbour algorithms to perform appropriate grouping. The authors propose two basic strategies to generate apps serving a specific intention. First strategy is to prepare new apps which themselves are ordered sequence of apps designed to serve specific intention (RepSeq). Second is to label multiple apps to each individual intention (input). Then, for every intention, a specific set of apps gets activated (MultLab). Clearly, RepSeq has an advantage that it is easy to execute, whereas once a sequence is initiated, the input language has no influence on the selection of apps, thus rendering it not actively interactive. Additionally, the authors in [19] highlight that for a clear execution of a task, it is necessary for the agent to communicate its understanding of the intention to the user. This allows for fault rectification, and prevents unnecessary faux-pas. It is here that the agent can use the output of clustering to summarise the intentions, by picking key phrases. The authors in [19] performed comparative analysis of cluster based vs. neighbour-based intention models, and RepSeq vs. MultLab by implementing standard statistical procedures, and conclude that under RepSeq strategy performed better with the neighbour based intention models, whereas MultLab strategy performed consistently irrespective of the type of intention model.

Building on the same paradigm of integrating intelligent assistants with other devices, Santos et. al [8], in a very recent work, redefined smart assistant as “Intelligent Personal Assistant (IPA)”, and performed a survey on IoT which can be used with a smart assistant, also providing the relevant IoT protocols. While listing the available IPAs (see Table 2.1), the authors [8] describes how IPA and IoT can be merged. The authors illustrate one scenario, where they explain how an IPA can get integrated with other objects in the surrounding. Specifically, the authors describe the scenario wherein the IPA responds to an alarm ringing in the morning by opening the curtains, and subsequently the coffee machine and toaster doing their jobs automatically as the person is getting ready for office. These IPAs can help the user to prepare for going to an office. The authors describe an achievable solution which can solve the network layer problem which causes
interruption when connected to devices. They also list the use of IPAs in a smart home and healthcare field with scenarios.

<table>
<thead>
<tr>
<th>Personal Assistants</th>
<th>Area</th>
<th>Specification</th>
<th>Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIRI, Google Assistant, Alexa</td>
<td>General</td>
<td>Use to do all kinds of useful things</td>
<td>Set alarm, Reminders, Preview calendar</td>
</tr>
<tr>
<td>HealthPal</td>
<td>Healthcare</td>
<td>Used on ambient-assisted living environments.</td>
<td>Reminds each time a task should be done</td>
</tr>
<tr>
<td>Adele</td>
<td>Education</td>
<td>Desktop based</td>
<td>Helps in learning course</td>
</tr>
<tr>
<td>Intelligent network-based assistant</td>
<td>Personal Communication</td>
<td>Maintaing, modifying and terminating real time sessions</td>
<td>Used to manage SIP based communications.</td>
</tr>
</tbody>
</table>

Table 2.1: IPAs [8]

Focussing on a specific environment, Kaliouby et al. [20] explore several issues that may arise when developing a family based Personal Assistants (PA) in the future. They illustrate few scenarios in a family environment where a PA could play an effective role. For instance, in a family gathering, this PA can help introduce other members of the family. The authors describe the communicative and inferential capabilities needed to design such devices, and then test the social and emotional capabilities of a hypothetical device built on such capabilities. While discussing the privacy, security, and ethical, aspects related to such design, the authors highlight the privacy and security of children when interacting with PA. The authors also discuss the philosophically relevant questions like “what these PAs can do? and whose assistant are they?” ([20], pp: 1033) and by suggesting that “in carrying on dialogues with people, such assistants would need to behave in fashion that people expect, per expectations of cognition and with norms of action over time” ([20], pp: 1034).

Pushing the boundaries of applicability of smart assistants, Isbister et al. [21]
introduce new application areas for smart agents in computer interface. Beginning with a discussion on an interface to assist human to human communication in a virtual environment, the authors propose a prototype which interacts with the users, and revives a lagged conversation by initiating a conversation by finding a common ground between the users. The authors found that the agent improved the users’ experience, and contributed positively to the users’ perception of each other.

2.2 Interaction and Identification

Before we discuss the specifics of speech-based interaction in the context of the smart assistant, there are many intelligent communication applications available, which are very hard to generalize. As basic, an action, it is for us building a machine capable of intelligent communication is an extremely intricate task.

In view of our project whose key component involves human interaction with an interface, which establishes a channel of communication between human and computer, we believe it is critical to examine various aspects of human-robot interactions. While listing the challenges overcome by the technological advancement in designing interactive robots, Mavridis in [17] stresses the need for following certain desiderata when designing any future device which involves interaction and communication with humans. The motivation for designing such a list stems from the need for flexibility in the utilisation of any automated machine/robot. Designing robots for specific tasks is passe, and the current need is for robots which can possibly learn and perform newer tasks, which calls for certain fluidity in the functionality of robots. The desiderata proposed by the authors in this paper tries to provide an empirical platform for designing robots with a wide range of functionalities including human interaction via speech. The desiderata dive deep into analyzing the basics of the science of linguistics. It begins with stating that the robot should ideally be programmed to follow speech besides following simple commands, and that it should be adept at handling multiple speech acts like requests, assertives, directives, commisives, expressives, and declarations. Additionally, the robot may sometimes also have to lead a conversation, as in FaceBots, rather than just following conversation. Another issue which is highlighted through the desiderata is that of encoding the “words” as specific objects, or encoding them by also augmenting them with their meaning, allowing the robot to possibly make intelligent decisions. However, the authors do note that designing such robot
would likely involve much advanced technology, and more research needs to be done to realize such a robot. Adding another layer to the semantic layer of an intelligent speech-based system, the next desideratum concerns with affected interaction, meaning to establish an emotional connection with the responder. For a robot to effectively communicate with humans, it is desirable to have the robot also understand non-verbal communication like gestures, eye gaze coordination. Another one of the desired desiderata suggests that the robots should act and interact with a specific purpose or objective. Such a design can be implemented via different methodologies like the Markov decision process, for instance. Another challenge to be overcome in order to be able to design a robot with fluidic verbal and non-verbal abilities is the formulate how and when does the robot learn. The author notes that online learning is clearly a difficult paradigm and requires immense care. This issue also resonates with formulating the principles of interaction of the robot with other online entities/services. The proposed desiderata can be used as a template that every future interactive communication device must conform to in order to achieve the objective of equipping the device with natural language capabilities.

Building further in the theory of human-computer interaction, researchers taking cure from the methods in machine learning, propose a natural foundation to build a robot capable of intelligent communication. Specifically, Andry et al. [18] proposes a neural network architecture design that encodes the idea of “learning by imitation” to design a machine to learn and communicate, at least at low levels. In order to understand the process of imitation, which seemingly comes naturally to many living beings, it is crucial to understand the theoretical underpinnings of imitation through a perspective of developmental psychology. The authors note that the early developmental psychologists classified imitation as immediate and deferred. The modern understanding of imitation among neonates and infants proposes imitation as a coupled process involving perception and action, which relates to immediate imitation. Based on these theories, the authors propose a model of an imitative system which primarily performs two types of tasks: should respond well in do as I do tests (see HAYES and HAYES); and also perform spontaneous imitation due to internal motivation, or due to high novelty factor associated with the action to be imitated. While illustrating an example, the authors highlight that imitation can be induced by perception ambiguity. Meaning, incapability to differentiate between objects and thus tricking the robot to synchronize its visual and motor sensors. The authors in [18], however, point that such an endeavour is
highly dependent on the experimental setup as minute changes in distances and/or angles can lead to different perceptions leading to different imitation. It is noted that integrating this paradigm of perception ambiguity with the concept of control, is at the heart of reproducing a specific trajectory of motor sequences. That is, the robot incorporates the visual input, and while reproducing the motor sequence it applies appropriate control to fit the motor sequence. This has three major component neuron groups: Time base, Time derivation and Prediction output. The three components refer to the three actions: keeping temporal account of past events, identifying new input event, and predicting a pattern which is different from the past, respectively. In the next level, the authors discuss the synchronization effects, when two robots enact the same motor sequence and they try to synchronize their movements. It is however noted that in such dynamics, a balance between independent (re)production and adaptive (re)production is necessary to be established. Requisite changes are introduced to the architecture to ensure that the pathway from perception to action is merged with simple production, without causing any interference. It is also concluded in simulated experiments that synchronization is an attractor of interactions. Finally, the authors suggest a similar architecture for rhythm prediction as an instance of learning new sensory-motion association without explicit reinforcement. It is suggested that the prediction mechanism can be used as a tool for detecting a change in the pattern of rhythm.

In the context of our thesis, the smart assistant being used and developed at Cybercom already used speech-based interaction to communicate with the user. In this aspect, we now present some of the works highlighting various issues involved in speech-based interaction. One of the main challenges in implementing a speech-based interaction is that speech is inherently a context-based medium of communication, and therefore no one rule fits all.

Building on this premise, Hone et al. [12] discussed different constraints on speech-based interaction with computers. One of the constraints they found is the “semantic constraint where users are permitted to express in their inputs to a system” ([12], pp: 640). Second is the dialogue constraint “dialogue constraint where users are permitted to express given the local dialogue context” ([12], pp: 641). The third is the “syntactic constraint which defined in terms of the number of different paraphrases which are allowed for expressing a given meaning” ([12], pp: 641). Fourth is the “lexical constraint which gives restrictions in terms of the individual words which are acceptable” ([12], pp: 642). Fifth constraint they enumerated is the “recognition constraint that refers to limitations at the level of
recognizing spoken words” ([12], pp: 643). Apart from finding constraints, the authors present two studies that aim to explore the issue of habitability in the speech-based system with the ATM machine. The first study tests the system with the users aiming to investigate the behavior of users. The system works in the following way, the user has to speak out the commands that are displayed on the screen, then the system is analyzed based on the constraints. Their second study was based on “telephone home banking” ([12], pp: 650), to investigate the effective level of constraints. From both the studies, the authors conclude that these results would give a direction to the developers in designing a speech-based interactive system.

Despite the challenges faced in developing speech-based socially interactive devices, we are surrounded by many such devices. From Siri [22] to Alexa [23] and many more such speech devices are the helping hands for many of us. These so-called assistants help in performing our tasks in a shorter period of time. For example, we do not have to use keypads or hands to book a train ticket or to search for the weather forecast. Instead, we just need to instruct the assistant to perform a task and the task is performed. On the other hand, we do come across many issues experienced by the users of such commercially available smart assistants.

In support of the usability study performed by us, it is worth mentioning that such studies have often been considered very practical, and fruitful in understanding the user experience with regards to a device, leading to improvement in future designs of similar nature. Specifically, Yankelovich et.al in [24] discuss the SpeechActs system, which is an experimental prototype developed at the MIT for speech applications. The authors, present the detailed functionality and methodology of the system, and also list several challenges and issues in speech-based interaction systems. The basic premise of the SpeechAct system relies on learning and redesigning with information gained from user feedback. The authors discuss several scenarios for applications. For instance, in one of the speech-based application the authors describe an “email application” ([24], pp: 369), where the actions include reading, sending and discarding of emails. The next application they described is the “calendar application” ([24], pp: 370), where the user asks the system about the schedule for the present day or the next day or it may be also some other day. In another application, the authors describe “weather application” ([24], pp: 370), which provides a detailed weather report. The authors propose to run these applications simultaneously, and the user can switch from one application to another application by simply saying goodbye to the current ap-
plication. In order to test their proposed system, the authors performed usability studies, and used the formative evaluation philosophy of Jakob Nielsen [25] in the context of changing and retesting their interface. We present a detailed overview of the challenges discussed by the authors in Section 2.4.2. Among several of their findings, the authors point that getting a response from the system after every interaction exudes confidence among the users about the quality and correctness of interaction.

Finally, in the context of a system of interacting devices, we are all very well aware of the developments in the area of Internet of Things (IoT). In a survey [5], various visions of IoT paradigm are found and the enabling technologies are reviewed. It is observed that the interest in the concept of the Internet of Things among the researchers is increasing. It is discussed about various domains in which IoT is in use including the smart home. Atzori e. al [5] have discussed various domains or applications based on IoT can be used, such as transportation, smart environment, health care, personal and social, futuristic. They have also discussed about the privacy and security issues. From the observation, they found that the IoT should be considered to be a part of the internet of the future.

This paper [7] discusses how the IoT based smart home automation system is built and what advantages one can get. They discussed about the sensors and how different types of sensor can help in controlling the required system. It is mentioned that sensors being part of IoT can be used to gather data and based on that data, information can be generated to make some intelligent decisions. It is observed that the IoT based system has been broadly used for building a smart home as it can say smart living or automated home but also used in developing smart cities where it is more concerned about developing smart transportation and smart buildings, smart offices and many more.

We now conclude this section with a brief overview of various identification methodologies in smart assistant. It is very clearly seen that in the present world, we value security, and a small error can have serious implications. In such a scenario, biometric based identification has shone itself apart with due to it being extremely conservative toward error. We see people getting identified by face, fingerprint, iris, retina, signature, DNA, and/or voice. There are many other ways of identifying people which can be stated as a biometric recognition system. The authors in [13] discuss different types of biometric recognition systems. The biometric recognition includes physical features for identifying persons. The authors describe a biometric system, which first verifies the user and then identifies. The
authors provide a great in-depth description of the overall functionality of the system. One of the most commonly used biometric recognition attributes is face recognition. There is also a multimodal biometric system. This means that it identifies a specific person by using two different features. In this, both iris and fingerprint are used as well as face and some other features maybe used to identify. All these intricacies help further reduce the possibility of an error in such a high stakes environment. We provide further discussion of various issues and challenges in adapting biometric identification involving face recognition in Section 2.4.1.

2.3 User-Centered Design (UCD)

Since a smart assistant by definition is designed to help the users, it is broadly categorized as user-centered design. In such a case, it is necessary to have clear guidelines in defining the basic constructs of a product. It is in this regard that we now discuss the status-quo of standardized guidelines of user-centred design.

When designing a product with specific usability, it is necessary to follow a set guidelines so as to achieve optimal user satisfaction. Keeping this precise principle in mind, international standards organization proposed ISO 9241-11 standardization which provides a universal definition of usability for product developers to follow [26]. As per this standardization, usability is defined as The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use. The keywords here are effectiveness, efficiency, satisfaction, and context of use, which are further clarified in ISO 9241-11. Additionally, it also defines crucial aspects of usability like goals and tasks. Often definitions like those given above are stated in an abstract way, however, in terms of measurable outcomes, the definition of usability can be used to determine: users of the system; goals of users; environments of usage; measures of effectiveness, efficiency, and satisfaction. Additionally, a year later in 1999, the organization proposed ISO 13407, which provides guidance on human-centered design activities. Broadly, it describes user-centered design by proposing to provide:

(a) a clear rationale for the design with the user satisfaction, reduced costs and training being put at the center of the design,

(b) basic principles which drive user-centered design, like involving the users at each stage of development so as not to lose sight of the target,
(c) efficient planning of the whole design by taking regular user feedback,

(d) a clear set of activities to be performed.

Since the above two standardizations pertain to user-centered design specifications, it is critical to examine these proposed guiding principles in view of the standardized definition of usability as proposed in ISO 9241-11. In the paper [26], Jokela et. al performed this precise exercise of carrying out an interpretative analysis of the latter by contextualizing it with respect to the former to test for consistency. The authors, while providing a striking comparative study of the two standards, observe that the guidance in ISO 13407 fails to provide a set definition of usability as it fails to provide descriptions of critical elements of usability like effectiveness, efficiency, and satisfaction. The authors also conclude that ISO13407 does not address the challenges related to the identification of different users with different goals. Finally, the authors suggest that although ISO 13407 provides several important guiding principles in the development of user-centered design, however, due to the aforementioned issues, it is a good practice to implement ISO 13407 and ISO 9241-11.

In a specific scenario of application of UCD, Afrianto et al. [27] while noting the problems with the usage of “online journal aggregator system”, argue in favour of the software development method to find the users need. Their objective being to execute the UCD method when developing the “online journal aggregator system”. UCD method can be also used to design game-based applications, health care system, and many more. The authors use the Likert scale method to determine user preference and their acceptability of the “online journal aggregator system”. “Accessibility, content aspect, and navigation” are the three categories of Likert method. The authors present their research method in two sections: The first section discusses the data collection method. The data was collected through various media: literature, conducted interviews with the journal managers, researchers and observation technique. The second section elaborates the five process involved in the UCD method:

- The first process is to identify user requirements.
- The next process is specifying the context of the use of such a product.
- The third process describes the functional and non-functional requirements of the user or organization. This can be presented using a UML diagram.
- The fourth process provides the design according to user needs.
The final step is to the evaluation of the design in order to find whether the requirements meet the user expectations.

The authors interviewed 30 participants, among which 20 were researchers and 10 were journal managers. The authors show their results in a tabular form containing three columns of data collection method, participant type, and results in the description. It is found that 82% of the participant believes that this system can further be developed to the next stage.

2.4 Issues and Challenges

2.4.1 Face Recognition

As pointed earlier, there are several issues to be dealt with when working on a face recognition based identification procedure. Specifically, the paper [28] discusses different types of challenges in facial recognition like age, expression, image quality, and background. As per their findings, one of the major challenges they discuss is aging. The authentication of the person varies if the image in the database is 10 years old., implying that the system needs to be updated frequently. A system that is not adapted to aging is likely to cause several misclassifications. Another challenge related to the image quality, which can affect the facial features. There are different types of problems that can affect image quality. The problem could be with the illumination, distortion and many other different factors. The authors also observe that apart from aging and image quality, facial expression can also be one of the causes in not recognizing a particular person. Based on these challenges, one can improve the facial recognition model.

Delving deeper into the intricacies of face recognition, Ohlyan et al. in [29] provide a survey of various problems and issues in face recognition. In addition to aging, image quality, and facial expression, the authors identify a few more potential issues while recognizing a particular individual. The authors begin with classifying facial features into two categories, which are the “intrinsic factors and extrinsic factors” ([29], pp:2534). The problems also can be with the direction or angle of a face. Another problem could be with low resolution when the resolution is less than some fixed value. “Occlusion” ([29], pp:2536) could be another issue that fails in recognizing a face. Apart from these core problems, they have also mentioned that the problems could also be in the system. If the system has faults then it will be difficult to analyze the result. The background noise, camera dis-
toration or low database storage. Network problem can also be one of the problems in recognizing.

2.4.2 Speech-based Interaction

In [24] the authors provide a detailed description of various challenges involved in speech-based interaction as a result of their usability study. They enumerate four major challenges that might affect the speech act system. The challenges are categorized in the following way:

- The first issue, the author described is the “Simulating Conversation” ([12], [24], pp: 371). In this category, the authors have discussed the “pacing” ([24], pp: 372) and “prosody” ([24], pp: 372) words which might affect the conversation. It is found that the pausing between the conversation results in delayed in recognizing. The use of prosody words are difficult for the user to understand if the system uses any kind of prosody words. Not only the difficulties for the users, but it is found that the system also faces the same difficulties in using such words.

- The next issue in the speech-based interaction is “the nature of the speech” ([24], pp:375). The nature of the speech means the cognitive behavior where the author thinks that these cognitive behaviors are important in speech user interface(SUI) interaction. Here the authors compare between Speech User Interface (SUI) and Graphical User Interface (GUI). It is mentioned that visual feedback in the SUI is deemed very useful and interactive from the user perspective. Since in GUI the user can visualize the interactive behavior, whereas in SUI it is only dialog based interaction, where the user has to wait for the response from the smart assistant, and during the time user has to look at the blank screen which might affect the productivity of the system. Another issue in the nature of the speech section is “speech and persistence” ([24], pp:375), which means a human can perceive written words easily as compared to spoken words. This may lead to misinterpreting the words which conclude in the wrong information. Another issue they listed in this category is the “ambiguous silence” ([24], pp:376), where it has a double meaning. The first could be that the speech is being processed or recognizing the speech. The second could be that the assistant is unable to hear due to the background noise or the user’s low voice.
• “Converting GUI into SUI” ([24], pp: 372) is the third category. In GUI, the information is followed sequentially and the user can clearly visualize the steps. On the screen, the user can move on the next step simply by clicking the next button. Whereas if GUI is transformed to SUI there might be some issues since every step will be dialogue-based. Sometimes this may be very time taking if the assistant doesn’t understand the speech, and if it asks repeatedly about the speech. This was included as a subcategory titled “information flow” ([24], pp: 373). The second subcategory is the “information organization” ([24], pp: 373), comparing to SUI the authors state that the information displayed in GUI the information are arranged in a sequence. Vocabulary is also one of the issues in dialogue based conversation. In GUI there is built-in vocabulary from where the user has to proceed to the next step. On the other hand in SUI, it is deemed difficult since the conversation is handled verbally.

• One of the major issues in speech-based interaction is “Recognition Errors” ([30], [24], pp:374). This is when the speech-based interaction system unable to understand user speech. This may lead to user dissatisfaction if they get a continuous rejection from the assistant saying I didn’t get you or I didn’t understand. Refusal from the assistant might also lead when the user is not speaking in a clear and normal tone. This might confuse the assistant and may start uttering I don’t understand what you said. The other kind when the assistant heard a similar pronounced word, which will show a different result. They provide one example to avoid this type of situation. It is said that instead of providing the answer directly, it would be better if the smart assistant confirms the message like by replying as did you say?. This means a verification is needed before providing the correct results.

2.4.3 IoT based Application

A survey was conducted in [31] and have identified challenges and security faced from the perspective of IoT based applications and technologies. Their survey was mainly focused on the technologies and application which includes a smart home system, health care system. Key challenges include the security, privacy, standards, trained workforce. It was also found that researchers face different problems like authenticity, interoperability, privacy, data confidentiality, low range of internet signal, power supply, power backup, fault tolerance, reliability, cost,
poor support, and most important awareness and skills. They listed different types of security and risk factors and also listed different types of challenges in a tabular form for different IoT based applications. They mention scalability, mobility, data protection. Regarding the technology, they found that the risk can be the front end sensors, can be a network issue or it can be the backend of the system. There are many possibilities and uncertainties in its application case scenarios. Thus, it is important to consider solutions and improve these challenges that create value for the stakeholder, industries, academics.

In related work, a holistic framework is proposed in [32] to integrate the smart home with a cloud-based platform. The authors identify tools that will help in managing the home. Besides proposing the framework the authors discuss the challenges that one can face while designing the application. Challenges found are Big data, that means when data is large it is likely going to be difficult to analyze. Networking issue can be one of the challenges when trying to connect with WSN. Finally, the most effected challenges are the security and privacy.

Challenges characterized by uncertainty in the sensor data are also pointed out in [33]. This needs certain representation after network processing procedures, which leads to another key research problem: how to reorganize and represent sensor data and provide effective integration of uncertain information. There is another challenge when it comes to resolving the apparent contradiction between large-scale heterogeneity and the dynamics of the IoT system, and the requirement of highly efficient data exchange. They propose mechanisms and methods of information integration and interaction of IoT, service delivery of IoT and verified the platform of IoT.

2.5 Summary

There are several theoretical aspects to be considered when designing a socially interactive device. Any interactive device, by definition, interacts with the users, and sometimes also interacts with the environment and other interactive devices. Each type of interaction is to be treated differently with different designs and structures. When dealing with socially interactive devices requiring human interaction, another layer that of identification gets introduced in almost all such instances. Our primary objective in this chapter was to provide an overview of the basic paradigms and various challenges involved in designing a smart assistant working on a platform of speech-based interaction, and integrated with biometric-based
identification. Specifically, our focus is on understanding the bare-bones structure of identification and interaction procedures, and highlight the challenges.

In order to understand identify, highlight and appreciate the issues and challenges involved in any design, we must first understand the fundamental design of the same. We, therefore, began with understanding the basics of socially interactive robots as explained in [10], which is an abstract umbrella involving all smart assistants, intelligent assistants, and personal assistants [19, 8]. These papers are necessary for our thesis as it helps us gain significant insight into the fundamentals of a smart assistant. Through these papers, we learned that the smart assistant being used and developed at Cybercom is a functionally designed device with a specific functional objective. On the practical side, it is stated that socially interactive devices must respect the environment. This is very relevant to the case at Cybercom as the smart assistant is working in an office environment, and thus several parameters involving privacy, confidentiality and respecting the workspace become extremely important.

Moving next to the interaction part of a smart assistant, we began with laying down the foundation of any communicative device as provided in [17], and discussed the basic heuristic of learning by imitation as documented in [18]. In case of future development of Smart Assistant, it is recommended that the device be allowed to learn from imitation as pointed in [18]. The proposed paradigm of learning by imitation is well grounded in evolutionary biology and linguistics.

In terms of basic requirements to be met by any speech-based interactive device, we learned that the literature supports developing smart assistants which not only follow conversations but also lead them. For instance, in the case of smart assistant used for scheduling, if the users instruct the smart assistant to book a specific meeting room, and if that room is unavailable, then the smart assistant should lead the conversation by suggesting alternatives based on the past usage and requirement of the individual user. It is suggested that a Markov decision process [17] can be integrated with a smart assistant to help do this. In [19], we also learned that clarity of communication is extremely necessary for the effective execution of the functionality of a smart assistant. The authors in [19] propose to use standard clustering algorithms to identify several words which mean the same, which also is very relevant to the development of the smart assistant at Cybercom.

Contextualizing various modes of interaction, the authors in [12] document relevant constraints for the two modes of interactions: screen-based input and speech-based dialogues. These constraints provide an excellent skeletal structure to
the basic design involving speech-based interactions. Although the smart assistant at Cybercom does not involve interaction with other smart devices, we presented some recent developments in the broad area of interacting devices, which is called IoT. Our primary source being an excellent survey authored by Atzori et al. [5], and [31] for a detailed description of the challenges involved in implementing such technology.

Further, delving into the potential issues and concerns in relation to speech based interaction, we documented several predicaments as elicited in [24, 30], wherein the authors present a clear comparison between GUI and speech-based interface, and point out pitfalls in both the technologies. Driving our discussion to the aspects related to identification in smart assistants, we provided a basic overview of several technologies [13], and highlighted the issues stated in [28, 29].

In order to integrate all our discussion in the context of a broad category of user-centred design (UCD), we presented the widely accepted international standards for UCD as recorded in [26], and we also discussed its applicability in certain scenarios as presented in [27]. We provide a brief summary of our discussion in the following figure 2.1 and table 2.2:

![Figure 2.1: Summary of Literature Review](image-url)
### Topics

<table>
<thead>
<tr>
<th>Topics</th>
<th>Outcomes</th>
</tr>
</thead>
</table>
| SA              | **Fundamental design:** Socially interactive SA and the challenges involving smooth integration of such an SA in a given environment including issues of privacy and confidentiality [10, 19, 8];  
**Fundamentals of communication:** Fundamentals of any communicating device involving interaction with user, and designing verbally interactive robots using the heuristic of learning by imitation; Use of Markov decision process in algorithmizing the process of leading a conversation by suggesting alternatives [17, 18];  
**Speech:** Having clear communication and using standard clustering algorithms to identify similar meaning words [19]; contextualising different modes of interaction [12] |
| IoT             | Applications and challenges in integrating SA with IoT enabled devices[5, 31]                                                                 |
| Face recognitions | The basics and the challenges involved in implementing face recognition interface for identification [13, 28, 29]                             |
| Speech-based interaction | Practical challenges in using a speech based interactive device[12, 24, 30]                                                               |

**Table 2.2:** Learned from the Literatures
Chapter 3

Research Methodology

This chapter begins with the research approach. It discusses all the activities that took place in the research work. Based on these literatures we have framed our questions. The chapter discusses about the data collection technique and ethical consideration in section 3.2 and section 3.3 respectively.

3.1 Research Methodology and Approach

Surveys [34] are one of the research methods used in research. Surveys are data collection exercises which help the researcher gain valuable insight into the proposed theory. Researchers typically look for patterns which align with their goals, and which can be generalized to a wider populace, and therefore can be shown as evidence of the validity of the proposed theory. This makes surveys very valuable for research. Qualitative research [35] is often time intensive, primarily because it requires collecting data by interacting with people over long periods of time. Subsequently, analyzing conversations and bringing out insights is also a time-intensive exercise.

To investigate the usability and productivity of the system, we used survey method. Referring to the research questions, the intention is to bring out the concerns with the smart assistant technologies. We described our research methodology in Figure 3.1.

We initially started with finding literature related to our keywords. Thereafter followed by a pilot study, to understand the users’ perception. After analyzing the data collected through a pilot study, we conducted an online questionnaire.

We collected the data from a diverse cross-section of the working population. Our sample included individuals from a wide variety of workplaces including En-
engineering, Information Technology, Space Research, Pharmaceutical Industry, Human Resources, law firms, and academics. The sample of individuals chosen for our survey was driven by the need to collect data from a wide section of the working population. Additionally, due to the technical theme of our questionnaire, we could not choose a random sample of individuals for our survey. Implying that many of our inferences are to be interpreted accordingly and generalizations are to be made with utmost caution.

It is expected to answer two of our research questions RQ1 and RQ2 as mentioned in section 1.4. A questionnaire is prepared thus addressing RQ1. Co-design approach is used to answer our RQ2. This will contribute knowledge on value adding aspects of smart assistants when integrated face recognition and speech-based interaction in different contexts. In order to fulfill our goal, we began with a pilot study and thereafter we followed by online surveys and co-design session, which are described in the sections 3.1.1, 3.1.2, and 3.1.3 respectively.
3.1.1 Pilot Study

We started a pilot study with a list of 12 questions. We used the Google form to post our questions and sent to our colleagues. We put two scenarios which help them understand filling out the form. The scenarios are described in section 3.1.1.1. Our pilot study consists of both numeric and descriptive data.

3.1.1.1 Scenarios

**Scenario 1:** This scenario involves in what ways one wants to be identified by a smart assistant. For instance, when the computer says, “Hi June,.....” it implies that the computer has already identified the individual. So, what kind of identification feature would you prefer to use? Traditional log in/-password? Identifying via some interface like identity card? Face recognition or other bio-metric technologies?

**Scenario 2:** In this scenario the objective is to book a room for specified purpose like meeting, conference, discussion, class, etc. Specially, the scenario deals with interaction of user with a smart interface using an interactive voice response system (IV-RS) for booking a room. In the scenario, you supply parameters required for booking a room through talking to the assistant, which responds in computer voice, i.e the interaction is completely speech-based. The parameters to be supplied can be listed as: user information – user’s name, etc. type of event – discussion meeting / class / conference, etc. time and duration of event – date(s), time and duration of the event logistical requirements – tables, chairs, computers, projectors, TV screen/monitor, recording facility, etc. After having supplied the above parameters, the smart interface looks into the system and assigns a room.

The pilot study was conducted in phase 1 of our work, to test the validity of our questionnaire and we opted to collect the responses from our fellow classmates. This choice was primarily driven by the intent to collect responses from a sample of individuals who are aware of the latest technologies and can be trusted with providing valuable feedback on the design and content of our questionnaire. We would like to mention here that we received very constructive feedback not only about the structure of our questionnaire but also the content. We, however, do intend to send the questionnaire out to the users of similar devices and wish to gather wholesome information for our main analysis. The online survey consists
of similar questions with a more structured way. Section 3.1.2.1 discusses online survey in a comprehensive way.

3.1.2 Survey

We conducted two online surveys, section 3.1.2.1 consists of people have no experience in using SA. Section 3.1.2.2 includes the employees at Cybercom who have experience in using SA. Further details of each surveys are described below.

3.1.2.1 Online Survey

In this thesis, the purpose of this online survey is to find out the privacy concerns of SA and peoples' view of such SAs. The online survey involved a sample size of 34, where the participants were from varied professions such as students, professors, IT employees, core engineers, human resource managers, and pharmaceutical industry professionals. The online survey was conducted by attaching a link of the Google form and sent via E-mail. The form had a description of our work and motivation to post these questions. Our survey consisted of both numeric and descriptive data. This survey results will help in answering the RQ1. The online survey questions are listed in Appendix A. 4.2.1.

3.1.2.2 Web-based Questionnaire

As recommended, we carried out a web-based questionnaire instead of interviews. A web-based questionnaire is conducted with Cybercom employees. A web-based questionnaire is also referred to as an online survey.

The purpose of this web-based questionnaire is to find out the privacy concerns and degree of comfortability of using SA. The study is conducted with the employees at Cybercom who have experience with the SA meeting planner. Our web-based questionnaire consists of both numeric and descriptive data. The questions for Cybercom employees are listed in Appendix B. The results from the web-based questionnaire will help in answering RQ1. In this thesis work, simple and short answered questions were designed based on the research question. Most of the questions are based on “Likert scale”, “Yes/No”, and few are open-ended.

Participants Our research relies on the data gathered from employees at Cybercom. This company has implemented a functioning prototype which is a “smart meeting planner”. The detailed working principle of the prototype is
described in section 1.2. The questions are related to the use of this prototype and the use of such technologies. The participants are the experienced user of the SA meeting planner.

**Conducting Web-based Questionnaire** Web-based Questionnaire is conducted by sharing a link of the Google form. Initial contact with the employees at Cybercom was established by contacting the concerned person over email. The content of the email included clear details of our research work and our purpose for conducting the questionnaire. A detail of the questionnaire was sent to the company for review and approval. Once approved we started with making the web form and ensure the link with the employees. The data were collected and stored in Google form. The results are discussed in chapter 4.

### 3.1.3 Co-design on different contexts of SA

Co-design approach is a collaborative design approach in which researchers, designers, or potential users are considered as “experts of their experience” [36]. The researchers, designers, or potential users share views and work as a team with the designer. The views of these users result in creating design guidelines.

To explore the usage of a smart assistant in different contexts, we considered four different scenarios and conducted the respective co-design sessions. The four different scenarios are as follows: Office environment, Public place (Eg: Shopping malls, bus stands), Health care Industry, Home. The co-design session took place in the Orkanen lab. The chosen participants are the experienced user of SA. Figure 3.2 shows the interaction between the hosts and the users. The host’s task is to describe all the scenarios and ask questions to the users. The results of the co-design approach is discussed in section 4.3 of chapter 4.

**Participants** We invited our participants through messaging protocol. We named our study as a design session for Smart Assistant. Participants are between 25 and 35 ages. Participants are from various occupations such as software engineers, researchers, and students. The total session took place between 5 and 6 hours. Each session took about 60 minutes and each scenario took 15 minutes. The sessions are mainly focused on “When, What, Which” of SA.

The steps followed to perform the co-design session are enumerated below:

- The first step is to decide the number of participants to include, what are the pre-requisites things for the session, and how to conduct it.
Secondly, we made an agenda, with all the details of the session, that includes date, time and location.

In the third step we provide the materials and explained the scenarios.

In the Fourth step, we conducted a questionnaire session for every scenario and record their views with the permission of the participants.

Final step we analyzed the collected data.

### 3.2 Data Collection Techniques

In this thesis work, we have three different types of studies. The first is the pilot study, second is the online survey, third is the co-design session. Questions are designed in a simple and structured way. Responses for most of the questions in our thesis are Likert scale, and few are with “Yes/No”, and “Agree/Disagree”. Short paragraphed questions and answers are also used. Table 3.1 shows the collected data from each study.

To answer the two research questions, data from each study were assembled and analyzed. The textual data was manually reviewed from the sheet and recorded system.
<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Data</th>
<th>Technique</th>
<th>Methods/ Approach</th>
<th>Subjects/ Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Literatures</td>
<td>Journal papers, Conference papers, and articles</td>
<td>Literature Review</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>Pilot Study</td>
<td>Questionnaire</td>
<td>Survey Method</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>Survey with who haven’t used SA</td>
<td>Questionnaire</td>
<td>Survey Method</td>
<td>34</td>
</tr>
<tr>
<td>4</td>
<td>Survey with Cybercom</td>
<td>Questionnaire</td>
<td>Survey Method</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>Co-design session</td>
<td>Transcripts and recordings</td>
<td>Co-design approach</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3.1: Data Collection Technique

3.3 Ethical Considerations

Ethical considerations are one of the vital aspects of research. Bryman et.al [37] enumerates ten important key points, which should be considered in every study. We considered those key points and discussed them in the following paragraph.

Since our study is a qualitative method, it requires a collection of textual data. Before sending an online questionnaire we had informed the participants that this study will not be used for any research or product which may cause harm in the future. We had assured all the participants of the security of their data, and we also ensured that we respect respondents’ schedules in participating in our survey. We ensured that our participants’ personal data remains undisclosed. We discarded a few of the responses in case of misleading information. Before performing the task we provided all the necessary information to our participants. We fully explained the inference of participation. We had given authority to each and every participant to withdraw from the study at any stage [38]. The study has been done with honesty and transparency.
Chapter 4

Results

The chapter focuses on the results for each study. First, we discuss the results from the pilot study which is in section 4.1. The second section 4.2, includes two sections which contain the results for online survey and web-based questionnaire section 4.2.1 and section 4.2.2 respectively. The section 4.3 includes the results for our co-design session with a table.

4.1 Pilot Study

For the pilot study, we sent the questionnaire to 7 individuals, and we received 6 responses. Since individuals with the different background may have different perceptions and opinions about Smart Assistants, we begin our questionnaire by inquiring about the professional status of individuals. We learned that two of the respondents are employed in the corporate sector and the rest are students.

Next, respondents were to rate their experience of having used a smart assistant. The verdict appeared to be split with 50% reporting a reasonably pleasant experience and the rest expressing their dissatisfaction with their experience.

Although we had very few respondents, the above information right at the onset is very intriguing as it reinforces the importance of learning about the various issues or disadvantages rendering the smart assistant from being widely accepted or adopted, and thus thrusts this thesis into the limelight. Despite their displeasure, all the respondents appeared to support the view that smart assistants indeed improve efficiency at their workplace.

Next, the respondents were asked to choose their preference for the way they would like to identify themselves to the smart assistant. Interestingly the figure 4.1 shows almost everyone (5 out of 6) was in favor of face recognition as their
chosen mode of identification. Since the respondents were given the choice to choose more than one mode of identification a small the percentage also chose to identify themselves via RFID or passcode (33% each).

Figure 4.1: Preference to identify themselves (Pilot study)

Of all the respondents, a majority seemed to prefer an app-based smart assistant for the purpose of booking the rooms as shown in figure 4.2. However, a web (computer) based smart assistant is also acceptable to many.

Figure 4.2: Preference of platforms (Pilot study)

The respondents as shown in figure 4.3 appeared to prefer touch screen-based medium of interaction as compared to speech-based. This information appears to
be at odds with the technological development in the general area of smart devices.

![Figure 4.3: Medium of interaction (Pilot study)](image)

When given the choice to interact through speech the respondents appear to prefer to end their interaction by uttering "OK" as shown in table 4.1.

<table>
<thead>
<tr>
<th>Word to end the conversation</th>
<th>No. of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK</td>
<td>4</td>
</tr>
<tr>
<td>Thank you</td>
<td>1</td>
</tr>
<tr>
<td>Over and out</td>
<td>2</td>
</tr>
</tbody>
</table>

*Table 4.1: Preference to end the conversation (Pilot study)*

Most of the respondents believe that smart assistant would work well in the offices, schools/universities, and corporate offices. Apart from these environments, two of the respondents suggested the applicability of smart assistant in homes, hotels, and cars. The respondents also suggested that smart assistants could be very useful in helping the aged and differently-abled people in our society.

The respondents primarily expressed major concern related to privacy, authentication, and security. Many were of the opinion that the “old-styled” keypad, was a safer and more efficient mode of communication with the smart assistant because
it minimizes various potential problems related to speech based interaction like interpretation; accent; language; diction; verbal and non-verbal skill mismatch; etc.

With regards to identification using a face recognition tool, the respondents expressed genuine concerns related to the reliability and safety of the tool. The respondents seemed to suggest that when the stakes are high then such safety concerns must be addressed with due diligence. It was also pointed out that the usability of such a high-end smart assistant can be improved by imparting appropriate training to the users. This one time exercise will surely help avoid foreseeable issues.

Eliminating the unclear answers by two responses, and considering the valid response from the respondents for voice UI is quite good. One of the respondents enumerated valid points such as having the option of multiple languages and multiple accents which could be appropriate for voice UI. Apart from this response, one respondent opinion for the voice UI, if don’t act accordingly to the said tasks, an option would be appropriate as a text input model. Another response was regarding notifying the person when arriving at the office if that person has a meeting.

4.1.1 Improvements in the Questionnaire

We learned through the pilot study that possible responses to the posed question did not allow for multiple correct answers, which we plan to rectify in our planned study.

As was observed in this pilot study people appear to prefer touch screen-based interaction as compared to speech-based, we feel it is relevant to identify the reasons for people to not prefer a speech-based interaction. Therefore we wish to add one more question: What are the potential problems/issues related to the usage of speech based interaction with a smart assistant?

Although the question was framed appropriately to first classify the respondents based on whether they ever used a smart assistant, there appears to be some room for improvement in the way the question is formulated because of the respondents appear to have and express their opinion of the usability of smart assistants in an office.
4.2 Surveys

4.2.1 Online Survey

The questionnaire consisted of 14 questions through which we wish to learn participants’ past experiences in interacting with smart assistants and face recognition interface and their perception about the usability of the smart device at their workplace. Through the questionnaire, we also compared participants’ preference for the two modes of interaction with a smart assistant. The primary conclusions to be drawn from this questionnaire concern with the public perception of new technologies with regards to common security and privacy issues. Broadly, the questionnaire is divided into three parts: Questions based on identification; questions based on the smart assistant’s platform and questions based on the mode of interaction with the smart assistant.

The questionnaire begins with collecting basic information viz name, designation (Occupation), etc. We received 34 responses to our online survey, out of which a majority (91%) of them had never used a smart assistant at a workplace environment. However, given the popularity and wide reach of products like Google Assistant, Amazon’s Alexa, and Apple’s Siri, we believe that many of the participants of our survey must have used or interacted with one of the above smart assistants in environments different from an office setting. Despite not having interacted with a smart assistant in the setting of a workplace, more than 78% of the participants expressed great confidence in believing that the introduction of smart

![Pie chart](image)

**Figure 4.4:** Have used smart assistant (Online survey)
assistant in an office may improve the overall efficiency in the workplace. It must be noted here that some of the 22% who were of the contrary opinion, expressed their misgivings of the usefulness of smart assistants at smaller workplaces with abundant resources.

The three participants in our survey with the experience of having interacted with a smart assistant in office expressed satisfaction of their overall experience with the device. Arguably, since the number of such participants is very small, it is impractical to draw any inference based on this information.

Next, we tested the level of users’ comfort with various technologies involved in identifying oneself with an autonomous device, with the options being: face recognition, RFID based identity card and password/passcode. Figure 4.5 the RFID based technology clearly was a winner with the highest support of almost 65% support, with face recognition based technology coming close second with 50% support of participants, and the password/passcode based identification came in a distant third with merely 38% support from the participants despite this being one of the most popularly used mode of identification currently in use in most of the internet based applications. We believe that the recent incidents of data theft may have influenced participants’ opinion of a password/passcode based identification method.

Since the smart assistant being developed at Cybercom involves face recognition based identification, we intended to further implore participants’ past experience of their interactions with the face recognition interface. Incidentally, a
mere 36% of our participants had ever interacted with a face recognition interface, which possibly alludes to the fact that the face recognition interface as a method of identification is yet to reach its full potential of wide usability.

When inquired about participants’ level of concern regarding their privacy when using a face recognition interface, a mere 15% expressed almost no concerns, whereas a significant number (33%) expressed moderate privacy concern, and finally, a whopping 52% expressed serious privacy concerns with face recognition based interface, as shown in figure 4.7.

Continuing on the theme of potential concerns with using face recognition interface we measured participants’ level of concern related to the security of users...
data on a scale of 1-10 with 1 being least concerned and 10 being highly concerned. Our data in figure 4.8 shows that 68% of participants expressed reasonably high level of concerned and almost 32% of the participants expressed a very low level of concern regarding the security of user data.

![Pie chart showing security concern levels](image.png)

**Figure 4.8:** Security Concern with face recognition (Online Survey)

Specifically, the participants’ concerns ranged from lack of trust in the technology involved in face recognition; to the exorbitant cost involved in the implementation of such technology; to concerns regarding the algorithm driving the face recognition interface. This pattern appears to suggest that although 50% of the participants support face recognition based technology, they do have serious security and privacy concerns. Additionally, the participants were of the opinion that face recognition based identification can possibly be used in highly secure areas like airports, banks and other security establishments. Juxtaposing the survey participants’ serious concerns regarding security and privacy with face recognition and their support for using the same at highly secure locations seems to paint a very contradictory picture. Logically put, this seems to suggest that users think of identifying false positives as high risk feature, i.e., falsely identifying one individual for another. This information can potentially be used by the designers and manufacturers of the device for testing its efficiency, and performing quality tests. The conclusion to be drawn from this data is that commercial implementation of face recognition based interface in any form needs to be preceded by a public information campaign to eradicate misconceived notions from public domain.

Moving on to the physical manifestation of smart assistant we inquired the participants about their preference for a computer-based or an app-based smart assistant. From our data, it is difficult to clearly conclude participants’ preference
for either of the two. However, in figure 4.9, the pattern of participants’ scores for the two interfaces, and their individual comments seems to suggest that an app-based smart assistant is likely to gain wide acceptability among potential users.

Next, we focussed on participants’ opinion about the mode of interaction with smart assistant. Our emphasis was to compare the touch screen-based interface with preset options and an interactive voice response system (IVRS). The verdict in our survey appeared to favor a touch screen-based interface as shown in figure 4.10.

However, since the smart assistant being developed at Cybercom is a speech-
based interactive device, we intended to further investigate participants’ concerns regarding the speech-based interactive interface. Many participants opined that incorporating different accents and dictions into the software is extremely essential for the device to be widely acceptable. It was also pointed out that validation of the conversation in a speech-based interactive interface must also be integrated within the system. This may possibly be achieved by supplying a print of the conversation at the end of every interaction. Table 4.2 shows the preference of words to end the conversation.

<table>
<thead>
<tr>
<th>Word to end the conversation</th>
<th>No. of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK</td>
<td>11</td>
</tr>
<tr>
<td>Thank you</td>
<td>16</td>
</tr>
<tr>
<td>Over and out</td>
<td>7</td>
</tr>
</tbody>
</table>

| Table 4.2: Preference to end the conversation (Online survey) |

We also noted that basic functions of speech, as the end of the conversation, have to be well programmed to allow for clear communication. Due to several issues related to basic aspects of speech-based interaction, many participants didn’t favor using the technology in any environment. However, some participants suggested that speech-based interfaces can be used to guide visually impaired users, and at quiet places where background noise is unlikely to corrupt the interaction.

Despite many concerns, there was broad acceptability for smart assistants, and many participants suggested the implementation of smart technologies at industrial plants.

4.2.2 Web-based Questionnaire with Cybercom Employees

The questionnaire consisted of 10 questions with the objective to learn the user experience of a smart assistant at Cybercom. Like the online survey where we measured general public response to face recognition and speech-based interaction, this survey also aims to do the same but with a different population. Since many of the employees at Cybercom have already had used a smart assistant, this questionnaire was modeled slightly different as compared to the online questionnaire. We received 18 responses out of which almost 89% had used smart
assistant at Cybercom.

![Image](image1.png)

**Figure 4.11:** Have used Smart Assistant (Online survey with Cybercom)

This is starkly different from the responses we received on an online survey where most of the respondents had not used a smart assistant. The two employees who had not used the smart assistant stated that the smart assistant being used at Cybercom has issues related to activation and interaction. It was suggested that the speech-based interaction need to be streamlined for the effective usage of a smart assistant. It was pointed that devices using identification interfaces should be placed in a private place instead of a corridor at workplace. Out of the 16 respondents who had used, almost 38% of the respondents expressed their displeasure by rating their experience on the lower half of the scale. None expressed complete satisfaction with their experience with the smart assistant, and the general perception can be defined to be just about average.

![Image](image2.png)

**Figure 4.12:** Comfort Vs Privacy (Online survey with Cybercom)
Since the objective of our thesis focuses on the identification and interaction aspects of a smart assistant, we measured the satisfaction level of the users of smart assistant at Cybercom. Through our survey we learnt that a whopping majority (78%) of users have expressed great satisfaction with the face recognition interface for identification. This implies a face recognition system in a SA can also simplify all the tasks. A small (22%) of the respondents did, however, express that they have strong privacy concerns about the face recognition interface as shown in figure 4.12.

Interestingly, all the respondents who expressed privacy concerns agreed that an anonymous identification, even using face recognition, will address their concerns. All the respondents who had used the smart assistant at Cybercom expressed relatively high trust in that the booking was indeed made their name.

![Figure 4.13: Trust Vs Experience with Speech-based interaction (Online survey with Cybercom)](image)

Moving on to the interaction part of the smart assistant, the satisfaction levels among the users was not as high as with the identification procedure, with a majority of the respondents preferring to give above average rating. Indeed a large percentage (72%) of the users admitting to have had frustrating moments while interacting with the smart assistant as shown in figure 4.14. Primarily the users stated very slow processing and unclear interaction protocols as major reasons for their dissatisfaction.

Despite these drawbacks, two-thirds of the respondents favored a speech-based interaction as opposed to touch screen based interface, and the same fraction of respondents favored an interaction interface which is an amalgamation of touch
When inquired if the users had any concerns identifying the end of their interaction with the smart assistant, majority of users expressed their satisfaction, and indicated that the end of interaction is clearly communicated by the smart assistant vocally and also by showing the acknowledgment of the booking.

However 22% of respondents gave a score of 5 or less on the scale of 10 in response to the question measuring the level of their trust in that the conversation
was indeed complete at the end of the interaction.

4.2.3 Summary of the Surveys

From the two surveys, if we compare the responses, we have listed few factors which are shown in table 4.3. We found that having no experience of using SA is more concerned about privacy and security. As Cybercom employees are using the app, so they are more comfortable as compare to the non-user SA.

<table>
<thead>
<tr>
<th>Key factors</th>
<th>Users who haven’t used SA (Survey with varied professions)</th>
<th>Users who have used SA (Survey with Cybercom employees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA in office</td>
<td>50-80% agreed</td>
<td>90% agreed</td>
</tr>
<tr>
<td>Concerned about privacy</td>
<td>More concerned</td>
<td>Less concerned</td>
</tr>
<tr>
<td>Identify via face</td>
<td>Not comfortable</td>
<td>Feels comfortable</td>
</tr>
<tr>
<td>Medium of interaction</td>
<td>both touch-based and speech-based</td>
<td>both touch-based and speech-based</td>
</tr>
</tbody>
</table>

Table 4.3: Summary of the Surveys
4.3 Co-Design on different contexts of SA

From the co-design session, we have found several opinions from each participant. There are some functional features that could be added in SA in order to make the SA more interactive. The detailed results for every scenario are explained below. For the co-design we mainly focused on the questions words like “when, what, which and why”. These are the questions: When do they feel that SA is good in use? What are the risks? What are the qualities? What new features can be added? Which context do they feel suitable to have face recognition and not speech?

A. Office Environment: Focusing on the usage of smart devices specifically meant for scheduling in an office environment, we investigated the usability and potential concerns emanating from the usage of smart applications. Among the interviewed individuals there was absolute unanimity in expressing their support for introducing such devices at the workplace. However, the verdict was split for, and against, the usage of face recognition for the purpose of identification. Privacy and discomfort were the primary reasons which stated for not favoring in face recognition. When inquired about their preferred choice of interface for interacting with the smart assistant, once again, the participants unanimously voted for touch screen-based interaction and voice-based interaction.

B. Home Environment: In home environment, most of the participants preferred SA which has the features of identifying the user via face and communicates verbally. But when arising the question about risks, few participants mentioned that they think no risks of having SA at home. Few of them mentioned about the storing of images could be a risk. It is also observed that it is not always an appropriate way of identifying a user via face. As asked about the new features, we got few responses about users mood which means that based on the users mood SA will start speaking. Apart from this new feature, the participants also preferred both touch screen-based interaction and voice-based interaction as a medium of interacting with the smart assistant.

C. Healthcare Environment: Contrary to participants’ views of risk-free usage of Smart Assistant in an office environment, the participants expressed serious concerns against the introduction of Smart Assistant in the health
care industry. The broad primary concern shared by all the participants relates to the handling of highly sensitive data, and high stakes involved in healthcare. The verdict for using face recognition as a mode of identification was once again split due to similar concerns as in the case of an office environment. Like in the case of the office environment, the screen based interaction was upvoted for the same reasons. It must be noted here that the concerns raised against the usage of face recognition by female participants. We recommend that this relationship between gender and aversion to face recognition be further investigated to address similar privacy concerns.

D. Public Space Environment (Shopping mall, Bus stop): Comparing with the other three scenarios participants showed more concern in having a SA in public spaces. They emphasize that no one will interest in having a conversation with SA in public. Here the concern arises in terms of privacy, as people will be around them and none will be interested in answering private questions. When asked about the risks everyone mentioned their personal details which are stored in SA. Apart from these privacy and risk factors, we then asked about new features that could be added. Summarizing all the responses, the new features they believe that would be suitable in SA are: 1) Navigation, extending this point to why navigation they conclude if the phone’s battery is dead, then SA in public space can be used for navigation. 2) To have speech-based and touch-based features. 3) Fascinatingly, one of the participants mentioned if SA is collecting real-time images, then based on the person’s mood it would offer discounts in foods or cloths if the SA is in a shopping mall or bus stops. From the above points, one participant mentioned that SA in the museum will be more applicable than in a shopping center or bus stop. Having SA in museum indicates that the user can ask SA about the history of the sculpture or painting. From this scenario, we summarize that SA with face recognition will have a negative impact than having only speech-based interaction.

From above all the scenarios, the participants preferred home and office environment are the most suitable and beneficial places to have a SA. Apart from the home and office environment, one participant mentioned the health care environment would also be a good idea to have a SA. Table 4.4 displays the summary of all the scenarios.
<table>
<thead>
<tr>
<th>Categories</th>
<th>Office</th>
<th>Home</th>
<th>Health Care</th>
<th>Public Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualities</td>
<td>Scheduling</td>
<td>Help in daily tasks</td>
<td>Monitoring, patient</td>
<td>Navigation</td>
</tr>
<tr>
<td>Usability</td>
<td>Reduce user effort</td>
<td>Reduce user effort</td>
<td>Reduce user effort</td>
<td></td>
</tr>
<tr>
<td>Risks</td>
<td>Stores private data</td>
<td>Stores private data</td>
<td>Stores private data</td>
<td>Stores private data</td>
</tr>
<tr>
<td>Added features</td>
<td>Multiple language</td>
<td>Multiple language</td>
<td>Multiple language</td>
<td>Multiple language</td>
</tr>
<tr>
<td></td>
<td>Notification to specified user</td>
<td>Notification to user in case of theft</td>
<td>Notification to the doctor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Both touch-based and Speech-based</td>
<td>Both touch-based and Speech-based</td>
<td>Both touch-based and Speech-based</td>
<td>Both touch-based and Speech-based</td>
</tr>
</tbody>
</table>

Table 4.4: Summary of the Co-design session
Chapter 5

Discussion

This chapter aims to gather all our findings and address the stated research questions. Later in this chapter, we also discuss the limitations and validity aspects of our study.

The results of our survey and co-design sessions can broadly be categorized as relating to...

• ... the general perception of SA among the population
• ... the fundamental design of SA which includes the physical manifestation.
• ... the identification of the user by SA.
• ... the interaction with SA.

There are several aspects within each of the above categories. For instance the issues concerning security and privacy appear in identification and interaction.

We begin with noting that despite many of the participants in our online survey having never had interacted with a SA, a majority of them expressed great confidence in the belief that the SA’s can truly improve the efficiency of work in an office environment. This implies that the potential users of SA perceive great comfort in using SA. In the parlance of technology acceptance model TAM, the participants of our survey are weighing heavily on the perceived usefulness. It is possible that with widespread usage of smart phones, many people have a positive opinion of the usability of smart devices.

The results of our online survey indicate that there are many people who have serious security and privacy concerns with regards to using a face recognition interface for the purpose of identification in an office environment. These concerns are...
in line with the problems and issues with the face recognition interface listed and discussed in [28, 29]. Many of the concerns possibly stem from the lack of knowledge about the algorithm used in the background by a face recognition interface. Specifically, some of the participants were concerned whether the face recognition interface could be tricked by showing an image of an individual instead of the individual physically presenting himself/herself to the face recognition interface. These issues are broadly related to the fundamental design aspect of SA, as every smart device should ideally be designed using the fundamentals of UCD as discussed in our literature and detailed in [26, 27]. It is very clearly stated that in a UCD the users should be involved at each stage of development to possibly gain the confidence and feedback from the users.

With regards to the interactive part of SA, the majority of the survey participants favored touch-based interactive interface as against a speech-based interface. It is here that, in the parlance technology acceptance model, the potential users of SA have influenced by the lack of perceived ease-of-use of speech-based interactive devices. The participants elicited very practical concerns like lack of clarity in communication (unclear interaction protocols); potential cause of distraction for fellow colleagues at the site of SA; and vice versa the background environment noise causing disruption in interaction. The issue of speech-based SA causing distraction to the co-workers in the vicinity of the device relates to violation of the idea of respecting the environment as detailed in [19, 8]. In many scenarios, this concern can be addressed by placing the SA in a secluded and private area in the office. This way, the potential users may also have reduced concerns with regards to the privacy of their conversation. The issue of the background environment noise causing disruption in communication is a design issue and can possibly be addressed by more sensitive sensors. Finally, the concerns raised by the survey participants which relate to the quality and clarity of communication have all been very well documented and researched in [24]. The participants of the web-based survey specifically mentioned many frustrating interactions with the SA which manifested due to unclear communication protocols. Such issues can be addressed by incorporating the fundamentals of verbal and non–verbal communication [17] during the design phase of speech-based interactive device, and using certain Markov decision process based algorithms which lead the conversation and simultaneously drive the communication protocols [18].

Reverting to the issue of security and privacy related concerns, the results of our co-design sessions indicate that implementation of SA in environments like health
care, office and public space should be taken forth with utmost care. Specifically, the concerns regarding the privacy of an individual (data) needs to be respected. It is interesting to note that identification through face recognition interface is perceived to be unwelcome by all of our female participants in all the environments. We are unaware of any literature which studies or addresses this gender aspect of face recognition interface. We believe that further research needs to be conducted exploring this line of research for a successful integration of face recognition with SA. Juxtaposing the above with the results of our online survey we note that many of the security and privacy related concerns may be alleviated by placing the SA in a secure environment. For instance, the participants in our online survey stated that they perceive higher comfort level in using SAs in highly secure environments like banks, airports.

One of the important take aways from our co-design session was that the participants expressed great hope that SAs could play a major role in improving the wholesome experience in the health care environment. However, the participants were extremely wary of the security of their personal data. We believe that by integrating several security measures and using the basics of the fundamental design of SA it is possible to use SA in certain areas of health care environment like inpatient wards.

5.1 Design Guidelines

In this section, we shall use the outcomes of our survey and co-design sessions, together with the relevant literature to formulate certain guidelines which can form a basic checklist and reference for future developers of socially interactive SA.

5.1.1 Design: Conceptual and Physical

The foremost conceptual aspect of any SA is to identify the functionality of the SA. This defines several related aspects of SA like potential user population, which in turn also identifies the domain specific issues and concerns. Thereafter one of the key characteristics of any socially interactive SA is that it must respect the environment. This includes

- Respecting the privacy of individual users
- Respecting the workspace environment
Despite the technology being very secure, since SA is to be based on UCD, all the relevant information about the security must be communicated to potential users. This can be achieved by performing the following actions:

- Organize training sessions for potential users to apprise them of the usability of the device.

- Respect and acknowledge the feedback of all the stakeholders, and gain the confidence of all the potential users.

The physical manifestation of the device should be welcoming to one and all without any bias. This includes the physical space where the device is placed.

5.1.2 Identification

The most commonly used methods of identification are:

- passcode
- RFID card
- face recognition or biometric based identification

Our survey appears to suggest that passcode and RFID card based identification are preferred over face recognition based method. However, our survey also proposes a solution to gain the favourable opinion of face recognition, which is by incorporating stringent security measures.

- Appropriate security arrangements like CCTV cameras could be used at the physical location of SA to further gain the confidence of users in favor of face recognition based identification interface.

We learned from our survey that female users of SA were extremely apprehensive of using a face recognition based interface.

- Female users of SA should be consulted, and their concerns should be addressed.

5.1.3 Interaction

We have learned from the theory that the design of speech-based socially interactive devices require extreme care due to the presence of several degrees of freedom.
It is therefore suggested that a lean and functionally focussed design be adopted for the speech-based interactive device. This can be achieved by setting stringent and closed communication protocols for the interaction which rule out any digression in the communication.

One way of achieving the goal of setting closed communication protocols is to design the SA in a way so that it always leads the conversation.

- A Markov decision process based algorithm as proposed in [19] can be adopted to achieve the objective.

Due to the diverse environment almost in every sphere of life, it is crucial that the SA be designed so as to understand the diction of a large section population [30].

In order to ensure clear communication, the SA may be designed in either of the following ways:

- by providing a written transcript of the conversation,
- or by showing SA’s understanding of users communication on the attached display screen,
- or by enunciating SA’s understanding of users communication

In order to neutralize the effect of diction on the interaction between SA and user, we suggest that...

- ... an alternative touch-based intercative screen may be used allowing the user to switch between speech and touch based interactions to avoid any confusion.

For broad acceptability among the users, we also suggest that...

- ...the SA be equipped to interact in different voices representing different identities like gender and linguistic identity.

This can be achieved by an advanced learning algorithm as proposed in [18], which is based on the premise of learning by imitation. However, designing such an algorithm is likely to involve complex machine learning algorithms.

### 5.1.3.1 Interaction in high stakes / sensitive environments

Our survey seems to suggest that the choice of interactive technology depends on the context of interaction. For instance, a touch-based intreactive SA is preferred
over a speech-based SA in environments where either the communication involves exchange of highly sensitive/personal information, or where the interaction involves high stakes.

5.2 Limitations and Validity Aspects

Surveys have many benefits and a few drawbacks. In our pilot study, we had a small number of participants and faced difficulties in generalizing the answers. We got a few responses which were not clear and some of the participants did not respond to all the questions leading to various problems in interpretation. As we performed an online survey, it was not possible to ask a follow-up question based on an answer. Surveys often face the problem of respondents not being honest in their responses. We also received many responses which were not only difficult to interpret but also irrelevant. Some of the respondents did not provide detailed information leading us to interpret their responses appropriately. Since our survey was centered around smart assistant, we expected the respondents to be aware of certain basic technologies like speech based UI and face recognition. However, despite our best efforts, we could not ascertain if all the respondents were familiar with the technologies and clear about the objective of our questionnaire.

In order to minimize bias, participants for an online survey (non-SA users) were selected from varied professions. The participants (SA users) for the web-based questionnaire, Cybercom employees were selected as they are experienced users of SA. A total of 4 participants were selected to conduct co-design sessions which might be a small sample to generalize results. In addition, the results obtained from co-design session are not validated with any practical implementations but can be considered as future work.
Chapter 6

Conclusion and Future Work

In this chapter, we conclude our study with a summary of our work. A thought for future research opportunities within the field of Smart Assistant and IoT has also been given.

6.1 Conclusion

We recall the first research question:

RQ 1. What are the potential issues in using face recognition and speech-based interaction with a smart assistant?

From our analysis of the available literature and the survey results we list the following issues which we deduced

- **Face recognition**
  - *Privacy*: Many users (female) feel violation of personal space when interacting with face recognition.
  - *Security (of data)*: lack of confidence in the SA keeping personal biometric data (image) safe and secure.
  - *Confidence in technology*: users appear to not trust the extant technology used in face recognition, even when they have never even interacted with one.

- **Speech based interaction**
– **Privacy**: users expressed that communicating with a speech-based interactive essay in certain environments, like crowded office settings, can cause the users specifically if the interaction if of personal nature.

– **Communication**:
  * **Communication protocols**: Lack of clear communication protocols can lead to unwanted scenarios.
  * **Clarity of communication**: users expressed concern if the SA indeed understood the conversation. For instance, simple actions like ending the conversation can become a bone of contention.
  * **Lack of diversity**: the language and tone adopted for the speech-based interactive device can have an unwarranted impact on the users.

We next recall the second research question:

RQ 2. What degree of comfortability can one achieve when smart assistant is integrated in different contexts?

The degree of comfortability one can achieve with SAs in different contexts clearly depends on the context. In public environments like shopping malls, bus stops and train stations, it is believed that SA can be of great help and users will appreciate the service, however in such environments a SA which is not speech based and which does not involve identification of the user will ensure high comfort levels among the users. In contrast, in highly secure environments like airports, banks and security establishments speech based interactive SAs which also include identification of the user using biometric data can achieve similar high comfort levels as in the previous case.

In high stakes environments that involve exchange of highly sensitive data, the SAs cannot blindly be implemented. In specific sub environments, the SAs may exude high levels of comfort among the users, whereas in certain different sub environments the SAs may not gain the confidence of the users.

It is interesting to note that there has been some research on the human tendency to not readily accept new technologies and be extra cautious when adopting latest cutting edge technologies [40]. In most cases, such negative perception stems from people’s distrust inability of new technologies in handling sensitive private data. In our survey, we witnessed similar behavior, and we believe that with continued efforts to bridge the trust gap, it is possible to lead people to overcome
their misgivings about new technologies, like the face recognition interface in our case. We also showed the possibilities of SA in different contexts and conclude our research as a significant contribution in this domain.

6.2 Future Work

There are several directions that are interesting to continue. Future development of SA may incorporate the guidelines discussed in 5.1 to build more inclusive SA enabling large acceptability of SA by ensuring high satisfaction levels among users. It is hoped that this study will lead to a knowledge contribution to the smart assistant integrated with face recognition and speech-based interaction implementation in IoT.
Appendix A - Questionnaires

Note: Considering ethical issues we will not disclose any kind of personal data (your name, and email address). All the personal data will be kept unknown.

Description: A Smart assistant, using the said technologies, will help to book a meeting room at your workplace. The smart assistants’ task is to book a meeting and room for a recognized user. The user is recognized by face recognition. And for this the user has to be in the database. So, the user will ask the smart assistant to book a room for meeting. In this face recognition and speech based interaction is needed. In general, we can think of smart assistant in Universities, School, and in home also. For example users interact with the smart assistant in ATM machine via speech based interaction. There are many smart assistant which is only speech based. Face recognition is one of the bio metric recognition which could also help whether the right person is booking the room or not. We would like to integrate the smart assistant with face recognition. Apart from this we have also found issues related to face recognition. Our motivation to post this question is to and how people think when smart assistant is integrated with face recognition. What is a Smart Assistant: a device/interface which assists in performing simple administrative tasks like booking rooms? The main aim for this is to explore to what extent smart assistant can be implemented.

Q1. Have you a ever used a smart assistant in your office?

1. a. If yes, how will you rate your experience with using a smart assistant in your office? Provide scale of 1-10

1. b. If no, do you think a smart assistant will help improve efficiency in your office? Agree, disagree, don’t know

Q2. What are the various ways in which you would like to identify yourself to the smart assistant? Face recognition; RFID based identity card; Passcode(multiple choice)
Q3. Have you ever had to identify yourself via face recognition interface? Yes, No, can’t recall.

Q4. Do you have any privacy concerns while using a face recognition interface to identify yourself? Scale 0 to 10. [0—highly unsafe; 10–safe.]

Q5. Do you have any security concerns while using a face recognition interface to identify yourself? Scale 0 to 10. [0—highly unsafe; 10–safe.]

Q6. What are the potential concerns when using face recognition interface?

Q7. Which of the following platforms would you prefer for a smart assistant? Please rate the following 1–high priority, 2–low priority. App based (on mobile devices); Web based (on a computer/laptop)

Q8. What does motivate you to prefer the platforms?

Q9. How would you like to interact with the smart assistant? Touch screen with options provided in different dropdown menus or Interactive voice response system (IVRS).

Q10. If the smart assistant uses a speech based interaction (IVRS), how would you like to end the conversation? OK; Thank you; Over and out; others—specify

Q11. Where do you think smart assistant technologies can work and what can be the problems?

Q12. Where or under what circumstances do you think that face recognition would be appropriate?

Q13. Where or under what circumstances do you think that voice UI would be appropriate?

Q14. What are the potential problems/issues related to the usage of speech based interaction with a smart assistant?
Appendix B - Survey Questions for Cybercom Employees

As interaction with computational technologies are becoming more and more complex the potential for digital assistants is very strong. Also, the possibilities for alternative interaction such as speech-based commands (Interactive voice response system), facial recognition, gestures etc has a strong potential while at the same time research questions around these technologies remains unclear. Cybercom is doing joint research on smart assistants together with the research centre Internet of Things and People at Malmö University and we are two thesis working students who are part of that research. It would be most valuable for us all if you could help the research out through answering some brief questions.

At Cybercom a smart meeting planner, using face detection for identifying users and with speech based interaction, has been implemented as a functioning prototype. The questions are related to use of that prototype and the use of the said interaction technologies. Please feel welcome to add comments to yes/no questions or 1-10 scale questions.

Thank you for your participation

Prakriti Dhang, (Master’s Program in Computer Science)
Manaswini Kolluru (Master’s Program in Computer Science)

Q1. Have you used the smart assistant in your office? [Yes, No]

a. If yes, how will you rate your experience with using a smart assistant in your office? Provide scale of 1-10 where 10 is the best experience.

b. If no, have you had any specific concerns about using it? [open answer]

Q2. If having used it; were you aware of being recognized through face recognition? [Yes, No]
2.a. Do you feel comfortable with such a way of being identified? Provide scale of 1-10 where 10 is the highest degree of comfortability.

Q3. If feeling uncomfortable; is this due to privacy concerns or other reason? [Yes, No or other answer]

3.a. If you could be granted anonymity in the booking despite the face recognition, would that change your degree of comfortability? [Yes, No]

Q4. Did you feel trust in that booking was made in your name? Provide scale of 1-10 where 10 is the highest degree of trust.

Q5. What was your experience of using speech-based interaction. Provide scale of 1-10 where 10 is the best experience.

Q6. Did you experience any frustrating moments while interacting with speech-based interaction? [Yes, No] If yes; please give some short comments on the frustrating factors? [open answer]

Q7. Would you prefer writing on a touch screen for the booking rather than speech-based interaction? [Yes, No]

Q8. How would you like to interact with the smart assistant? Touch screen with options provided in different dropdown menus, speech-based interaction, or a combination of both? [provide the alternative and add a comment if you like]

Q9. Did you feel trust in that the booking had been completed and that the “conversation was over”? Provide scale of 1-10 where 10 is the highest degree of trust.

9.a. If high trust; was it any specific feedback that made you realize that the booking was completed? [please provide a short comment]

Q10. Would you like to add other comments on using the meeting planner? [optional open answer]
Appendix C - Co-Design

**Different Scenarios** The different scenarios of the co-design session are described below. The idea is to have a smart assistant in the office for booking a meeting room for a registered user verbally. The office scenario below is the way it works in Cybercom. From this aspect, we explored many different contexts such as home, healthcare, and public space.

A. Dialogues for the Smart Assistant in Office Environment:

SA: Hello, May  
User: Hello SA  
SA: How may I help you?  
User: I want to book meeting room.  
SA: What time?  
User: Is it possible this afternoon at 14:00?  
SA: Sorry, no room is available at specified time.  
User: Is it available at 15:00?  
SA: Yes, it is available.  
User: Please book at 15:00  
SA: Your meeting room is booked at 15:00 and your room number is B1096  
User: Thanks  
SA: You are welcome.

B. Dialogues for the Smart Assistant in Home Environment:

You entered your house and your smart assistant starts talking with you.  
SA: Hi Riya! How was your day, you look very tired? Would you like to listen to any soft music? (In the device your face was visible)[As this is your own house, SA knows your name]  
Riya: It was fine. Yes, I would like to listen to a instrumental music.
C. Dialogues for the Smart Assistant in Health care Environment:

A patient is admitted to a hospital and has been assigned a room. There are several activities involving interaction between the patient and many other health workers in the hospital. For example, the patient gets timely medicine from a nurse; patient’s vitals are constantly monitored by a team of doctors/nurses and every once in while the patient gets visited by a senior doctor responsible for the patient. All these activities need special scheduling and almost all hospitals use scheduling tools specially tailor-made for their needs. We wish to introduce a smart assistant as an interface between the patient and the internal scheduling software of the hospital. Through the smart assistant, the patient will be able to find out the exact time for visits meals, medicines, and any other medical procedure. In the following scenario, we exhibit an instance of exchange between the patient and the smart assistant.

SA: Hello, Lora.
User: Hi, SA.
SA: How may I help you?
User: When is the doctor’s visit scheduled next?
SA: Doctor’s visit is scheduled between the 13:00-14:00.
User: Thanks for the information.

Many other services can also be integrated with the smart assistant like SOS, Housekeeping etc.

D. Dialogues for the Smart Assistant in Public Space:

You are roaming around a mall or walking near a bus stop. There is a smart assistant device and it suddenly starts talking. From this we can say that few may answer and few may not. It depends on person to person. Based on this scenario, we have related questions.

SA: Hi! How’s your day? (In the device you are identified by face and there are many people around) [What would be your reaction?]
Riya: It is good. [few may have ignore it and may not answer.]
SA: What did you buy? [If you were near mall] OR What is your destination? [If you were in the bus stop]

| Participant No.1 |  
| Contexts | Answers |
| SA in Office | Qualitites: reduce human effort Added things: increase the room if it is unused. |
| SA in Home | Good idea 2-3 time and after that find another solution Good qualities can recognize face and see your mood. |
| SA in Healthcare | Useful but risks health records if misused, misscheduling and wrong inject. Good qualities: helps in monitoring |
| SA in Public | Dont want to talk in public, too busy Risk: data stored. Added features: usefull in big mall can ask which shop should I go. I am hungry where should I go. only voice not face kids its in museum park or burger king this can help Navigation (if the phone is dead then it is good to have in public) |

| Participant No.2 |  
| Contexts | Answers |
| SA in Office | Even your employee recognize by face, it is good. Yes it is usefull in office. Added feature SA will itself will say next available time is at 15:00. Risks: mmm... no risks or my be my face or if it storing my face Qualities: it is communicating, hmmm, I think it is good, you dont need to type. |
| SA in Home | Maybe 50% good, Risks: if not turning off, if it is on always, and some one enters the house. I will control my SA, if others controls may be my data will be theft. Added features: multiple language. Actually it will be good for elder people |
| SA in Healthcare | Yes it is helpfull, in sick people room, not in all rooms, may who needs, limited to specific patients. Risks: may be SA can effect patients. Added features: patient can call for doctor through SA. Its good I like this one. |
| SA in Public | Not usefull, I will be more angry if it recognizes by face. Risks: my face, it is in public. Added features: for navigation. |
### Participant No.3

<table>
<thead>
<tr>
<th>Contexts</th>
<th>Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA in Office</td>
<td>Not comfortable with face recognition, annoying to show my face. risks: my face is stored</td>
</tr>
<tr>
<td>SA in Home</td>
<td>I feel like the SA in home is good idea and also comes with privacy risks. Qualities: make daily life easier, and you can focus on other tasks, can spend time with childrens. Risks: are mainly are the information, data that is stored in SA. It can be hacked by someone, and can used it for benifits and can hurt that person. People can be obsessed by SA, that they will stop doing normal things for themsevles. SA can also help for find this and that, these kind of SA is addicted also. And if you lost your SA, you will be also lost in finding somethings.</td>
</tr>
<tr>
<td>SA in Healthcare</td>
<td>Quality: usefull for well knowing the schedule, the patient know I will get meal in this time. If the patient needs something, patient can ask SA and SA can convey the information to the specific person. Risk: If SA gets stupid and tells wrong things to another person. It has to be more accurate, and must work perfectly. Added features: No limit to add new features, patient can tell SA to send a email to the specific person for info or bring me this medicine. Multiple language can be added, not all patient knows english.</td>
</tr>
<tr>
<td>SA in Public</td>
<td>I feel it is not good idea since everyone will be around you, know your name, what you bought, like in general not always good. For navigation it is good to have such SA and just ask SA for your way.</td>
</tr>
</tbody>
</table>

### Participant No.4

<table>
<thead>
<tr>
<th>Contexts</th>
<th>Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA in Office</td>
<td>Good qualities: its fine affiliated to that organization, added features: make relevence to context. make more interesting and interactive, background info whether the same person has booked anywhere else. risks: if all the room is book and how SA will help that person</td>
</tr>
<tr>
<td>SA in Home</td>
<td>For home it is ok. Risk: Dont think about the mood, and starts talking. Added features: identify the mood and then starts talking.</td>
</tr>
<tr>
<td>SA in Healthcare</td>
<td>Qualities: yes to have in hospital, Added feature: if patient have urgency instead of going to reception can ask SA happy that the doctor visit is near. in a particular time SA can provide info and questions or SA start first talking or vice versa. In remote areas no unified system, to go each and every hospital. Risks: seriousness of a person gives more convieint, supplementary things that can minimize the fear by identify. Insecurity: some diease dont want to speak out loud? Authentication is needed, yes I am this person.</td>
</tr>
<tr>
<td>SA in Public</td>
<td>Instead of saying identifying, it will ask is this you? As per GDPR face is not good idea. If there is only voice: then it is ok.</td>
</tr>
</tbody>
</table>
References


Master Thesis Project: Interaction with Smart Assistant

[22] https://www.apple.com/siri/


