RESEARCH ARTICLE

A Framework to Enhance User Experience of Older Adults With Speech-Based Intelligent Personal Assistants

MUHAMMAD USAMA ISLAM AND BEENISH MOALLA CHAUDHRY

School of Computing and Informatics, The University of Louisiana at Lafayette, Lafayette, LA 70504, USA

Corresponding author: Beenish Moalla Chaudhry (beenish.chaudhry@louisiana.edu)

ABSTRACT Speech-based intelligent personal assistants (sIPAs) promise to improve quality of life in older adults, but they pose various usability barriers that limit their adoption by older adults. We conducted a semi-structured interview study with fourteen older adults to understand their experiences with these devices. The collected data were analyzed using inductive and deductive coding, resulting in the identification of two broad themes: usage of sIPA and concerns regarding sIPA. "Usage of sIPA" highlights different ways in which participants were currently using and wanted to use their sIPAs in the future. "Concerns regarding sIPA" explains different types of usability challenges that participants were facing with these devices. Based on our findings, we suggest that sIPAs for older adults should focus on privacy improvements, interpersonal skills and contextual awareness. In addition, we provide practical suggestions for implementing permission-based data storage, explainable artificial intelligence (XAI) principles, dialect and accent recognition, and humanized communication behaviors within sIPAs. This research provides both design and implementation directions to accelerate improvements in sIPAs aimed at older adults.

INDEX TERMS Older adults, virtual assistants, speech based intelligent personal assistants, usability, design, implementation.

I. INTRODUCTION

Speech-based Intelligent Personal Assistants (sIPAs), also known as Voice Assistants (VAs), have gained immense popularity in recent years. A sIPA or a VA can be described as a hardware or software agent that is powered by artificial intelligence and uses input such as the user's voice and contextual information to assist by answering questions, making recommendations, and performing actions using natural language in a spoken format [1]. Amazon Alexa, Google Nest Mini, Apple HomePod, Apple Siri, and Microsoft Cortana are a few examples of popular sIPAs in the market today [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino.

There is an assumption that, compared to other technology devices, sIPAs might be more acceptable to older adults, due to the device's natural interaction style, i.e. speech [3]. The empirical evidence for this claim, however, is still quite scant. Indeed, lack of technical skills is commonly reported by older adults, increasing the likelihood of low adoption of any technology in this age group. In a recent Pew Survey (2017), 34% of the surveyed older adults (65+ years) indicated that they have little to no confidence in learning to use a new technology device on their own, while 73% stated that they would prefer to learn to use one with someone's help. In the case of sIPAs, older adults still have to learn what to say and how to speak with the device, which might pose a barrier if this information is not readily available, or deviates from older adults' expectations [4].

The field of human computer interaction (HCI) [5] is concerned with making interactive systems easy to use and useful so that they are accepted and adopted by the target users [6], [7]. This requires an explicit focus on user interface design, user experience expectations [8], [9] and usability [10], [11], [12], [13], [14]. These areas have become popular in HCI with the emergence of the user-centered design approach, which, according to the Usability.Gov website, is "based upon an explicit understanding of users, tasks, and environments; is driven and refined by user-centered evaluation; and addresses the whole user experience". Given that old age is often accompanied with various physical and cognitive limitations and may also correspond with low technical skills, user-centered design approach mandates incorporating an explicit understanding of older users' point of view in the design and deployment of any technology product that is geared towards them [15]. Therefore, acceptance and adoption of sIPA in older adults cannot be determined without an explicit investigation of the views of this user group [16], [17], [18], [19].

Our literature review revealed that the HCI research in sIPAs and older adults has so far encompassed two main domains. The first domain covers specific uses of sIPAs by older adults, which includes entertainment, health management, reminder generation, control of home devices, and to some extent, companionship [20], [21]. These works generate expectations from sIPAs in promoting well-being of older adults in everyday (non-smart) and smart living environments, necessitating further research in improving usability and user experience of sIPAs for older adults. The second domain includes studies that describe adoption challenges of older adults with sIPAs such as dissatisfaction with interpretation of complex questions [22], understanding of diverse dialects and accents [23], security concerns and negative notions about artificial intelligence [24].

Overall, the existing work indicates that sIPAs are yet to meet expectations of older adults, especially in terms of learnability, security and conversation capabilities. Further investigation is needed for generalization of existing findings and to identify additional issues and concerns that can, ultimately, be addressed to improve acceptance and adoption of sIPAs in older adults. This is desirable because sIPAs have immense potential to support aging-in-place, lower the caregiver burden and improve older adults' quality of life [15], [20], [21], [25].

Based on this justification, we seek to advance the current research landscape by understanding usability of voice assistants from the perspective of older adults. In particular, in this work, we investigated older adults' experiences with different aspects of sIPAs, particularly effectiveness, and utility using a qualitative approach. The data was analyzed using an exploratory thematic analysis to propose a future research framework aimed at improving sIPAs according to older adults' perspective. The contributions of this work are two fold. First, we present design implications of our findings to inform future work. Second, we present a theoretical research framework that discusses implementation details using recent advances in the computing field.

The remainder of the paper is organized as follows. Section 2 provides an overview of related literature. The next section describes the study in terms of materials, methods, design, analysis, findings and discussion. In section 4, we propose a research framework based on the study findings and implications. We conclude the paper by presenting study and framework limitations, and future research directions.

II. RELATED WORKS
A. OLDER ADULTS AND VOICE ASSISTANTS
There is a rising interest in exploring how older adults perceive, interact with, and integrate sIPAs into their everyday lives. Kim [23] conducted task-based semi-structured interviews with older adults (74+ years) who had never used a virtual assistant. The authors note that VAs have limited follow-up natural conversational capabilities, whereas older adults in the study followed up their interactions with VAs with polite responses. The author recommend including common human conversation styles in VA's conversational capabilities so users are able to easily understand and make follow-up human-like conversations. Shalini et al. [26] echo the need to improve humanness of VAs. They compared various voice-assistants, including Amazon echo and Google assistant, to accomplish the objective of developing an easy to use health interface for older adults. They report that automated health alerts from these systems resonate well with the older adults but these alerts need to be more humanized.

In another study, Kim and Choudhury [27] conducted a sixteen week long longitudinal study to understand how older adults' perceptions towards VAs change as they progress from being novice to advanced users. The authors report that, over time, participants changed their behavior from simply enjoying the speech based interaction to consciously learning what to say to relaxing and appreciating the simplicity and freedom of the interaction modality. Many participants appreciated simple conversations with the device and found them effective for building companionship. The authors discuss some ideas around enhancing conversational capabilities of VAs to support older adults' companionship building behaviors with the device. Kowalski et al. [28] conducted two co-design workshops with older adults to brainstorm potential uses of VAs, barriers to use and solutions to address those barriers. Similar to Kim and Choudhury, they also received, an overall, positive feedback, about VAs, from the participants. Moreover, suggestions made by the older adults indicated that VAs need to expand their sensing and feedback capabilities to become more user friendly and useful for older adults.

Corbette et al. [25] report that older adults struggle with learning new technology based on a user study carried out for 60 days. The issue of learnability of VA has also been raised by Pradhan et. al. [29], who also mention speech non-recognition. Combined, the researchers recommend modifications to VA designs and conversation

capabilities for promoting learnability of the technology. Hanley and Azenkot [30] report that older adults were able to master VA use by attending formal technology training workshops and informal support groups.

Bonilla et al. [24] reiterate privacy, credibility, and security concerns raised by Kim [23] and Pradhan et al. [29], and also added reliability and credibility of information as a concern pointed out by older adults. Similarly, Nallam et al. [21] found that low-income older adults had concerns about confidentiality of personal data and receiving information that is trustworthy. These works show that the lack of knowledge, in addition to fake knowledge and misinformation about artificial intelligence takeover [31], have generated disinterest among older adults about using sIPAs. The authors recommend pursuing objective teaching of the subject matter to prospective users for an informed understanding of sIPAs, which would lead to informed choice of using a VA [32], [33].

Trajkova and Martin-Hammond [34] explored why older adults would like to use and avoid Alexa. They noted an interesting pattern aligned with [23], [24], and [35] regarding concerns on usable cases after a year of qualitative research with older adults. The findings show that older adults perceive Alexa as a toy companion that can decipher simple instructions such as playing music but malfunctions when it comes to sophisticated use cases for assisted living and well-being. In contrast, Nallam et.al [21] explored the perceived benefits of VA usage by low-income older adults. The authors commented that participants were inclined to search health related information using VAs. Corbette et al. [25], O'Brien et al. [20] and Hanley and Azenkot [30] also found that older adults and their support personnel see tremendous benefit of this technology in terms of supporting aging in place, providing companionship and lowering caregiver burden. The participants in their study made several recommendations around enhancing caregiving [36] and health knowledge providing capability of VAs [37]. The latter two studies were, however, based on short-term exploration of VAs by older adults.

B. OTHER VIRTUAL ASSISTANT STUDIES

Besides older adults, researchers have also investigated challenges of voice-based interactions with general users. The main areas of investigation have centered around understanding voice recognition, contextual understanding, and interaction capabilities of sIPA. For example, Tulshan and Dhage [2] compared various voice assistants and concluded that Google assistants excelled in voice recognition capability and hand free interactions, followed by Siri, while Cortana and Alexa had substantial room for improvement.

Chung and Lee [38] showed that the data collected by sIPAs can be used to generate many insights about a user's lifestyle such as sIPA usage patterns, user's interests and user's sleeping/wake-up patterns. Needless to say, this has many privacy and security implications, such as Liao and his team [39] showed that privacy and trust issues can impact

TABLE 1. HCI components explored in this study.

Usability Component	Short Description
Utility [53]	How practical and useful an object is.
	Usefulness of a function.
	Object can reach pre-defined goal/ task.
	Functionality of a system.
Effectiveness [54], [55]	Issues in completing task or action accurately.
	Component of usability.
User Experience [56]	User's overall experience and satisfaction at interaction time.
	User experience encapsulates emotional reactions to usability.
	Investigates the issue of emotional fulfillment.

user's decision to adopt sIPAs. Lau and his colleagues [40], however, showed that users had an incomplete understanding of privacy issues. Chung and his team [41] further explain that user's data in sIPA's cloud servers can be compromised by malicious sources and software. Chiu and their colleagues [42] pointed out the importance of emotionally aware personal assistants and recommended a neural network approach to develop an emotionally aware assistant.

Katharina Weitz and her team [43] focused on integrating interaction based design with explainable artificial intelligence (XAI) and found that it led to a significant improvement in user's trust. A similar research work was carried out in [44], where authors experimented with voice enabled assistant to disseminate the explanations of AI decisions. Exploration in this area was further carried out by [45], where the authors report potential of XAI in addressing explainability issues of various recommendations. Dialect and accent understanding is another active area of research in speech-based interactions [23]. Several researchers have proposed improved architectures to improve this capability of VAs for users. For example, Matani et al. [46] propose an additional layer of dialect and accent understanding. The users with accents were studied in [47] to understand the usability, acceptability, and satisfaction of the participants with sIPAs. Similar research work was carried out in [48], where researchers found low effectiveness of sIPAs in accent recognition.

Other applications of XAI in speech-based systems include Ahmed et al.'s work [49], in which authors utilized natural language processing to propose a tool for psychologists to use in mental health studies. Similar research work was done for hate speech by the Ahmed et al. [50], where the authors paired explainability and active learning. Furthermore, explainable sentimental analytics for mental disorders were performed in [51]. Djenouri et.al [52] explored amalgamation of IoT and XAI for tangible supervised learning that enhanced the body of research remarkably in this field.

III. STUDY

The aim of this study was to investigate older adults' experiences with sIPA from their own perspectives. We used a qualitative approach for the study because it uses techniques like open interviews, focus groups, or questionnaires for gathering information and subjective experiences of the users, after they have used the target product/system. The main advantage of

TABLE 2. Study's interview protocol.

Category	Interview Questions
Utility	How do you presently use your voice assistant?
	How would you like to use your voice assistant in the future?
Effectiveness	What challenges have you encountered while using the voice assistant?
User Experience	Do you have any concerns or fears about using your voice assistant?
	What are your thoughts about using a voice assistant in the future?
	What kind of impact has the voice assistant had on your life?

this data collection method is that it allows researchers to obtain rich information and insights into different aspects of the product/system under investigation that cannot be predicted beforehand.

A. METHODS

We obtained approval to conduct this study from the institutional review board of the University of Louisiana at Lafayette. The inclusion criteria of the study were being ≥65 years, and owning and using (or owned and used) a sIPA. Participants were excluded if they could/did not (want to) meet via phone or a video conferencing tool such as Zoom or Skype, or if they had a (self-perceived) cognitive impairment. The recruitment was done via online advertisements on various social media platforms and personal referrals. The online advertisement clarified the eligibility criteria, and encouraged interested individuals to contact the principal investigator via email for further information.

Twenty-six participants emailed us to learn more about the interview. However, only fourteen participants ended up scheduling a study session. All the interviews were completed online via Zoom, with one researcher conducting the interview. Each interview was approximately 60 minutes in length and was recorded for later reference and analysis. At the conclusion of the interview, participants were compensated $20 via PayPal.

B. STUDY DESIGN

The interview started with a brief description of the study. We then collected participants' demographics using an online Qualtrics survey link. After this, we began the actual interview, whose goal was to understand older adults' adoption experiences and usability challenges with their sIPAs. The interview protocol consisted of some pre-decided questions in certain categories (Table 1) but participants were probed further in each category according to their responses. The Utility category was explored further to gain a deeper understanding of the specific uses of sIPAs mentioned by participants. The Effectiveness category was explored further to understand the usability issues. The User Experience was investigated

by probing participants' personal feelings, emotions and attitudes towards sIPA based on their experiences and specific uses. Table 2 lists the main interview questions.

C. PARTICIPANTS

The study was conducted with fourteen (females = 10, males = 3 and non-binary = 1) participants with an average age 70.2 years (standard deviation = 4.4 years). Nine participants identified themselves as European Americans, one as Middle Eastern American, three as African Americans, and one as a Southeast Asian American. Everyone was living in the United States. Five participants lived alone, three lived with children, and six lived with significant others. In terms of education, eight participants had advanced degrees, four had undergraduate degrees, and the remaining two held high school diplomas. The demographic information of participants is summarized in Table 3.

In Table 4, we have summarized what kind of sIPA participants owned, how long they owned it for, and usage frequency. Everyone owned at least one VA and had used one within the last six months. Ten participants were actively using their sIPAs at the time of the study, while four had limited their usage of the device, or not used one for the past few months. Eight participants also used VAs on their smartphones along with their sIPA. The most popular smart speaker was Amazon Alexa (n = 11) and the most popular smartphone VA was Apple Siri (n = 8). Two participants rated themselves as novice users of the smart devices, and eight indicated that they would require help to use their devices more effectively. The remaining considered themselves expert users. On average, participants had been using their sIPAs for 3.1 years (standard deviation = 1.6 years).

Six participants owned more one VA, explaining that they were keeping VAs in different rooms but using them all for the same purposes. Everyone had placed at least one VA in the living room, hallway, kitchen, or any other area they considered as the main area in their house. The reason for choosing this location was convenience and accessibility for everyone in the house. One participant indicated that she was planning to purchase more VAs to put in other rooms.

Four participants had received their sIPAs as gifts from their family members, one participant had won his in a raffle, and the remaining had purchased on their own for different reasons. For example, one participant had purchased her sIPA to help her with post-surgery recovery, another was inspired by her friends, another liked trying out new gadgets, etc.

D. DATA ANALYSIS

We performed thematic coding of the interview transcripts [57] by applying six phases of manual coding (inductive and deductive) [58]. Two researchers were involved in the coding process, which included text interpretation, comparison with other codes, and, finally, clustering similar codes into categories, which were then linked back to the aspects of usability that were investigated in our interviews [57].

TABLE 3. Demographic survey questions and responses.

Demographic Information	Answer Choices (number of responses)
Please indicate your age range (Age)	65-70 (n=7)
	71-75 (n=6)
	76-80 (n=1)
Please indicate your gender. (Gender)	Female (n=9)
	Male (n=4)
	Non-binary (n=1)
Please indicate your race. (Race)	African-American (n=3)
	Asian (n=1)
	European American (n=9)
	Middle Eastern American (n=1)
What is the highest degree or level of education you have completed? (Education)	High School diploma (n=1)
	Some college experience (n=1)
	Bachelor's degree (n=4)
	Master's degree (n=6)
	PhD or higher degree (n=2)
What is your annual household income? (Income)	Less than 25k (n=1)
	25k-50k (n=2)
	50k-100k (n=7)
	More than 100k (n=3)
	Retired (Dependent on Savings) (n=1)
What is your living situation? (Living)	Living alone (n=5)
	Living with children (n=3)
	Living with a significant other (n=6)
How would you rate your ability to use a smartphone? (Ability-Smartphone)	Beginner (n=1)
	Intermediate (n=4)
	Expert (n=9)

Before we justify our choice of thematic coding and analysis, we report that there are several qualitative analysis methods that are available at our disposal. This includes content analysis, narrative analysis, grounded theory analysis and discourse analysis [59]. Thematic coding was the best choice for our study because we wanted to map our findings back to the explored usability components. Content analysis and narrative analysis are best choices when the intention is purely constructive, i.e. understanding experiences from interviewers' perspectives. Grounded theory and discourse analysis are more suitable for qualitative work that is scalable and longitudinal in nature. Hence, we opted for thematic analysis. The detailed aspects of thematic analysis with relation to our data is discussed herewith to provide a distinct understanding of data analysis that has been done by the researchers.

1) INDUCTIVE AND DEDUCTIVE CODING
Inductive coding can be defined as a form of bottom-up or ground-up approach where we try to find the codes by diving deep into data. In this case, we don't have any preconceived ideas of what the codes would be and thus allow the narratives to emerge from the raw data transcript. This is beneficial in performing exploratory research when we aspire to explore new theories, ideas, or concepts [58]. Deductive coding is a top-down method where we start by having preconceived codes in our mind which we materialize into a codebook with an initial set of codes. This set of codes is usually based on interview research questions or it might be a pre-existing framework of the researched theory. The next step is to read

through the data and take important excerpts to assign to codes [58].

We used both inductive and deductive coding to analyze the interview data. This is a normal practice in the qualitative research domain [58], where deductive strategies are utilized to explain causal relationships between concepts and variables, and inductive strategies to understand what is happening in the data, freeing the process from the researcher's preconceived notions.

2) PHASE 1: DATA FAMILIARIZATION
The first step of thematic coding is familiarizing oneself with the data that has been curated. For our analysis, this step included transcribing the interview, reading the whole interview to see if any code emerges or not. For example, P1, at certain point, when asked, "How do you presently use your voice assistant", replied, "Use it twice per day. When I am eating lunch. To enjoy lunch with music or podcast". In data familiarization, for this example, we map the question back to "Utility" aspect, and the use of the word "presently" in the question lead us to create "Current Usage" sub-aspect of "Utility" aspect. Similarly, the answer gives us an insight into three specific questions, a) Frequency of use, b) when it is mainly used, c) what it is used for, we note this information for our next phase.

3) PHASE 2: INITIAL CODE GENERATION
In this step, we generate first-order codes from the first glimpses of data. So, we start to code interesting features of the data in a systematic fashion across the data-set to identify

and collate words that can be associated with each code. For the earlier example, at first, we try to find interesting features of the data based on phase 1 questions and generate some codes. Apart from "music" as a code, we find that, other participants discussed listening to music with several words other than music such as, songs, opera, tune, tracks, pop, jazz that are all referenced as a single code, i.e. "music".

4) PHASE 3: SEARCHING FOR THEMES

In this step, we decide what themes are emerging from the generated codes. For our example, we categorize, the codes, "music", "podcast", "news", "personal search", "weather" with a working (first-order) theme named "leisure activities". After some loops of back and forth between phase 1 and phase 2, we finally named the theme "Activities" which can be categorized as second-degree theme.

5) PHASE 4: REVIEWING THE THEMES

In this step, we start checking all the codes have been covered under the identified themes or additional themes can be generated. The goal is to generate a thematic map where the themes are mapped to codes and each code might belong to one or more theme or a different theme altogether. In our example, we see that the code "news" and "personal search" do not conform to mere enjoyment or entertainment so we again go back to loops in phase 1 through 4 to create another second degree theme named "seeking information", which maps perfectly to our codes.

6) PHASE 5: DEFINING AND NAMING THEMES

This step is about making the somewhat arbitrary first or second order themes from phases 3 and 4 more concrete, so that the labels become more self-explanatory. For example, "Entertainment" was generated as a distinct theme from somewhat arbitrary second-order theme, "Activities" within the sub-aspect "Current Usage". This allows us to clearly see that "Older adults in our study are "Currently" using VAs as a source of "Entertainment".

7) PHASE 6: PRODUCING THE REPORT

The thematic analysis recommends using appropriate extracts as the final opportunity for and a discussion of analysis that would relate us back to research questions, or literature review. Our findings and discussion sections utilize phase 6 guidance to deliver proper insights on the thematic analysis.

As depicted in Table 6, we have analyzed the utility aspect to emerge into seven different themes, namely, entertainment, companionship, cognitive aid, communication, healthcare uses, home control, and information, with respect to sub-aspect current usage. The current usage sub-aspect shows that, in the entertainment theme, music is often enjoyed through VAs by older adults as well as weather, news, and information acquisition. The insight from these codes was that even though older adults would use a VA for medication reminders, a very few use would use the VA for that purpose

(code frequency is only 3). However, VA is mainly used for generic reminders and alarms, with combined frequency 10. As noted in the future usage sub-aspect, the users aspired to customize VA's voices, have more human-like casual emotional companionship-type conversations with their VAs, and receive automated insights and reminders from VAs. The challenges faced were mainly in terms of VA's inability to recognize speech (code frequency: 7), and privacy issues (code frequency: 13) that led the users to abandon and/or limit the usage of sIPAs.

The implications theme Table 7 encompasses better privacy settings (code frequency: 6) and permission-based recording (code frequency: 5). These codes generated ideas for addressing security concern required to be addressed by future research. Another relevant code of this category is improved speech recognition (code frequency: 11) and human-like conversation (code frequency: 11), respectively. The implications theme, overall, allowed us to brainstorm design and implementation frameworks to improve user experience of older adults with sIPAs.

E. FINDINGS

The qualitative data analysis resulted in two broad themes: usage and concerns. We discuss the main categories and sub-themes below:

1) USAGE OF sIPA

We identified seven distinct uses of sIPA by participants that have been summarized in Table 6. However, not everyone had used their sIPA for each identified purpose (Table 5). On average, each participant could perform at least three out of the seven different types of tasks we had identified. We describe the main usages of sIPA below.

a: ENTERTAINMENT

Reflecting on their experiences with sIPAs, participants mentioned that, overall, they had found their sIPA to be very helpful, useful and easy to use. Everyone who participated in the interview indicated that major reasons for owning a VA was convenience and fun.

> "It is a luxury. It is going to improve life, whoever will own it" - P3.

Eight participants indicated that they mainly used sIPA for entertainment, such as music and podcasts. Participants also indicated that sIPA was a very useful device to entertain guests. About four participants mentioned hearing jokes or playing games with sIPA when guests were visiting. P1's comment summarizes this nicely:

> "When we are eating or cooking dinner and lunch, we ask it to play music or news or podcast. Sometimes when we have company, we use it for music or jokes" - P1.

Other participants played with their sIPAs when they were alone by themselves. For example, P12 explained that he found it entertaining to ask sIPA certain questions about

TABLE 4. Details about VA ownership by participants.

Participant ID	VA Type	Ownership Time	Usage Frequency	Ways of Using
P1	Siri, HomePod	6 months	2 times/day	5
P2	Siri, Alexa	1 year	4-5 times/day	3
P3	3 Alexas	1.5 years	almost every time	5
P4	Alexa, Show, HomePod, Google Home	2 years	several times/day	4
P5	5 Alexas	3 years	almost every time	4
P6	Siri, Alexa, Google Assistant	4 years	1-2 times/day	4
P7	Alexa	4-5 years	2-3 times/day	3
P8	2 Alexas	5 years	3-4 times/day	3
P9	Alexa	4 years	3 times/week	2
P10	Alexa	5 years	several times/day	4
P11	Alexa	6 months	2-3 times/day	2
P12	Alexa	2 years	almost every time	3
P13	Echo Dot and Show	3 years	several times/day	4
P14	Alexa and Echo Dot	4-5 years	several times/day	3

herself. P6 asked her sIPA to play cat sounds for her when she was feeling down. Some participants mentioned that they liked switching voices on sIPA because listening to different accents was fun for them.

> "I sometimes will switch to another accent. I have used British and Australian accents, just for the fun of it" - P5.

Overall, participants looked at their VAs both as an entertainment device and a toy that offered a few tweakable features to keep the user engaged.

b: COMPANIONSHIP

Some of the participants hinted that they had formed close bonds with their VAs and often talked to them when feeling lonely. At least 6 respondents alluded to this type of usage. The following conversation is noted when P3 was asked to describe her relationship with her sIPA.

> "I call it out while sitting on my bed. I ask it to put the lamp on. I ask it what time is it? It is helpful. I ask it what the weather is like right now? Or what is going to be like? How much snow is it expecting? How long will it snow? It is always there. Sometimes I fall asleep, and I do not realize what time it is when I wake up. When you are old, you do not understand how time is passing. Alexa is helpful" - P3.

Participants disliked the rigid and rugged conversational tone of sIPA, and thought that it would be nice to introduce more emotions and human-like attitude. Specifically, P3, P8 and P12 thought that sIPAs should be more courteous and should allow their users to be more courteous as well.

> "I am a good person. I want to be able to say thank you. I mean, right now it is just loaded with volumes of information. Some characteristics it can have, it should have a good personality in the voice" - P8.
>
> "P12: She is too straightforward, very to the book. A little humor should be a part of it.

> Interviewer: So, you are saying that future Alexas should be more humorous?
> P12: Be more human, so to speak".

When asked what type of voice (male versus female) they would prefer to listen to, most participants did not indicate a preference. They indicated that whoever is talking, the sound should be clear and easy to understand. However, two female participants indicated that they would prefer to have female voices because they believed in female empowerment. One female participant had switched to male sound because she wanted her sIPA to be different from her friends'. Two male participants also preferred female voices but no one had preference for male voice.

c: COGNITIVE AID

Participants referred to their sIPAs as their "second brain". Six participants indicated that they had used sIPAs to receive different kinds of reminders ($n = 6$) such as paying bills, and receiving meeting reminders. Everyone concurred that using sIPA to receive reminders for certain things was a very good idea. Some participants acknowledged that their reliance on sIPA would increase as they would get older and become more forgetful.

> "As we are getting older, our memories are not as good, okay, and so you do not really have to try and remember all these facts, all these things. You can just go on to a voice activated device, and it will give you the answer so you are not under so much pressure to remember things" - P7.

Two participants were also using their sIPAs to create shopping lists and placing online orders. Four participants also indicated that they had solicited sIPA's help to obtain meal suggestions by sharing the content of their refrigerators. Participants thought that sIPA could be a very useful to supporting everyday cooking.

A few participants indicated that in the future they would like VAs to automatically detect their needs. For example, participants thought that VAs should be able detect what items

were running low or missing from the refrigerator and then, consequently, order them online.

> "It might be nice if it came up with ideas without me having to tell it" - P2.

Overall, participants thought that sIPAs should have more intimate knowledge of their needs and wants, and it should be able to accomplish a lot independently.

d: HOME CONTROL

Seven participants indicated that they had used a sIPA to control smart home features such as switching on and off appliances and lights. Some thought that using sIPA to control home devices was not a necessity, but everyone agreed that it was a great way to increase everyday convenience.

> "I have 19 connections to my home Internet and about 13-14 are Alexa and basic devices" - P5.

> "It is very beneficial, when I come into my room to turn the lights on from several feet away, so I do not run into furniture and things like that, so I find it a great little tool" - P13.

P6 had recently purchased a robotic vacuum and wanted to connect it to her Alexa.

> "I just got the floor cleaner, it is a robot, but you can pair it with Alexa and you can have her do all kinds of things" - P6.

Nine participants, who were not currently using their sIPAs for home device control, were very excited to learn more about this feature, and indicated that they would be interested in implementing it in the future.

> "In the future, my house is enough for me and what it does is fine. I think, actually, I would like to know more about what it can do that I may not be accessing right now" - P7.

Overall, participants thought that, in the future, they would prefer a VA that was able to do more household chores.

e: INFORMATION

Nine participants told us that they had used VAs for learning new information, be it news, stock market, weather, general knowledge, or word spellings. Four participants also indicated that they had used sIPA to search for recipes. Information seeking via sIPAs was an integral part of some of the participants' lives as summarized by P4.

> "Sometimes I need help with finding an answer to something and I go to it for help. Sometimes I am having conversation with my husband, and I have questions and I go to the device on the counter. It is like a little encyclopedia" - P4.

Participants indicated that they tend to avoid using their sIPAs for information purposes when they were with their friends or guests. Overall, participants thought that sIPA was very efficient with performing online searches and they would prefer to use it over doing the search online by themselves. There was a recognition among participants that their reliance

on sIPAs is going to increase with age. They also believed that sIPAs can help them stay connected to the world and age successfully.

f: COMMUNICATION

A few participants had also used their VAs as an intercom to communicate with other people. For example, P5 was using it to communicate with his wife during the day when she was in another part of the house, working in her art room. Also, P5 received calls on his Alexa from their disabled step-daughter, who lived in an assisted living situation. He thought that the Alexa helped his step-daughter communicate her needs when she was unable to get sufficient help from people in the assisted living facility.

> "And then we have got a two level house. The first Alexa is upstairs, there is one in the living room, that is the second one, and the third one is in my wife's art room and that is in there, because she sometimes is in there with the door closed" - P5.

Similarly, P3 was living upstairs in her daughter's house and was regularly using her VA to communicate with her daughter.

> "Because it is a great way to contact each other. My daughter can ask me do you want me to bring something. It is useful. It has been worthwhile all this time I have had it" - P3.

Overall, participants noted the hands-free and accessible nature of sIPA favored its use as an intercom. P13 had bought Amazon Show for her 92 year old mother and herself. She thought that operating buttons was challenging and burdensome for her mother but making calls with sIPA was simple and natural. Moreover, they could also see each other on the Show's screen. Everyone thought using VAs as intercom would help them greatly, if they were to become less mobile with age.

g: HEALTHCARE USES

A few participants specifically mentioned that VAs could be very useful for promoting self-care. A few participants ($n = 3$) were using their VAs for medication reminders.

> "I ask Alexa to remind me to take my medications, set appointments, drink water, to exercise" - P8.

There was, however, some differences in terms of using VAs for medication reminders, as some participants completely relied on their sIPAs while others preferred other compensation strategies such as a pill-box.

> "I do not use Alexa for medication reminders, I have a pill box and it has morning and afternoon boxes for my diabetes pills" - P6.

Participants mentioned that VAs should be designed to promote other healthcare and self-care usages in the future. For example, P8 exercised every morning and she thought it would be nice, if Alexa were giving her specific directions. She also mentioned that it might also be used to leave instruc-

TABLE 5. Theme-based usage of sIPAs by participants.

Participant ID	Entertainment	Companionship	Healthcare	Home Control	Information	Cognitive Aid	Communication
P1	X				X	X	X
P2					X		X
P3		X		X	X	X	X
P4		X		X	X	X	
P5	X	X		X			X
P6	X		X		X	X	
P7	X				X	X	
P8	X		X		X		
P9			X	X			
P10	X			X	X	X	
P11		X		X			
P12	X	X		X	X		
P13	X	X	X				X
P14	X		X			X	
n=14	9	6	5	7	9	7	5

tions for caregivers, who might be taking care of the elderly at specific times of the day.

2) CONCERNS REGARDING sIPA
The interviews followed by thematic coding indicated that participants had certain concerns related to privacy, technology behaviors, accent recognition, and complex query processing.

a: PRIVACY
Participants expressed concerns about VAs recording their conversations and invading their privacy. Everyone stated that they did not know that VAs were capable of recording conversations, when they first started to use their VAs. They learned about this through various means such as family, friends and articles. A few participants were not completely certain that a VA could record and analyze their conversations. But, they indicated that if this were true, then they would be concerned about others listening in to what they are saying and doing.

> "No, the only concern I have about Alexa is um …if she is recording conversations and I do not know that they say that she does, but I do not know that for a fact" - P8.

Four participants had considerably reduced sIPA usage because of privacy concerns, with two completely abandoning or putting sIPA out of their minds after a few months of use. P9 considerably limited use of his sIPA when he learned that it might be recording and monitoring his conversations.

> "I just feel like it is my privacy and I do not feel like a computer system or, you know, Alexa or the manufacturer or software company needs to know everything I see, do and say. I do not want it to record it in that regard" - P9.

Others were questioning and did not know what would happen if someone were to listen to their conversations. Some participants claimed that they do not have a very private life or they do not have anything to hide, therefore, the thought of someone listening to their conversation was not exceedingly worrisome.

> "If it is listening to what I am saying, like private conversations, and I am concerned about it I will just unplug it or turn it off" - P1.

A few participants had specific reasons that lead them to be concerned about their privacy. For example, P4 did not want to have a sIPA in her bedroom because she was concerned that she might say something embarrassing at night that would get broadcasted to her friends. P10 was concerned that she might say her passwords out loud or unwittingly reveal her bank account information to sIPA. A few participants had actually experienced problems with sIPA listening in to their conversation. For example, P6 recalled the following incident:

> "One time I was talking to my sister over the speaker phone. And she told me something like she has ordered a necklace from some jewelry store. A few days later, I received a notification that my order from that jewelry store is arriving in a few days. Alexa had placed the order for me without me knowing about it. When I found out, I had to go in and cancel it" - P6.

Another participants shared the following:

> "I read an article that someone was saying something about her mother-in-law that Alexa recorded and sent to her mother-in-law. I thought it was interesting" - P13.

Participants hinted that they needed clarity about privacy settings on sIPA, and whether they were sufficient to have safe interactions with the device.

> "Privacy is so critical. I will feel so secure if I can learn the privacy settings for HomePod. I am under the assumption that whatever privacy setting are on my iPhone, it is translated to HomePod but I need to be sure" - P1.

When asked whether explicitly controlling the recording permission on sIPA would alleviate concerns about privacy,

TABLE 6. Sub-themes, categories, codes and code frequencies associated with "Usage" theme.

Sub-Theme	Category	Code	Code Frequency
Current Usage	Entertainment	Music	7
		Podcast	2
		Weather	6
	Companionship	Daily Companion	4
		Conversation	4
	Cognitive Aid	Medication Reminder	3
		Alarm	4
		Generic Reminder	6
	Home Control	Home Devices Command	7
	Information	News	6
		Personal Search	8
	Health Care	Healthcare Reminders	2
	Communication	Disabled Calling	2
		Hands-free	2
Future Usage	Entertainment	Music, Podcast, and Weather	7
	Companionship	Human-Like Conversation	3
		Emotional Companionship	4
		Custom Voices	2
	Home Control	Additional Devices	3
	Information	Unbiased Recommendation	2
	Cognitive Aid	Automated Reminder	1
		Insightful Intelligent Reminder	1
	Healthcare	Health Coaching	2
		Caregiver Instructions	1
		Personalized Reminders	1

everyone responded with an affirmative. They thought that this would increase their trust in the device and lessen their fears around using sIPA.

b: BIASED INFORMATION

A few participants had witnessed some unexpected behaviors from their sIPAs. For example, P4 and P13 mentioned that, their sIPA often started talking without any prompt and they could not understand why their sIPA was providing them certain information.

> "One time, she just started talking about the stock market. I did not even ask it to give me that information" - P4.

P9 thought that this was sIPA's way of doing targeted advertising. He suspected that the information provided by sIPA is biased, and, hence, was hesitant to trust it. He mentioned that he did not appreciate targeted digital marketing, in general, because he felt it was trapping him inside a bubble and not exposing him to the broader, unbiased content.

> "I will give you a current example. I am looking to buy a new set of golf clubs and I want unbiased opinions and because I have searched for a couple of brands. Every time I search or try something different, it always takes me back to the same things that I already had looked at. So I am not sure if it is telling me those things because I already researched it and got information about, or is it taking me to these brands because it is the right way to go" - P9.

In essence, it appeared as if participants wanted greater clarity around how the device operated as well as assurances that the information it generated was unbiased.

c: OUT-OF-CONTEXT BEHAVIOR

A few participants also mentioned that Alexa was too sensitive and wakes up simply on hearing the mention of her name. Some participants even refrained from mentioning their VAs during our interview because they were afraid of their VAs becoming activated and breaking the interview flow.

> "Even if I whisper, it can listen" - P1.

P11 suggested that VAs should have "extra sensory perception" (ESP) and should be able to understand the context better. Participants wanted to avoid having repeated interactions with the device to make it understand their needs.

> "I have heated debate with Alexa. She does not understand" - P6.

While some participants mentioned that VAs should have some out-of-context functionality to increase its visibility in their lives. P10 mentioned that she often forgot to use her VA and thought that it would be nice if the VA could occasionally announce its presence.

> "Sometimes, I do not think of it, I do not think it is there anymore. I just like to forget that Oh, I have that. It can have some flashing lights or something, or just say, hi. How are you doing? I am still here. Then, I can use it as a tool more often" - P10.

While P10 made this comment because they were occasional user of sIPA, other participants who used sIPA more

TABLE 7. Categories and codes associated with "Concerns" theme.

Category	Code	Code Frequency
Privacy	Better Privacy	6
	Permission Based Recording	5
Speech Recognition	Dialect Understanding	3
	Accent Understanding	3
	Voice Recognition	5
Complex Queries	Unstructured Queries	4
	Complex Questions	5
	Cognitive Overload	3
Technology Behavior	Biased Information	2
	Out-of-Context Behavior	3

regularly had made similar comments, saying that it would make sIPA appear more human to them.

d: SPEECH RECOGNITION

Three participants were immigrants living in the United States. They reported that their VA was having problem understanding their accents, as summarized by P2.

> "The accent issues, I have heard the same things from other friend. Alexa does not relate to you. Alexa was not getting used to my accents and it was not understanding my patterns. Other friends complained the same" - P2.

Even a few participants who had North American heritage but lived in specific parts of the country, complained that their VA was not very good at picking up their dialect. Specifically, P7 and P8 from the Southern part of the United States, mentioned that their VA often misunderstood their commands.

> "It used to misinterpret when I wanted to say 'on' and 'off' because both of them start with "O". What is the weather today? It used to get confused with whether" - P8.

Overall, this theme shows that participants expected their VAs to better reflect their own personalities and styles.

e: COMPLEX QUERY PROCESSING

Many participants echoed their frustration with VAs not being able to understand more complex queries. Participants reported these struggles while mentioning that sIPA worked better with simpler commands, though even those simple queries could use a lot of improvement. To have their complex queries processed, participants had to reformulate their questions several times before they received an answer or gave up. P1's comment below sums up this problem concisely:

> "Sometimes with complex things, it has been frustrating. I asked it to help me with some technology, but it was not able to help me find the answer" - P1.

Similarly, P7 explained that her sIPA was not very good with processing questions without context. She had discovered that it is better to provide more specific prompts, but it meant that the user must have sufficient knowledge about the topic beforehand.

> "Let's say I would say, I would like to visit a restaurant near me. And then it does not give me

anything that I want, but then, if I say I would like a restaurant in Ipswich, Massachusetts, it would give me a better response. If not, I have to be even more specific" - P7.

P7 indicated that she would prefer to have her VA provide her with a simple straight forward answer rather than have her engage in continuous trial and error to find what she was looking for.

> "If I would like it to understand me better, I guess, on the first try, rather than for me to ask a few questions and I would rather it say I do not understand, rather than give me soup, you know of erroneous information" - P7.

Furthermore, participants indicated that sometimes the response given by the device was too long or complex. It made it harder for participants to pay attention and process the information. Many indicated that their sIPA "loaded their brain" with information, and expected them to "keep it all in their minds". This shows that VAs output should be designed with the recognition of the limited and reduced age-related cognitive processing capability of its users.

IV. DISCUSSION

The purpose of this study was to investigate how older adults used VAs and what was their experiences with this device. The qualitative analysis led to the identification of two broad themes consisting of various categories. The "Usage" theme is an amalgamation of seven main ways in which participants were currently using and wanted to make future use of their VAs. The main usage categories were entertainment, companionship, cognitive aid, home control, information, communication and healthcare. The second theme covers concerns around using VAs, including privacy invasion, out-of-context behavior, unexplained behavior, speech recognition, and complex query processing. Based on our findings, we discuss some consideration for improving usability and usefulness of VAs for older adults.

A. PRIVACY IMPROVEMENT

Similar to previous studies, we also found that privacy is an important consideration for older adults. Although, our study has a small sample size, the concern about privacy invasion was a recurring theme causing at least four participants to abandon or limit their VA use. It is important to note that these participants did not use their sIPAs for anything else besides entertainment. Previous research has also found that older adults who do not discover the value and use of VAs beyond entertainment, ultimately, abandon them due to privacy concerns [34]. Our findings, further show that those participants who had continued to use their sIPA seemed to be exercising caution and questioning the dangers of invasion of their "not-so-private" privacy.

Two branches of discussion, perception of privacy and actual security threat emerges from this concern [24], [60]. The existing literature explains that older adults are,

generally, skeptical towards adopting a new computerized or technology-based product due to privacy concerns [24]. Solutions such as teaching older adults about security and privacy is an effective method to inculcate knowledge of this issue [61]. Hence, we recommend preparing teaching manuals to improve older adults' awareness around privacy and security concerns related to VAs.

The second concern indicates that privacy has to be front and center in VA's design. Even though commercial VAs claim to provide privacy settings, it is clear from our study and existing literature that the existing system lacks transparency. Apparently, sIPAs allow users to choose a privacy setting but the option labels are confusing and contradictory to the point that the user does not want to bother and opts for a choice that ultimately benefits the manufacturers [62]. According to our study findings, older adults perceive invasion of privacy as someone listening in to their conversations all the time and recording them for some ulterior motive. Therefore, user has to be given the power to control and manage these concepts i.e. conversation recording and listening in the privacy settings. Moreover, user has to be made aware of how each privacy setting might help or protect them. Overall, we advocate for revision of the privacy setting and increasing their accessibility for the older adults.

B. INTERPERSONAL SKILLS
Our results confirm previous studies reporting older adults' interest in using VAs for companionship and entertainment [33]. According to our findings, while the interactivity created through natural language is accepted and welcomed by older adults, they expect VAs to engage in a higher-order interpersonal communication for tackling loneliness and receiving information. Particularly, study participants alluded to the need for a VA with effective communication skills, passionate listening ability, and caring and warm personality. In addition, trust and relatedness were recurrent codes in our analysis, and are also echoed in related works [63], [64].

Published literature calls for building improvisation capabilities within VAs [65] and moving away from structured and prescribed behaviors. We, on the other hand, want to draw research community's attention to the concept of intimacy, which is an aspect of interpersonal relationships and, according to the psychology literature, a powerful determinant of health and well-being. To elucidate this concept, we refer to the model of intimacy proposed by Reis and Shaver (1988) based on a review of psychological theories and literature. This model clarifies the nature of intimacy within a single interaction between two entities (A and B). According to this model, entity B attempts to understand and formulate a response to entity A's disclosure of motives, feelings and needs by sending it through its interpretative filter and informed by its own goals, desires, needs and fears. Entity A uses its own interpretive filter (again informed by its motives, goals, needs and fears) to process entity B's response to the disclosure in question, and to align the given feedback with its needs and goals. Currently, it appears that VA's interpretive

filter is informed by the manufacturer's goals as opposed to older adults' goals. Future research should investigate how the interpretive filters of VAs should be designed to align with older adults' expectations of their interactions with VAs.

Moving forward, researchers might also be interested in investigating how these intimate interactions can be used to create long-term relationships between older adults and VAs. Reis and Shaver explain that an intimate relationship between A and B can be viewed as an accumulation of a series of their one-time, intimate interactions, but it is greater than the sum of its parts. In the future, researchers should explore how intimate relationships between older adults and VAs could be derived from a history of a series of one-time intimate interactions.

C. CONTEXTUAL AWARENESS
Our results show that the out-of-context and unexplainable behavior of VAs is negatively impacting older adults' trust towards devices, causing some of them to even abandon these devices. There are two aspects of this result, firstly, future design of VAs needs to focus on building user's trust with these devices. We have discussed some ideas about improving trust in improving privacy section. The second aspect is improving contextual awareness of these devices. Participants in our study indicated that, in the future, they would expect sIPAs to have a better understanding their contexts and improved capability to produce output aligned with their needs. In fact, participants described this concept as sIPAs having ESP. Researchers such as Chung and Lee [38] have demonstrated that it is possible for sIPAs to develop a level of awareness about users' lifestyle by collecting various types of data. With the advent of the internet-of-things, it is now possible for different devices, including sIPAs, to exchange data with each other. Therefore, through data exchange and communication with other home and personal devices, sIPAs can develop even better understanding of its users and generate more relevant information for them. Hence, there is a need to understand what kind of user models can be created to meet older adults' needs and improve their quality of life.

V. FUTURE RESEARCH FRAMEWORK
The implementation of the proposed design suggestions are possible to some extent through the combined expertise of practitioners and researchers from various backgrounds and fields such as social sciences, psychology, natural language processing, machine learning, software engineering and human-computer interaction. Here we append a framework that derives knowledge from implementation based fields to generate ideas for enacting our design suggestions and more. The framework consists of the following components:

- Obtaining permission from the human to record conversation.
- Explaining generated insights from recorded conversation.
- Installing variable voice (dialect, accent, complex voice) recognition mechanisms.

- Building a multi-component module to generate humanized conversations.

A. PERMISSION-BASED RECORDING

Some interviewees had noted that while having personal conversation with others, sIPA would turn on and start recording if it were indirectly mentioned in the conversation. This made older adults cautious of their speech and sometimes even caused inconvenience. To counteract this, older adults should be provided with a simple and transparent mechanism [66] to preserve their privacy. It is clear that older adults value hands-free nature of sIPA, therefore, we caution against adding physical switches to control privacy. A possible suggestion could be implementing a permission-based recording [67] mechanism that would function as follows: sIPA's data collection engine would be off by default, on being awakened with a prompt (i.e. Ok, Google, Alexa), sIPA would have to ask the user for permission to record (Figure 1). The recording status of sIPA should be physically visible to the user at all times, possibly with the help of an external representation such as a light emitting device (LED).

One downside of this suggestion is that it has the potential to disrupt user's focus or divert their attention away from the primary task. This should be an important consideration when designing for older adults as they often tend to experience reduced cognitive ability with aging. On the other hand, it could increase user's sense of empowerment and control over the device. But this needs further investigation and refinement. As such, we urge the research community to explore this area further.
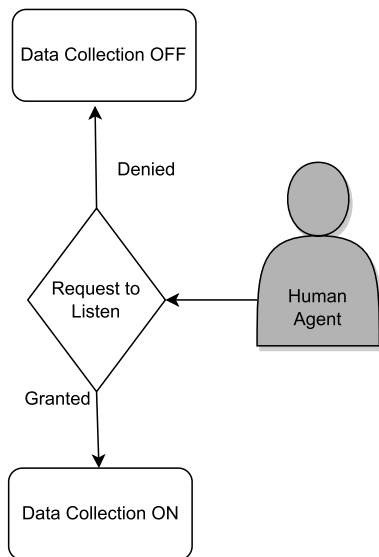

FIGURE 1. Mechanism for permission-based recording.

B. INSIGHT EXPLANATION

Here we discuss one possible mechanism to promote trust and transparency in the interpersonal relationships and, possibly, conceptualize properties of the interpretive filter for older adults and VAs. It is evident that sIPAs manufacturers generate profits by providing recommendations based on user's online search patterns and purchasing behaviors [68]. We suggest that such recommendations should be accompanied with explanations clarifying how such information is being formulated for the user (insight explanation). This concept is referred to as transparent AI and / or explainable AI (XAI) [69]. Techniques such as feature importance, feature dependencies, what-if analysis, partial dependence, feature interaction, correlation, and decision trees, AI / machine learning (ML) models can generate explanations of their results for the users [70], [71].

While the technical details of these techniques can be found in other publications, [70], [72] we give a brief overview here. Partial dependence plots (PDP) are utilized to visualize and analyze the interaction and dependence between the target response and features of interest. Similarly, individual conditional expectation (ICE) offers explanatory visualizations by showing an explanation for each feature separately, whilst PDP provides an average value of the explanation. Furthermore, accumulated local effects (ALE) provides dependence plots that can be used to understand the average impact of new inputs on classification or regression models. Local interpretable model-agnostic explanations (LIME) can be used to explain individual predictions.

We hypothesize that when recommendations (insights) generated by VAs are accompanied with explanations, older adults' trust in VAs will improve. Future research is needed to investigate this claim. We provide an implementation architecture for insight generation and explanation in Figure 2.
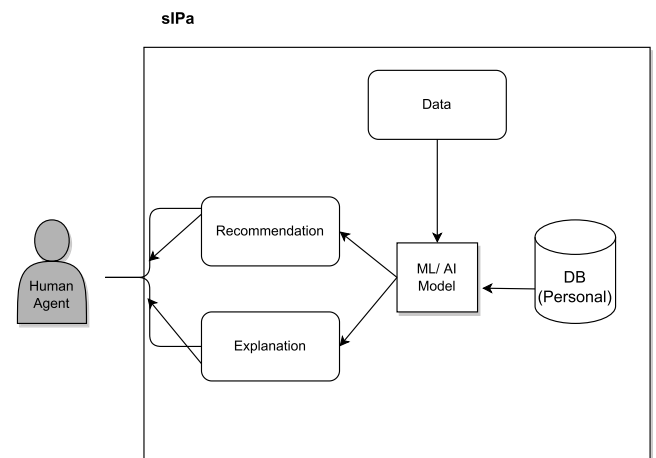

FIGURE 2. A possible architecture for insight and explanation generation.

C. VARIABLE SPEECH RECOGNITION

Based on older adults' annoyance and complaints regarding sIPA's failure to understand speech due to dialect and accent issues, we suggest that future VAs should place special emphasis on improving its speech recognition capabilities. Particularly, we emphasize inclusion of a pre-installed dialect and accent recognition module in sIPAs [73]. In the following subsections, we describe two approaches to developing such a module.
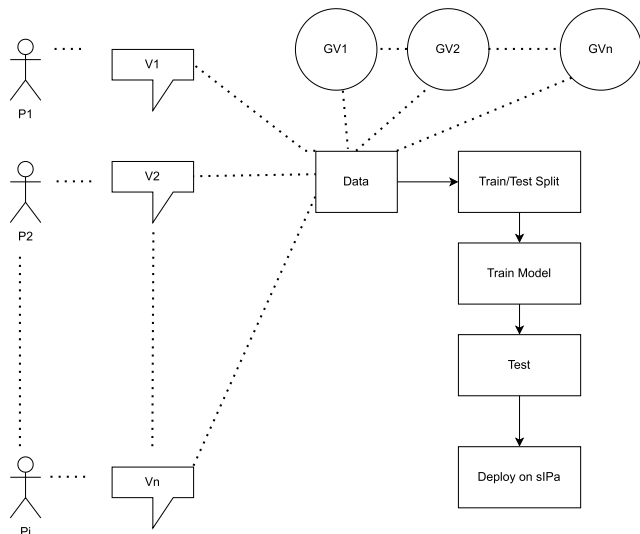
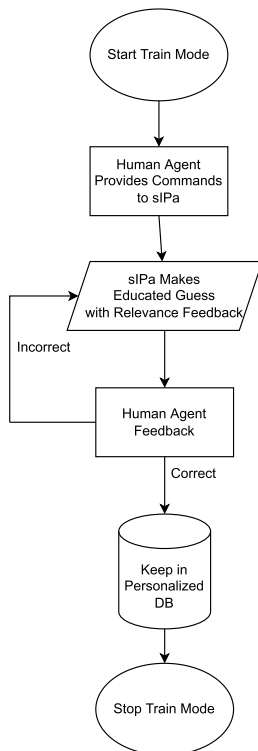**FIGURE 3.** Top-down approach for dialect and accent recognition.

FIGURE 4. Bottom-up approach for dialect and accent recognition.

1) TOP-DOWN APPROACH

In the top-down approach, there would be n persons $(P_1 \ldots P_n)$ and n text transcripts $(GV_1 \ldots GV_n)$. The n persons would read aloud in their voices $(V_1 \ldots V_n)$ each one of the n transcripts $(GV_1 \ldots GV_n)$, generating a $N \times N$ matrix, that is, $((P_1V_1GV_1 \ldots P_1V_1GV_n) \ldots (P_2V_2GV_1 \ldots P_2V_2GV_n) \ldots (P_nV_nGV_1 \ldots P_nV_nGV_n))$. This matrix would be used against a ground value to train a variable speech recognition model to be deployed on sIPAs. This explanation assumes that data cleaning, pre-processing, and noise

removal steps have already been performed. We have summarized this description in a diagrammatic format in Figure 3.

2) BOTTOM-UP APPROACH

In the bottom-up approach, sIPA would attempt to develop an understanding of user's dialect and accent in an unsupervised way, after successfully obtaining recording permission from the user. Upon listening, the sIPA would make educated guesses through trial and error and relevance feedback loop (human-in-the-loop). A detailed step by step description of this process has been shown in Figure 4.

D. HUMANIZING CONVERSATION

Given that older adults expect to interact with a more humanized entity, here we discuss how existing tools can be leveraged to provide such an experience to older adults. Since humanization is a broad concept with metrics and benchmarks still in development [74], based on our findings, we propose the following framework for humanizing sIPA's conversation ability: a) emotion recognition [75], b) sentiment analysis [76], c) casual non-objective conversation [77], and (d) cognitive load checking module [78]. We have summarized the proposed humanized conversation framework in Figure 5.

The emotion recognition engine would generate emotional insights by understanding text (acquired through transcription of speech to text using automatic speech recognition), analyzing vocal features of speech and detecting changes in facial expressions through its data collection engine. Furthermore, this data would be used for sentiment analysis and opinion mining to allow sIPAs to have casual and less rigid conversations. Our findings indicate that older adults are prone to long conversations that sIPAs process as a complex query, which can result in undesirable and confusing outputs for older adults. This negative user experience can lead to limited usage and, ultimately, abandonment of VAs. The authors recommend the inclusion of a casual non-objective conversation processing mechanism that would allow sIPAs to process long unstructured conversations and queries. A cognitive load check module should also be developed to monitor output size and ensure its suitability for older adults. Since cognitive decline with aging can make it challenging for older adults to maintain information in their short term memory.

One assumption built into this framework is that the sIPAs have video recording capability, which is currently not available most sIPAs. We recognize limitation of this framework and recommend further refinement based on additional research.

VI. LIMITATIONS AND FUTURE WORK

We acknowledge the small number of participants as a limitation and intend to improve by running a larger study to further analyze and evaluate usability and user experience of sIPAs. Future research also includes subjective tests such as System Usability Scale (SUS) with more participants to support the usability perceptions noted in this research work with
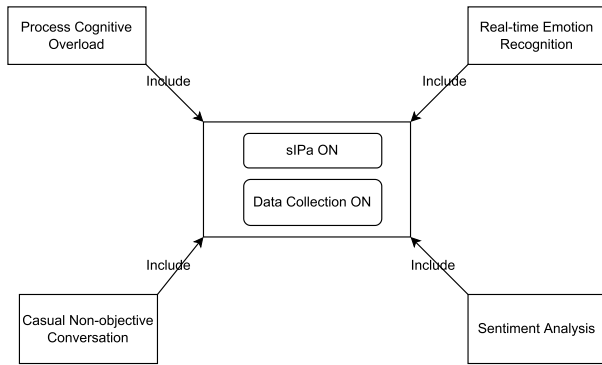
FIGURE 5. Recommended framework for generating humanized conversation.

objective evaluation. Exclusion of special needs personnel is noted as another limitation of our study. Inclusion of participants from diverse backgrounds, although challenging, is another future possibility in this line of work. The future work should also focus on prototype and actual implementation followed by user evaluation to assess effectiveness of the proposed suggestions.

VII. CONCLUSION

Though many older adults view sIPAs as an entertainment and information generating device, there is a recognition among this age group that such devices can play an important and significant role in their lives. In this regard, there are at least seven basic ways in which older adults might be interested in using sIPAs. However, current design of sIPA is insufficient to meet older adults' needs and expectations. Specifically, privacy considerations, interpersonal skills and contextual relevance of these devices require significant improvements to meet older adults' expectations. The design community should work closely with the developer community to implement the existing works in permission-based recording, XAI, variable speech recognition, and humanized conversations to accelerate the improvement of these devices. We are in the process of implementing a prototype of the proposed framework and evaluating it with the target users.

REFERENCES

[1] S. Lovato and A. M. Piper, "'Siri, is this you?' understanding young children's interactions with voice input systems," in *Proc. 14th Int. Conf. Interact. Design Children*, 2015, pp. 335–338.

[2] A. S. Tulshan and S. N. Dhage, "Survey on virtual assistant: Google assistant, Siri, Cortana, Alexa," in *Proc. Int. Symp. Signal Process. Intell. Recognit. Syst.* Singapore: Springer, 2019, pp. 190–201.

[3] D. S. Zwakman, D. Pal, and C. Arpnikanondt, "Usability evaluation of artificial intelligence-based voice assistants: The case of Amazon Alexa," *Social Netw. Comput. Sci.*, vol. 2, no. 1, pp. 1–16, Feb. 2021.

[4] A. Schlomann, H.-W. Wahl, P. Zentel, V. Heyl, L. Knapp, C. Opfermann, T. Krämer, and C. Rietz, "Potential and pitfalls of digital voice assistants in older adults with and without intellectual disabilities: Relevance of participatory design elements and ecologically valid field studies," *Frontiers Psychol.*, vol. 12, p. 2522, Jul. 2021.

[5] S. K. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*. Boca Raton, FL, USA: CRC Press, 2018.

[6] J. Zimmerman, J. Forlizzi, and S. Evenson, "Research through design as a method for interaction design research in HCI," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Apr. 2007, pp. 493–502.

[7] G. Cockton, "Value-centred HCI," in *Proc. 3rd Nordic Conf. Hum.-Comput. Interact.*, 2004, pp. 149–160.

[8] P. Ketola and V. Roto, "Exploring user experience measurement needs," in *Proc. 5th COST-MAUSE Open Workshop Valid Useful User Exp. Meas. (VUUM)*. Reykjavik, Iceland: Citeseer, 2008, pp. 23–26.

[9] N. Bevan, "What is the difference between the purpose of usability and user experience evaluation methods," in *Proc. Workshop UXEM*, vol. 9, no. 1. Princeton, NJ, USA: Citeseer, 2009, pp. 1–4.

[10] P. W. Jordan, *An Introduction to Usability*. Boca Raton, FL, USA: CRC Press, 2020.

[11] S. Bødker and Y. Sundblad, "Usability and interaction design—New challenges for the Scandinavian tradition," *Behav. Inf. Technol.*, vol. 27, no. 4, pp. 293–300, Jul. 2008.

[12] J. R. Lewis, "Usability testing," in *Handbook of Human Factors and Ergonomics*, G. Salvendy, Ed., 4th ed. New York, NY, USA: Wiley, 2012, pp. 1267–1312.

[13] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, "Usability measurement and metrics: A consolidated model," *Softw. Qual. J.*, vol. 14, no. 2, pp. 159–178, Jun. 2006.

[14] H. Petrie and N. Bevan, "The evaluation of accessibility, usability, and user experience," in *The Universal Access Handbook*, vol. 1. Boca Raton, FL, USA: CRC Press, 2009, pp. 1–16.

[15] N. Thakur and C. Y. Han, "A review of assistive technologies for activities of daily living of elderly," 2021, *arXiv:2106.12183*.

[16] A. W. Smith, "User experience design for older adults: Experience architecture and methodology for users aged 60+," in *Proc. 37th ACM Int. Conf. Design Commun.*, Oct. 2019, pp. 1–9.

[17] J. D. Portz, E. A. Bayliss, S. Bull, R. S. Boxer, D. B. Bekelman, K. Gleason, and S. Czaja, "Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: Descriptive qualitative study," *J. Med. Internet Res.*, vol. 21, no. 4, Apr. 2019, Art. no. e11604.

[18] M. Caboral-Stevens, M. V. Whetsell, L. S. Evangelista, B. Cypress, and D. Nickitas, "Usability framework for older adults," *Res. Gerontol. Nursing*, vol. 8, no. 6, pp. 300–306, 2015.

[19] N. Giunta and M. L. Thomas, "Integrating assessment and evaluation into partnership initiatives: Lessons from the community partnerships for older adults," *J. Appl. Gerontol.*, vol. 34, no. 5, pp. 609–632, Aug. 2015.

[20] K. O'Brien, A. Liggett, V. Ramirez-Zohfeld, P. Sunkara, and L. A. Lindquist, "Voice-controlled intelligent personal assistants to support aging in place," *J. Amer. Geriatrics Soc.*, vol. 68, no. 1, pp. 176–179, Jan. 2020.

[21] P. Nallam, S. Bhandari, J. Sanders, and A. Martin-Hammond, "A question of access: Exploring the perceived benefits and barriers of intelligent voice assistants for improving access to consumer health resources among low-income older adults," *Gerontol. Geriatric Med.*, vol. 6, Dec. 2020, Art. no. 2333721420985975.

[22] A. Pyae and T. N. Joelsson, "Investigating the usability and user experiences of voice user interface: A case of Google home smart speaker," in *Proc. 20th Int. Conf. Hum.-Comput. Interact. Mobile Devices Services Adjunct*, Sep. 2018, pp. 127–131.

[23] S. Kim, "Exploring how older adults use a smart speaker–based voice assistant in their first interactions: Qualitative study," *JMIR mHealth uHealth*, vol. 9, no. 1, Jan. 2021, Art. no. e20427.

[24] K. Bonilla and A. Martin-Hammond, "Older adults' perceptions of intelligent voice assistant privacy, transparency, and online privacy guidelines," in *Proc. 16th Symp. Usable Privacy Secur. (SOUPS)*, 2020, pp. 1–5.

[25] C. F. Corbett, E. M. Combs, P. J. Wright, O. L. Owens, I. Stringfellow, T. Nguyen, and C. R. Van Son, "Virtual home assistant use and perceptions of usefulness by older adults and support person dyads," *Int. J. Environ. Res. Public Health*, vol. 18, no. 3, p. 1113, Jan. 2021.

[26] S. Shalini, T. Levins, E. L. Robinson, K. Lane, G. Park, and M. Skubic, "Development and comparison of customized voice-assistant systems for independent living older adults," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, Jul. 2019, pp. 464–479.

[27] S. Kim and A. Choudhury, "Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study," *Comput. Hum. Behav.*, vol. 124, Nov. 2021, Art. no. 106914.

[28] J. Kowalski, A. Jaskulska, K. Skorupska, K. Abramczuk, C. Biele, W. Kopeć, and K. Marasek, "Older adults and voice interaction: A pilot study with Google home," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, May 2019, pp. 1–6.

[29] A. Pradhan, A. Lazar, and L. Findlater, "Use of intelligent voice assistants by older adults with low technology use," *ACM Trans. Comput.-Hum. Interact.*, vol. 27, no. 4, pp. 1–27, Aug. 2020.

[30] M. Hanley and S. Azenkot, "Understanding the use of voice assistants by older adults," 2021, *arXiv:2111.01210*.

[31] L. Pepino, P. Riera, L. Gauder, A. Gravano, and L. Ferrer, "Detecting distrust towards the skills of a virtual assistant using speech," 2020, *arXiv:2007.15711*.

[32] M. Kraus, N. Wagner, Z. Callejas, and W. Minker, "The role of trust in proactive conversational assistants," *IEEE Access*, vol. 9, pp. 112821–112836, 2021.

[33] K. Gupta, R. Hajika, Y. S. Pai, A. Duenser, M. Lochner, and M. Billinghurst, "Measuring human trust in a virtual assistant using physiological sensing in virtual reality," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2020, pp. 756–765.

[34] M. Trajkova and A. Martin-Hammond, "'Alexa is a toy': Exploring older adults' reasons for using, limiting, and abandoning echo," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2020, pp. 1–13.

[35] D. Pal, M. D. Babakerkhell, and X. Zhang, "Exploring the determinants of users' continuance usage intention of smart voice assistants," *IEEE Access*, vol. 9, pp. 162259–162275, 2021.

[36] J. Balsa et al., "Intelligent virtual assistant for promoting behaviour change in older people with T2D," in *Proc. EPIA Conf. Artif. Intell.* Cham, Switzerland: Springer, Sep. 2019, pp. 372–383.

[37] L. M. Bowen, "Resisting age bias in digital literacy research," *College Composition Commun.*, vol. 62, no. 4, pp. 586–607, 2011.

[38] H. Chung and S. Lee, "Intelligent virtual assistant knows your life," 2018, *arXiv:1803.00466*.

[39] Y. Liao, J. Vitak, P. Kumar, M. Zimmer, and K. Kritikos, "Understanding the role of privacy and trust in intelligent personal assistant adoption," in *Proc. Int. Conf. Inf.* Cham, Switzerland: Springer, Mar. 2019, pp. 102–113.

[40] J. Lau, B. Zimmerman, and F. Schaub, "Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers," *Proc. ACM Hum.-Comput. Interact.*, vol. 2, pp. 1–31, Nov. 2018.

[41] H. Chung, M. Iorga, J. Voas, and S. Lee, "Alexa, can I trust you?" *Computer*, vol. 50, no. 9, pp. 100–104, 2017.

[42] P.-S. Chiu, J.-W. Chang, M.-C. Lee, C.-H. Chen, and D.-S. Lee, "Enabling intelligent environment by the design of emotionally aware virtual assistant: A case of smart campus," *IEEE Access*, vol. 8, pp. 62032–62041, 2020.

[43] K. Weitz, D. Schiller, R. Schlagowski, T. Huber, and E. André, "'Do you trust me?': Increasing user-trust by integrating virtual agents in explainable AI interaction design," in *Proc. 19th ACM Int. Conf. Intell. Virtual Agents*, Jul. 2019, pp. 7–9.

[44] K. Sokol and P. Flach, "Glass-box: Explaining AI decisions with counterfactual statements through conversation with a voice-enabled virtual assistant," in *Proc. IJCAI*, Jul. 2018, pp. 5868–5870.

[45] K. Weitz, D. Schiller, R. Schlagowski, T. Huber, and E. André, "Let me explain!': Exploring the potential of virtual agents in explainable AI interaction design," *J. Multimodal User Interfaces*, vol. 15, no. 2, pp. 87–98, 2021.

[46] J. Matani, P. Gervais, M. Calvo, S. Feuz, and T. Deselaers, "Matching language and accent in virtual assistant responses," Technical Disclosure Commons, Tech. Rep. 1239, Jun. 2018. [Online]. Available: https://www.tdcommons.org/dpubs_series/1239

[47] D. Pal, C. Arpnikanondt, S. Funilkul, and V. Varadarajan, "User experience with smart voice assistants: The accent perspective," in *Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2019, pp. 1–6.

[48] A. Palanica, A. Thommandram, A. Lee, M. Li, and Y. Fossat, "Do you understand the words that are comin outta my mouth? Voice assistant comprehension of medication names," *NPJ Digit. Med.*, vol. 2, no. 1, pp. 1–6, Jun. 2019.

[49] U. Ahmed, G. Srivastava, U. Yun, and J. C.-W. Lin, "EANDC: An explainable attention network based deep adaptive clustering model for mental health treatment," *Future Gener. Comput. Syst.*, vol. 130, pp. 106–113, May 2022.

[50] U. Ahmed and J. C.-W. Lin, "Deep explainable hate speech active learning on social-media data," *IEEE Trans. Computat. Social Syst.*, early access, Apr. 20, 2022, doi: 10.1109/TCSS.2022.3165136.

[51] U. Ahmed, R. H. Jhaveri, G. Srivastava, and J. C.-W. Lin, "Explainable deep attention active learning for sentimental analytics of mental disorder," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, Jul. 2022.

[52] Y. Djenouri, A. Belhadi, G. Srivastava, and J. C.-W. Lin, "When explainable AI meets IoT applications for supervised learning," *Cluster Comput.*, pp. 1–11, Aug. 2022.

[53] M. Toomim, T. Kriplean, C. Pörtner, and J. Landay, "Utility of human-computer interactions: Toward a science of preference measurement," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, May 2011, pp. 2275–2284.

[54] D. Datcu, S. Lukosch, and F. Brazier, "On the usability and effectiveness of different interaction types in augmented reality," *Int. J. Hum.-Comput. Interact.*, vol. 31, no. 3, pp. 193–209, Mar. 2015.

[55] Q. G. To, C. Green, and C. Vandelanotte, "Feasibility, usability, and effectiveness of a machine learning–based physical activity chatbot: Quasi-experimental study," *JMIR mHealth uHealth*, vol. 9, no. 11, 2021, Art. no. e28577.

[56] M. Kuniavsky, "User experience and HCI," in *The Human-Computer Interaction Handbook*. Boca Raton, FL, USA: CRC Press, 2007, pp. 923–942.

[57] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006.

[58] A. J. Bingham and P. Witkowsky, "Deductive and inductive approaches to qualitative data analysis," in *Analyzing and Interpreting Qualitative Data: After the Interview*. Newbury Park, CA, USA: Sage, 2021, pp. 133–146.

[59] C. Burck, "Comparing qualitative research methodologies for systemic research: The use of grounded theory, discourse analysis and narrative analysis," *J. Family Therapy*, vol. 27, no. 3, pp. 237–262, Aug. 2005.

[60] J. Lien, M. A. Al Momin, and X. Yuan, "Attacks on voice assistant systems," in *Security, Data Analytics, and Energy-Aware Solutions in the IoT*. Hershey, PA, USA: IGI Global, 2022, pp. 61–77.

[61] B. Chaudhry, D. Dasgupta, M. Mohamed, and N. Chawla, "Teaching tablet technology to older adults," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, Jul. 2021, pp. 168–182.

[62] The Neeva Team. (Apr. 2021). *Is My Phone Listening to Me?* [Online]. Available: https://neeva.com/learn/is-my-phone-listening-to-me

[63] D. Fisk, N. Charness, S. J. Czaja, W. A. Rogers, and S. Sharit, *Designing for Older Adults*. Boca Raton, FL, USA: CRC Press, 2004.

[64] J. Balsa, I. Félix, A. P. Cláudio, M. B. Carmo, I. C. E. Silva, A. Guerreiro, M. Guedes, A. Henriques, and M. P. Guerreiro, "Usability of an intelligent virtual assistant for promoting behavior change and self-care in older people with type 2 diabetes," *J. Med. Syst.*, vol. 44, no. 7, pp. 1–12, Jul. 2020.

[65] M. R. Ali, K. Van Orden, K. Parkhurst, S. Liu, V.-D. Nguyen, P. Duberstein, and M. E. Hoque, "Aging and engaging: A social conversational skills training program for older adults," in *Proc. 23rd Int. Conf. Intell. User Interfaces*, Mar. 2018, pp. 55–66.

[66] G. Liu and N. Malkin, "Effects of privacy permissions on user choices in voice assistant app stores," *Proc. Privacy Enhancing Technol.*, vol. 2022, no. 4, pp. 421–439, Oct. 2022.

[67] D. Dasgupta, A. Roy, and D. Ghosh, "Multi-user permission strategy to access sensitive information," *Inf. Sci.*, vol. 423, pp. 24–49, Jan. 2018.

[68] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (XAI): Toward medical XAI," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4793–4813, Nov. 2021.

[69] A. B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020.

[70] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–10.

[71] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, "From local explanations to global understanding with explainable AI for trees," *Nature Mach. Intell.*, vol. 2, no. 1, pp. 56–67, 2020.

[72] M. U. Islam, M. Mottalib, M. Hassan, Z. I. Alam, S. M. Zobaed, and F. Rabby, "The past, present, and prospective future of XAI: A comprehensive review," in *Explainable Artificial Intelligence for Cyber Security*. Cham, Switzerland: Springer, 2022, pp. 1–29.

[73] F. Faisal, S. Keshava, and A. Anastasopoulos, "SD-QA: Spoken dialectal question answering for the real world," 2021, *arXiv:2109.12072*.

[74] L. Leahu, S. Schwenk, and P. Sengers, "Subjective objectivity: Negotiating emotional meaning," in *Proc. 7th ACM Conf. Designing Interact. Syst.*, 2008, pp. 425–434.

[75] M.-C. Lee, S.-Y. Chiang, S.-C. Yeh, and T.-F. Wen, "Study on emotion recognition and companion chatbot using deep neural network," *Multimedia Tools Appl.*, vol. 79, nos. 27–28, pp. 19629–19657, Jul. 2020.

[76] A. Alm and L. M. Nkomo, "Chatbot experiences of informal language learners: A sentiment analysis," *Int. J. Comput.-Assist. Lang. Learn. Teach.*, vol. 10, no. 4, pp. 51–65, 2020.

[77] J. Simpson, F. Gaiser, M. Macík, and T. Breßgott, "Daisy: A friendly conversational agent for older adults," in *Proc. 2nd Conf. Conversational User Interfaces*, Jul. 2020, pp. 1–3.

[78] Y. Sun, S. Li, and L. Yu, "The dark sides of AI personal assistant: Effects of service failure on user continuance intention," *Electron. Markets*, vol. 32, pp. 17–39, Aug. 2021.

MUHAMMAD USAMA ISLAM received the B.Sc. degree in computer science from the Islamic University of Technology, Bangladesh. He is currently pursuing the Ph.D. degree in computer science with the School of Computing and Informatics, University of Louisiana at Lafayette, Lafayette, LA, USA. He worked as a Lecturer with the Asian University of Bangladesh, Bangladesh. His research interests include human–computer interaction, affective computing, and UI/UX research.

BEENISH MOALLA CHAUDHRY received the Ph.D. degree in health informatics and human–computer interaction from Indiana University at Bloomington, Bloomington, IN, USA. She worked as a Postdoctoral Research Associate with the Interdisciplinary Center for Network Science and Applications, University of Notre Dame. She is currently an Assistant Professor with the School of Computing and Informatics, University of Louisiana at Lafayette, Lafayette, LA, USA. She uses user-centered design methodology to design and evaluate mobile applications for chronically ill and low literacy populations. She also uses a number of theoretical and empirical tools, such as community-based participatory research, participatory action research, and social determinants to study how socio-ecological factors involved in technology use impact health of people from low socioeconomic backgrounds. Her research interests include pervasive/ mobile computing, human–computer interaction, and health informatics.

• • •