Voicing Suggestions and Enabling Reflection: Results of an Expert Discussion on Proactive Assistants for Time Management

Jovan Jeromela
jeromelj@tcd.ie
ADAPT Centre, Trinity College Dublin
Dublin, Ireland

Owen Conlan
owen.conlan@tcd.ie
ADAPT Centre, Trinity College Dublin
Dublin, Ireland

ABSTRACT

While voice-controllable Intelligent Personal Assistants (IPAs) have become widespread in recent years, they remain primarily reactive with rather constrained calendaring capabilities. Anticipating more adaptive and complex assistants in the future, we organised a multidisciplinary expert discussion investigating potential use cases, interaction principles, and user modelling challenges within proactive IPAs for time management. This paper presents the identified themes and deliberations on enticing self-reflection, longitudinal task assistance, interaction modality, dialogue design, perception of the system, usage willingness, onboarding, and explainability. These findings outline a framework of advanced IPAs for time management.

CCS CONCEPTS

• **Human-centered computing** → **Personal digital assistants**; *User models*; *Empirical studies in HCI*; *User centered design*; Natural language interfaces.

KEYWORDS

adaptation, context, data representation, personalisation, control, time management

ACM Reference Format:
Jovan Jeromela and Owen Conlan. 2023. Voicing Suggestions and Enabling Reflection: Results of an Expert Discussion on Proactive Assistants for Time Management. In *ACM conference on Conversational User Interfaces (CUI '23), July 19–21, 2023, Eindhoven, Netherlands.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3571884.3604317

1 INTRODUCTION

Intelligent Personal Assistants (IPAs) – also referred to as virtual, digital, or, when it is the dominant interaction modality, voice assistants[1] – became practically omnipresent in the day-to-day lives of users through integration in personal computers, smartphones, smartwatches, smart TVs, and smart speakers [5, 9, 20]. Prior works [8, 9, 15, 21] foretell that future IPAs are to be *proactive* and more *personalised*, which would enable the assistants to predict and perform actions most appropriate for the domain, user's context, and user's intent. Researchers have begun to examine proactive assistant-initiated interactions [29], yet the proactivity of commercial assistants remains limited to presenting relevant informative content, notifying users about events, and smart-home routines [11, 21, 22, 29]. Furthermore, Murad and Munteanu [19] observed that, regardless of the ever-advancing underlying artificial intelligence, voice assistants offer "marginally more than an information kiosk or automation control" for many user groups due to the lack of specific interaction guidelines.

This work seeks to contribute to establishing such interaction guidelines and, more concretely, to outline a framework of a voice-enabled and proactive IPA for time management. Help in managing time is a desired feature of IPAs [25] that has the potential to save the user's time [7] and facilitate collaboration [23]. Therefore, we organised a multidisciplinary expert discussion to devise a preliminary framework of IPAs for time management inspired by Meurisch et al.'s study on general-purpose proactive IPAs [16]. This paper presents the results of the thematic analysis [4] of the discussion that ensued.

2 RELATED WORK

Yorke-Smith et al. [28] described a framework for proactive task management and meeting scheduling for CALO, a project on building a personalised assistant for knowledge workers [2]. More recently, Cranshaw et al. [7] presented Calendar.help, a mail-based meeting scheduling system employing microtasks and crowdsourcing. Jain et al. [10] implemented a context-aware auto-response system that notifies contacts of the user's situational awareness, e.g. that they are in a meeting.

To study the appropriate proactive dialogue behaviour of a voice assistant, Zargham et al. [29] designed six groups of storyboards. Groups "reminder" (wherein the assistant warns the user that if they continue snoozing the alarm, they would miss a meeting) and "health risk" (the assistant proposes to schedule an appointment with a physician after hearing the user cough) received the highest median ratings for both appropriateness and usefulness by the study participants. It was reported that participants generally welcomed IPA interventions in urgent situations or in multi-user scenarios if they save time. Similarly, publications on mental health interventions [1, 3] indicated that users value adaptive notifications and the option of getting insights from data. Yet, group "advice" in [29], in which the assistant nudges the user to reduce TV-watching, was received far less favourably. Zargham et al. connect this outcome in part to Luria et al.' study [15], which reported a wariness of surrendering decision-making agency to the assistant and a desire for the assistant to apprehend the user's social relations.

[1] In line with [6, 8, 21, 22, 25], we chose the name IPA because it indicates that the envisioned assistant's features are personalised and enabled by artificial intelligence.

CUI '23, July 19–21, 2023, Eindhoven, Netherlands

ACM ISBN 979-8-4007-0014-9/23/07.
https://doi.org/10.1145/3571884.3604317

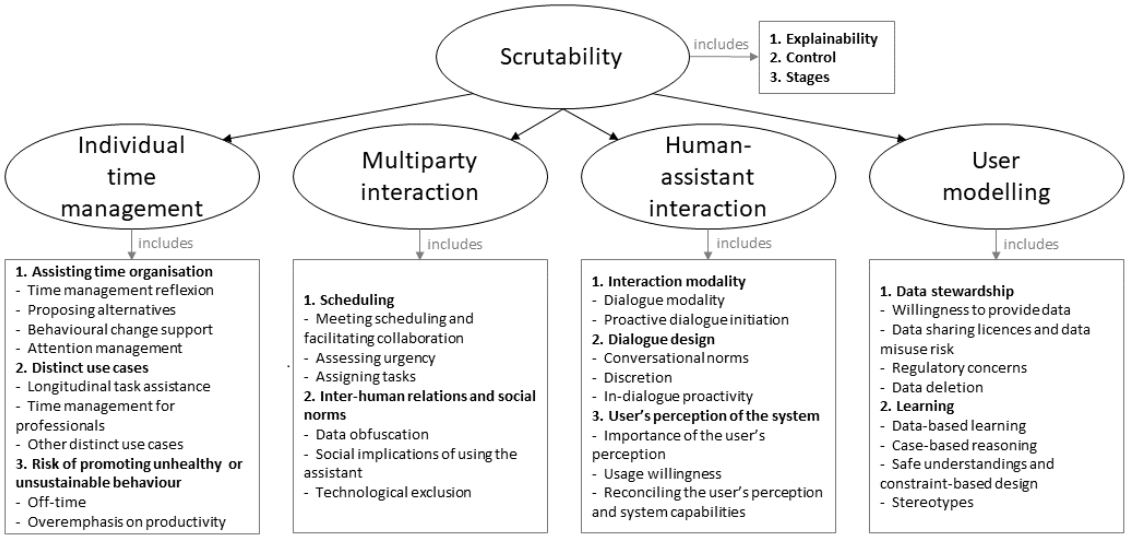

Figure 1: The four extracted themes and their subtopics; scrutability is seen as an overarching meta-theme of the discussion

One of the most well-known early proactive assistants was Clippy, which offered situational advice in Microsoft Office '97 [26]. Whitworth [26] suggested that Clippy ultimately failed because it had not followed the four rules of polite computing: respect user choice; disclose yourself; offer useful choices; and remember past choices. Similarly, Xiao et al. [27] hypothesised that the usefulness and relevance of offered proactive assistance might explain why the participants in their study reacted positively to their WoZ prototype while having a negative opinion of Clippy.

Prior works uncovered that usage barriers for both regular [14] and infrequent [6] users of commercial IPAs include misjudging system capabilities. The authors of these works both question whether human-like conversation style is something the IPAs should emulate and instead suggest having the voice assistant reveal its capabilities. Roger K. Moore [17] argued for the whole interaction design to be aligned to the component with the lowest level of performance within a voiced computer system. He maintained that only then the dialogue can be coherent to the user and illustrated the argument by proclaiming that a speech-enabled robot that is still too limited to provide a human-level dialogue flexibility should have a robot voice. Similarly, Moore and Nicolao [18] proposed a needs-based architecture that would help mitigate the limited depth of communications capabilities IPAs currently possess, while Viswanathan et al. [24] presented a WoZ IPA prototype that mimics interaction principles of graphical interfaces.

Another manner in which the user could understand the capabilities of an assistant is *scrutability*. This term, originally introduced by Judy Kay [12], describes adaptive systems that allow the user to study, interact with, and modify the system's user model. As presented in [13], through active scrutiny the user is able to understand how their data is used, to what extent the system is personalised to them, and if there are errors in the system's user model. Jeromela [11] proposed investigating how principles of scrutability relate to anticipated proactive IPAs and outlined a scrutable-by-design assistant model.

3 DISCUSSION STRUCTURE

Two professors, two associate professors, and three PhD candidates participated in the discussion. Their self-reported research interests are: 1) personalisation, adaptive systems, and proactivity; 2) data integration, databases, and the semantic web; 3) trustworthy AI, data governance, and standardisation; 4) conversational user interfaces, psychological and cognitive aspects of human-machine dialogue, and the impact of design on speech interface interaction; 5) empathetic chatbots; 6) simulation-based development of proactive IPAs; and 7) explainability and counterfactuals in reinforcement learning. Unfortunately, due to the unavailability of some initial invitees, there was only one female participant. The discussion took place in person and lasted approximately 90 minutes. Akin to the study in [16], during the first 25 minutes, the participants were asked to individually think of and present use cases and scenarios for the time management assistants. The main discussion followed, with the moderator presenting three discussion prompts, roughly 20 minutes apart. The prompts included discussing the presented use cases; human-assistant interaction; and graphically representing the envisioned IPA. The moderator (first author) kept the discussion open, only speaking to present and clarify prompts and to make sure all participants spoke. The second author participated in the discussion as an expert. The event was organised in line with the institutional ethical guidelines. The anonymised transcript and participants' initial use cases can be found here: https://drive.google.com/drive/folders/1-crFEKo2jyr7d-_5OgBSLlSF7B9SASP3?usp=sharing.

4 THEMATIC ANALYSIS

We performed a *thematic analysis* of our data following the five steps from the seminal paper by Braun and Clarke [4]. Contriving and revising themes was done collaboratively, with the two authors meeting weekly to discuss them over four rounds of revisions. As can be seen in figure 1, four main themes were identified. In addition,

scrutability, i.e. the manner in which the users can study, understand, and control the personalisation of the IPA was identified as a meta-theme that spans through all themes. *Stages* emerged in the discussion as an onboarding mechanism through which the user would be gradually introduced to the system's capabilities. This paper focuses on the findings that relate to the conversational interface and human-assistant interaction with anticipated future work to elaborate further on points regarding ethics and scrutability.

4.1 Individual Time Management

The theme of individual time management comprised ideas and comments on how the IPA could help an individual manage their time for tasks that do not directly affect other people, such as *behavioural change*. For instance, it was suggested that the IPA could nudge the user to meet their physical activity goals or stay within the desired screen time through proactive notifications asking the user to get back to the main task or take a break. The ideal mode of delivery of such disruptions was discussed extensively (cf. subsection 4.3), especially when the user is potentially involved in safety-critical tasks such as driving or performing medical procedures. Furthermore, as *p*articipant *B* in *s*tatement *46* – shortened as *PB-S46* in the remainder of the text – explained, there are professions, such as lawyers and consultants, where the IPA could potentially propose a more goal-appropriate schedule. Moreover, PD-S136 suggested that longitudinal tasks – spanning a long time or comprising multiple steps – would be a specific use case where the assistant could propose concrete plans by, for instance, finding a suitable time in the calendar for practising a skill the user would like to master and then allocating the appropriate number of time slots.

More generally, the IPA was imagined by the participants to be a catalyst for self-reflection by offering several distinct levels, or *stages*, of time management. These stages can be thought of as a set of concentric circles, where the inner-most one would be the most rudimentary support requiring the least amount of data about the user. That would mean, for instance, only allowing the user to enunciate their goals for the week and having the assistant repeat those back or create a corresponding to-do list. Notably, this would not even require the assistant to access the user's calendar. As PA-S252 put it: "Then the next step is an agent that will help you reflect on what's going on in your calendar, give you maybe some data-driven analysis of it, but you're still fully in control of any action to do with that calendar. It may not even make direct suggestions. All it's doing is helping the reflective capability. The next step may be a suggestion, yeah, but you're still the actor, the performer on it". A yet further stage could then conceivably include an even greater IPA autonomy, so that the agent may, for instance, automatically postpone the evening jog reminder if the user's location shows that they left their workplace late. However, some participants were concerned that the IPA could overemphasise productivity thus negatively affecting the user. As PG-S74 put it: "it could be very unhealthy for people who have control over their time in some sense, right, but who are 'list-people' and who are perfectionists and want to get everything done and just become overwhelmed by /. . . / constantly being told 'No, hang on, you didn't do that' or 'We will reschedule it to that slot tomorrow'".

4.2 Multiparty Interaction

The theme of multiparty interaction comprised of discussed scenarios where the IPA would interact not only with its main user but with other humans and their own IPAs as well, in order to schedule meetings or facilitate collaboration. Thinking again in terms of stages, it was envisioned that a simpler stage could be to offer an obfuscated view of the calendar with approved people. A more data-intrusive and technologically more challenging feature would be allowing the IPA to infer the social relationships of the user and to adjust both what information to share and with whom. The importance of model explainability for such potentially socially consequential features was repeatedly stressed. PB-S207 cautioned that users might refrain from having the IPA delegate arranging their meetings in order not to give the other party the impression that they were not worthy of a personal invitation. Contrastingly, should the system become widespread, those unable or unwilling to use the IPAs for time management might be disadvantaged through technological exclusion. Lastly, PD-S166 warned of accidental multiparty interaction. The given example was that the IPA might notify the user of a sensitive medical appointment via loudspeakers while the user has guests in the house. Proposed remedies included privacy-sensitive or – more challengingly and intrusively – context-sensitive adaptive modality selection.

4.3 Human-Assistant Interaction

A significant part of the human-assistant interaction discussion focused on dialogue design and modality selection. PD-S46 proposed that the assistant could look for contextual cues when choosing the output modality. The proposed cues for voice interaction were situations when the user's hands or eyes are focused on another task (e.g. driving, cleaning, assembling furniture). Moreover, PD-S200 underlined that personalisation should ideally also allow the user to specify when they would prefer the assistant to communicate with them using voice output or any other modality. Regarding dialogue design itself, PC-S199 believed the assistant "should also keep track of the social norms where the agent is deployed" and that it "shouldn't be lying". PD-S204 believed communication norms between humans and IPAs qualitatively differed from interhuman conversation.

PD-S200 proposed to differentiate between proactivity within and outside of ongoing dialogues. The former is reified in the IPA's responses within ongoing dialogues, whereas the latter comprises the actions the assistant performs from a "dormant state", such as rescheduling or initiating a new dialogue itself. The participant also noted that both forms of proactivity would ideally be personalised to the user and adapted for the relevant context. The importance of scrutability in this aspect was stressed by multiple participants (e.g. PB-S46, PE-S151). Furthermore, PD-S86 emphasised that the user should be able to adjust how proactive the IPA is in adjusting their calendar: "So imagine, say, for instance, if you're writing a paper [and] you're thinking about something you're doing, and then your agent is going 'Your calendar has been changed!' and it takes you out of your primary task and into that secondary task there".

The perception the user would have of the assistant and its capabilities was discussed. The participants (statements 233–251) explicitly associated the work of Roger K. Moore with the role of

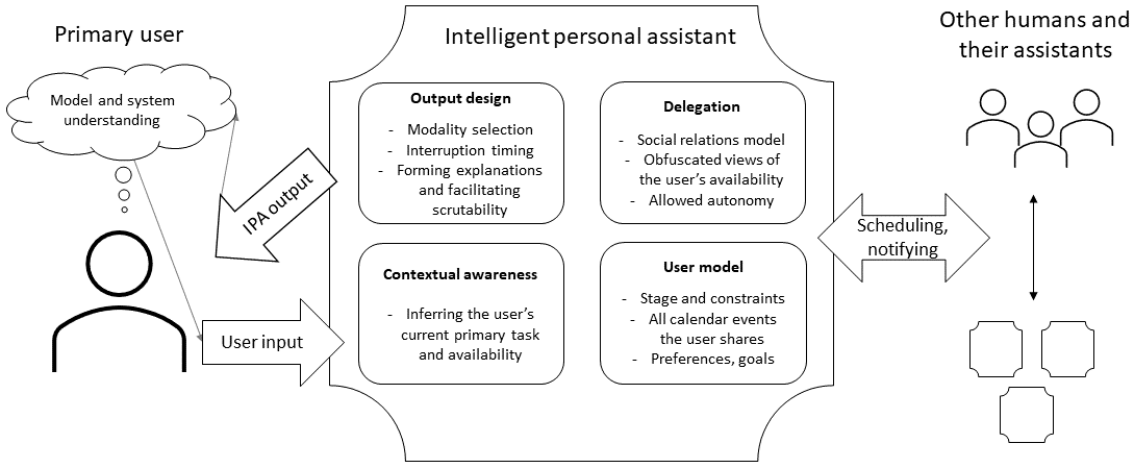

Figure 2: Diagram of the envisioned IPA and its interactions

voice in human-assistant interaction. Specifically, they noted that the degree of common understanding depends both on the assistant's user model and the user's "internal model" of the assistant. The participants also noted that the user's perception of the IPA would change with usage and improve or deteriorate depending on how successful it was in understanding and fulfilling tasks. Furthermore, bad communication design (inappropriate voice, modality choice, dialogue design) could lead to misaligned expectations and abandoning of the technology, as "we assume that it will do very little [for us]" (PD-S249). Regarding the usage willingness itself, PG-S96 hypothesised that it would also depend on one's general openness towards new technologies and that it would span to encompass those who would likely eagerly adopt an automated time management system for both private and personal tasks, over those who would prefer to keep their off-work calendars completely private, to those "who would really need to be convinced that this will help". The *stages* combined with scrutability were offered as a potential remedy for both the perceived lack of system utility and the general misalignment of perceived and actual system capabilities. As described by PA-S192, the user could "test-drive" more advanced system capabilities through hypothetical scenarios showcasing possible functionalities.

4.4 User Modelling

The points the participants raised regarding user modelling can be split into two categories: data stewardship and learning mechanisms. Regarding the former, a recurring question (e.g. PG-S72, PB-S103, PF-S154) was the user's willingness to provide data, given the complex capabilities that were envisioned for the system (urgency and task complexity approximation, modality switching, inferring time constraints, understanding social and physical context, tracking attention and cognitive load) as well as the misuse risk. In that sense, PG-348 proposed all data collected about the user be tied to licences, which would also limit what the data may be used for. Participants agreed (e.g. PB-46, PD-S137, PG-S188) that the scrutability of data collection, sharing, and use becomes more pronounced when the assistant is making decisions on the user's behalf, e.g. by rescheduling events or filtering incoming or outgoing information. As PB-S46 put it: "If there's some sort of preference or there's some model being built – who and how is that managed and owned? Because that's potentially very intimate information. So, the actual act of managing or scrutability or transparency there or explainability is really important".

There have been several proposed mechanisms through which the assistant could learn about the user's time management preferences, ever-changing social structures, and the broader context. PE-S202 and PA-S203 contrasted data-driven learning and modelled understanding, noting that a degree of prescribed instructions would likely be needed and debating the appropriateness of constraint-based learning for scheduling. Specifically, it was suggested that the assistant may keep a list of "hard" (i.e. never to be done) and "soft" (preferred) constraints. PB-S191 believed sharing sets of constraints to be a potent control mechanism, as it would allow control of both data usage and time assistance preferences. However, PB-S185 also accentuated that constraints alone are limited by humans' incapability to imagine all potential situations (not) bound by the constraints. An alternative learning mechanism dubbed "proactive explainability" was proposed by PF-S320, whereby the assistant would proactively propose an action to the user and offer an explanation for its suggestion. Through case-based reasoning and after enough similar situations, the IPA would potentially act without asking for prior validation. However, PG-322 deemed such an approach too risky, given how complex the context of interhuman interactions may be.

5 PROPOSED INTERACTION DIAGRAM

Based on the sketches the participants made in the last phase of the discussion, diagram in figure 2 was created. It emphasises the main functionalities discussed and highlights the importance of the user's perception of the IPA.

6 LIMITATIONS AND FUTURE WORK

The rather open-ended structure of the study had been intentionally chosen to allow the experts to voice and debate the points they

considered most important given their respective backgrounds. This indeed led to intriguing insights. Nevertheless, this structure did not allow for a targeted probing of specific ideas. Future work with a more in-depth focus on individual time management is thus desirable. Moreover, since all participants were academics, further studies might consider using time management IPA prototypes and prompts to investigate the views of the general public, akin to studies in [6, 14, 15, 25, 29] that investigated existing or imagined general-purposed IPAs.

7 CONCLUSION

This work presents the results of an expert discussion on proactive IPAs for time management thereby supplementing guidelines on proactive general IPAs [16, 29] and contributing towards a set of general principles for conversational interfaces [19]. Specifically, the proposed onboarding in stages and scrutability may help remedy both the misalignment of user perception of the system capabilities as well as the concern about data use. Due to partially distinct sets of challenges and use cases, our findings suggest that individual time management may be considered separately from assisting in multi-party interactions. Affecting both, however, are unresolved challenges concerning data stewardship, interruption timing, and modality selection.

ACKNOWLEDGMENTS

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number 18/CRT/6223. This research was conducted with the financial support of Science Foundation Ireland under Grant Agreement No. 13/RC/2106_P2 at the ADAPT SFI Research Centre at Trinity College Dublin. ADAPT, the SFI Research Centre for AI-Driven Digital Content Technology, is funded by Science Foundation Ireland through the SFI Research Centres Programme. For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

REFERENCES

[1] Andreas Balaskas, Stephen M Schueller, Anna L Cox, and Gavin Doherty. 2022. Understanding users' perspectives on mobile apps for anxiety management. *Frontiers in Digital Health* (2022), 178.

[2] Pauline M. Berry, Thierry Donneau-Golencer, Khang Duong, Melinda Gervasio, Bart Peintner, and Neil Yorke-Smith. 2017. Evaluating intelligent knowledge systems: experiences with a user-adaptive assistant agent. *Knowledge and Information Systems* 52, 2 (2017), 379–409. https://doi.org/10.1007/s10115-016-1011-3

[3] Judith Borghouts, Elizabeth Eikey, Gloria Mark, Cinthia De Leon, Stephen M Schueller, Margaret Schneider, Nicole Stadnick, Kai Zheng, Dana Mukamel, Dara H Sorkin, et al. 2021. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. *Journal of medical Internet research* 23, 3 (2021), e24387.

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa doi: 10.1191/1478088706qp063oa.

[5] Leigh Clark, Philip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schlögl, Jens Edlund, Matthew Aylett, João Cabral, Cosmin Munteanu, Justin Edwards, and Benjamin R Cowan. 2019. The State of Speech in HCI: Trends, Themes and Challenges. *Interacting with Computers* 31, 4 (2019), 349–371. https://doi.org/10.1093/iwc/iwz016

[6] Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What can i help you with?": infrequent users' experiences of intelligent personal assistants. , Article 43 pages. https://doi.org/10.1145/3098279.3098539

[7] Justin Cranshaw, Emad Elwany, Todd Newman, Rafal Kocielnik, Bowen Yu, Sandeep Soni, Jaime Teevan, and Andrés Monroy-Hernández. 2017. Calendar.Help: Designing a Workflow-Based Scheduling Agent with Humans in the Loop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2382–2393. https://doi.org/10.1145/3025453.3025780

[8] Allan de Barcelos Silva, Marcio Miguel Gomes, Cristiano André da Costa, Rodrigo da Rosa Righi, Jorge Luis Victoria Barbosa, Gustavo Pessin, Geert De Doncker, and Gustavo Federizzi. 2020. Intelligent personal assistants: A systematic literature review. *Expert Systems with Applications* 147 (2020), 113193. https://doi.org/10.1016/j.eswa.2020.113193

[9] Eduardo Islas-Cota, J. Octavio Gutierrez-Garcia, Christian O. Acosta, and Luis-Felipe Rodríguez. 2022. A systematic review of intelligent assistants. *Future Generation Computer Systems* 128 (2022), 45–62. https://doi.org/10.1016/j.future.2021.09.035

[10] Pranut Jain, Rosta Farzan, and Adam J. Lee. 2022. Laila is in a Meeting: Design and Evaluation of a Contextual Auto-Response Messaging Agent. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1457–1471. https://doi.org/10.1145/3532106.3533493

[11] Jovan Jeromela. 2022. Scrutability of Intelligent Personal Assistants. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization* (Barcelona, Spain) *(UMAP '22)*. Association for Computing Machinery, New York, NY, USA, 335–340. https://doi.org/10.1145/3503252.3534355

[12] Judy Kay. 1998. *A scrutable user modelling shell for user-adapted interaction.* Doctoral dissertation. Basser Department of Computer Science, Faculty of Science, University of Sydney.

[13] Judy Kay and Bob Kummerfeld. 2013. Creating Personalized Systems That People Can Scrutinize and Control: Drivers, Principles and Experience. *ACM Trans. Interact. Intell. Syst.* 2, 4, Article 24 (jan 2013), 42 pages. https://doi.org/10.1145/2395123.2395129

[14] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. , 5286–5297 pages. https://doi.org/10.1145/2858036.2858288

[15] Michal Luria, Rebecca Zheng, Bennett Huffman, Shuangni Huang, John Zimmerman, and Jodi Forlizzi. 2020. Social Boundaries for Personal Agents in the Interpersonal Space of the Home. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376311

[16] Christian Meurisch, Maria-Dorina Ionescu, Benedikt Schmidt, and Max Mühlhäuser. 2017. Reference model of next-generation digital personal assistant: integrating proactive behavior. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 149–152.

[17] Roger K. Moore. 2017. *Is Spoken Language All-or-Nothing? Implications for Future Speech-Based Human-Machine Interaction.* Springer Singapore, Singapore, 281–291. https://doi.org/10.1007/978-981-10-2585-3_22

[18] Roger K. Moore and Mauro Nicolao. 2017. Toward a Needs-Based Architecture for 'Intelligent' Communicative Agents: Speaking with Intention. *Frontiers in Robotics and AI* 4 (2017). https://doi.org/10.3389/frobt.2017.00066

[19] Christine Murad and Cosmin Munteanu. 2019. "I Don't Know What You're Talking about, HALexa": The Case for Voice User Interface Guidelines. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 9, 3 pages. https://doi.org/10.1145/3342775.3342795

[20] Martin Porcheron, Joel E. Fischer, Stuart Reeves, and Sarah Sharples. 2018. *Voice Interfaces in Everyday Life.* Association for Computing Machinery, Paper 640. https://doi.org/10.1145/3173574.3174214

[21] Yu Sun, Nicholas Jing Yuan, Yingzi Wang, Xing Xie, Kieran McDonald, and Rui Zhang. 2016. Contextual Intent Tracking for Personal Assistants. , 273–282 pages. https://doi.org/10.1145/2939672.2939676

[22] Yu Sun, Nicholas Jing Yuan, Xing Xie, Kieran McDonald, and Rui Zhang. 2017. Collaborative Intent Prediction with Real-Time Contextual Data. *ACM Trans. Inf. Syst.* 35, 4 (2017), Article 30. https://doi.org/10.1145/3041659

[23] Carlos Toxtli, Andrés Monroy-Hernández, and Justin Cranshaw. 2018. Understanding Chatbot-Mediated Task Management. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3173574.3173632

[24] Sruthi Viswanathan, Fabien Guillot, Minsuk Chang, Antonietta Maria Grasso, and Jean-Michel Renders. 2022. Addressing Hiccups in Conversations with Recommender Systems. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1243–1259. https://doi.org/10.1145/3532106.3533491

[25] Sarah Theres Völkel, Daniel Buschek, Malin Eiband, Benjamin R. Cowan, and Heinrich Hussmann. 2021. *Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants.* Association for Computing Machinery, Article 254.

https://doi.org/10.1145/3411764.3445536

[26] Brian Whitworth. 2005. Polite Computing. *Behaviour & Information Technology* 24, 5 (2005), 353–363. https://doi.org/10.1080/01449290512331333700 arXiv:https://doi.org/10.1080/01449290512331333700

[27] Jun Xiao, Richard Catrambone, and John T Stasko. 2003. *Be Quiet? Evaluating Proactive and Reactive User Interface Assistants.* Technical Report. Georgia Institute of Technology.

[28] Neil Yorke-Smith, Shahin Saadati, Karen L Myers, and David N Morley. 2012. The design of a proactive personal agent for task management. *International Journal on Artificial Intelligence Tools* 21, 01 (2012), 30. https://doi.org/10.1142/S0218213012500042

[29] Nima Zargham, Leon Reicherts, Michael Bonfert, Sarah Theres Voelkel, Johannes Schoening, Rainer Malaka, and Yvonne Rogers. 2022. Understanding Circumstances for Desirable Proactive Behaviour of Voice Assistants: The Proactivity Dilemma. In *Proceedings of the 4th Conference on Conversational User Interfaces* (Glasgow, United Kingdom) *(CUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 3, 14 pages. https://doi.org/10.1145/3543829.3543834