

Designing Proactive Voice Interfaces: Key Factors for Office Settings

Shashank Ahire¹, Michael Rohs¹

¹Human-Computer Interaction, Leibniz University Hannover, Germany

Abstract

Although proactive voice assistants (VAs) have proven beneficial for health and well-being, they face several challenges, primarily intrusiveness, lack of user agency, and privacy concerns. In office settings, knowledge workers often encounter health and well-being issues while at work, making it essential to design VAs that address these needs. However, how to best design proactive VAs for formal, sensitive and social settings like offices, remains an open question. Drawing from our experiences and findings in developing a VA for workplace settings, this paper presents eight key factors for designing VAs for office environments.

Keywords

voice assistant, office, proactive, voice interface, health, well-being, knowledge worker

1. Introduction

Conversational agents have been studied in multiple use cases in the office, such as to-do lists [1], scheduling [2], work reflection [3], performing focused work [4], managing distractions [5], scientific work [6], and health and well-being [7, 8, 9, 10]. VA for health and well-being are impactful, but also have challenges such as maintaining user's privacy, not disturbing colleagues, identifying the availability of users, and initiating interaction only during their availability.

Current mainstream VAs such as *Google Assistant*, *Alexa* and *Siri* support passive interaction, i.e., users have to explicitly initiate an interaction with the VA. However, in recent years there have been attempts by researchers to develop proactive VAs for various use cases such as automotive [11], home [12], and work [10, 9, 8, 13] settings. Previous research has found negative experiences with proactive voice assistance especially with regard to privacy [12, 9], intrusiveness [14], loss of control [14], frustration, and abrupt termination [15].

In office settings, which are often sophisticated, sensitive, and formal, voice interventions for the health and well-being of knowledge workers must be designed with consideration for office protocols and etiquette. Consequently, designing proactive voice interventions for the work environment is challenging.

Based on our experience in designing and developing a VA for the health and well-being of knowledge workers [8, 9, 10] we outline eight key factors for designing VAs to support the health and well-being of knowledge workers: (1) Agency over voice interventions; (2) Prioritizing interventions; (3) Structuring voice interventions; (4) Opportune and inopportune moments for voice interventions; (5) Identifying and adapting to user context (6) Harnessing multi-modality for proactive interventions; (7) Developing context-aware multi-level interventions; (8) Relationships with colleagues. In the following we discuss each of these factors. We believe that considering these factors reduces the likelihood of negative experiences with proactive VAs in work environments.

Joint Proceedings of the ACM IUI Workshops 2025, March 24-27, 2025, Cagliari, Italy

*Corresponding author.

† These authors contributed equally.

✉ shashank.ahire@hci.uni-hannover.de (S. Ahire); michael.rohs@hci.uni-hannover.de (M. Rohs)

🌐 <https://shashankahire.com> (S. Ahire); <https://www.pi.uni-hannover.de/de/hci/prof-michael-rohs> (M. Rohs)

🆔 0000-0001-5930-79533 (S. Ahire); 0000-0003-4599-729X (M. Rohs)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

2. Background

2.1. VAs in Office Settings

In our initial work [10], we conducted semi-structured interviews with knowledge workers, who expressed a desire for a conversational agent (CA) to assist them with health, well-being, and productivity. Based on these insights, we developed a *Ubiquitous Work Assistant* (UWA) comprising a stationary CA (S-CA) and a wearable CA (W-CA). The UWA was designed to offer flexibility, enabling knowledge workers to move around their work environments while maintaining seamless interaction with the system.

Subsequently, we developed *WorkFit* [9], a voice assistant designed to support the health and well-being of knowledge workers. WorkFit monitors the behavior of knowledge workers and delivers interventions in cases of sedentary activity, inconsistent hydration, and stress-related issues within office environments. The findings highlighted challenges such as providing users with a sense of agency, minimizing intrusiveness for colleagues, and managing privacy concerns.

To address the lack of user agency in proactive voice interactions, we introduced a novel approach called *dual-mode intervention* [8]. This technique begins with non-intrusive primary interventions using different modalities (such as graphical interface, vibration, auditory icon, and ring), resorting to voice interventions only when the user does not respond to the initial prompts. Our evaluation revealed that the dual-mode technique was strongly favored over direct voice outputs in proactive VAs. The UEQ ratings indicated significant improvements in dimensions such as enjoyment, supportiveness, efficiency, pleasantness, and practicality.

Lastly, many knowledge workers struggle with insufficient hydration during busy work hours. We developed the *H2Office* system [13], which includes a water gauge and a smartwatch app. The system utilizes a combination of auditory icons and vibrations to deliver hydration interventions. This approach effectively prompted users to hydrate without causing distractions during work.

2.2. Proactive VAs

A survey by Schmidt and Braunger [16], with 1,550 participants, revealed that users highly value proactivity as a key feature of voice assistants. Similarly, a study conducted by Völkel et al. [17] on user's envisioned interactions with an ideal voice assistant highlighted that many participants favored proactive behavior in voice assistants.

Proactive VAs have been studied in limited home settings [12, 14, 15]. To build a proactive VA, researchers have augmented current VA with microphones and cameras. Such augmentations increase a VA's capability to identify user context, user movements, and social presence, but leads to privacy concerns among users [12]. Reicherts et al. [14] investigated the user perceptions of proactive VAs. Participants in the study raised concerns related to privacy, intrusiveness, loss of control, and interference with social settings.

Similarly, to determine opportune moments for proactive conversational interactions in domestic settings Cha et al. [12] conducted a study with 40 participants. They concluded that opportune moments are dependent on user busyness, complexity of the primary task, user mood, and social availability. They found that opportune moments for smart speaker interaction are those when users are engaged in repetitive activities that require low attention.

Zargham et al. [18] used storyboards illustrating various scenarios to explore user perceptions of proactive smart speaker interventions in daily life. The study focused on eight common scenarios during interviews, with participants expressing that proactive health-related interventions were particularly valuable to them. In another study Zargham et al. [19] investigated use of humor for in proactive voice interaction, they found use of humor to be complex and subjective. They emphasized on thoughtful and personalized humor integration to enhance the desirability of proactive interventions.

Komori et al. [20] designed a prototype to identify opportune moments for delivering voice information services without disrupting user's activities. The system detected user activity transitions by monitoring location and body movements, then assessed notification acceptability at specific moments.

Their findings indicated that participants were more receptive to notifications while in bed, eating, doing household chores, or working. In a separate study, Wei et al. [21] observed that participants were more available for conversations during entertainment activities compared to work or study tasks. Furthermore, self-reported measures of boredom and mood were found to have a significant correlation with participants' availability.

3. Key Factors

Below, we outline eight key factors for designing proactive voice interventions aimed at promoting the health and well-being of knowledge workers.

3.1. Agency Over Voice Interventions

The importance of agency is widely acknowledged across human-computer interaction guidelines [22, 23]. However, in voice interactions, users have reported the issue of minimal or no agency during interaction [14]. In passive smart speaker interactions, such as with music, search, or controlling smart home devices, agency may not matter as much since no personal information is at stake. However, in the case of health and well-being in office settings, users should have agency during voice interactions. Serious privacy issues may arise if the information is delivered in inappropriate situations, which could make users uncomfortable to engage in voice interactions altogether. Moreover, in formal and sensitive settings it becomes crucial for knowledge workers to have agency over voice interactions, which would allow them to stop, delay or discard the proactive voice intervention. In our previous work on a dual-intervention approach [8] we attempted to give agency to knowledge workers by utilizing multiple modalities in combination with voice interventions.

3.2. Prioritizing Interventions

We propose prioritizing health interventions based on urgency parameters. In the work setting it is essential to not distract knowledge workers during inopportune moments. Delivering multiple non-urgent notifications could annoy the knowledge worker and spoil the user experience.

In the domain of health and well-being there are multiple types of interventions addressing stress, sedentary behavior, hydration, and stretching [9, 8, 13]. Although all of these health and well-being interventions are essential and should be acted upon eventually, some interventions can be acted upon later, while some should be acted upon immediately. For instance, stress-related interventions should be delivered as voice interventions to urgently gain the user's attention. However, interventions related to hydration, sedentary behavior, and stretching do not need to be acted upon immediately. Prioritizing the delivery of interventions based on urgency would allow the VA to gain knowledge workers attention only if necessary.

3.3. Structuring Voice Interventions

In WorkFit [9], we designed voice interventions based on three principles: (1) user-centered dialogue; (2) reflection on health; and (3) motivation. The voice interventions were as follows:

- Stress intervention: *Your heartbeat is too high/low. Would you like to perform a breathing exercise? If so, please say yes.*
- Sedentary behavior: *You haven't moved for a while. How about taking a break? Say yes to confirm.*
- Hydration intervention: *It's time to drink a cup of water.*

However, during the evaluation, we found that while these interventions were suitable for delivering daily health interventions at home, participants noted that in workplace settings, the interventions was

obvious and direct. As a result, they could be easily interpreted and understood by colleagues, leading to concerns about privacy. Moreover, these interventions were lengthy and attention consuming. We resolved these issues of privacy, intrusion, and obviousness by restructuring the voice statements in the following manner:

- **Friendly:** *“Hey, would you like to perform a short mindfulness exercise?”* Friendly interventions suitable to the social setting in the office. It is polite, indirect, and non-intrusive.
- **Suggestive:** *“Do you feel stressed?”*, *“Do you want to relax for a while?”* These interventions are short and direct, hence can be used at one’s personal desk. Without consuming too much attention.
- **Encouraging:** *“Take a deep breath and unwind.”*, *“Make yourself a cup of tea and relax.”* These statements are private, indirect and encourage the user to act on the recommendation, without delivering private information to the colleagues. The statement is crafted to resemble a general daily recommendation suitable for a casual social setting.
- **Directive:** *“Time for a break—get up and move around!”*, *“Time to Stand!”*, *“Stand up and take a step!”* These interventions can be delivered when the user has ignored the health interventions multiple times. However, these interventions sound very direct and may come across as impolite. Thus, they should only be used for urgent and emergency situations.
- **Goal:** *“It’s time to drink your 200 ml cup of water.”* Goal-suggestive interventions can be used when the knowledge worker need to be told about the recommendation to be performed. Information provided in the statement could be encouraging for some users, whereas it could also be considered as obvious and direct regarding privacy.

3.4. Opportune and Inopportune Moments for Voice Intervention

During a proactive interaction, it is important to consider if the moment is opportune or inopportune. ‘Opportune moments’ are those that are suitable for having a voice interaction with a user. In contrast, ‘inopportune moments’ are ones that are not suitable for having a voice interaction. During an opportune moment, the likelihood that users engage in a voice interaction is high. Kang et al. [24] found that opportune moments are dependent on three user-related factors: personal state (mood, busyness), movement (entrance, departure, physical transitions), and social presence (presence of other persons).

Studies have attempted to find opportune moments for initiating proactive interactions [12, 15]. We advocate that it is not only important to consider opportune moments, but to also have a mechanism to avoid voice interactions during inopportune moments for example, during spontaneous meetings). Considering inopportune moments is crucial in work settings. Not considering inopportune moments in office settings during a proactive interaction may lead to awkward moments. Due to instances of embarrassment, awkwardness, and privacy concerns, it is essential to consider inopportune moments.

3.4.1. Techniques for Identifying Opportune and Inopportune Moments

Below, we outline widely available tools that can assist in identifying opportune and inopportune moments.

- **Calendar: Monitoring the user activity.** In our previous work we found that knowledge workers’ daily activities are influenced by events in their calendar. Their availability and unavailability in a particular moment is strongly related to the event in the user’s calendar at that particular moment. Hence, the calendar is an important source of information for identifying opportune and inopportune moments to initiate a proactive interaction. However, the calendar is not sufficient, because not all events are available in the calendar, such as spontaneous meetings with colleagues or phone calls.

- **Wi-Fi fingerprinting: Identifying user location.** Work settings have a strict division of places for casual chats (pantry, social places/couches, corridors), formal discussion (conference room, personal desk), private one-to-one discussions (cabin). Using Wi-Fi positioning techniques, such as fingerprinting [25] can be used to track the position of a person in an office. The user's availability and mood (e.g., relaxed in the pantry or focused at their desk or in a conference room) are often associated with specific locations. As such, tracking the user's location within the office can provide insights into their likelihood of being available for interaction.
- **Voice recognition: Identifying spontaneous interactions.** Spontaneous meetings, conversations, and discussions are challenging to schedule in advance on a calendar. Therefore, a microphone and voice detection of 3-5 seconds can help to detect if the user is engaged in a conversation. If so, the proactive interface can refrain from engaging in an interaction.
- **Bluetooth: Monitoring the number of colleagues.** Bluetooth can be used to estimate the number of colleagues present at a particular setting in the office. Based on this information, the VA can decide to deliver the voice intervention immediately or to delay it. Such tracking can help to minimize annoying and embarrassing moments in proactive interactions.

To effectively identify various contexts, a combination of these tools can be utilized. Below, we discuss how a combination of these tools can be applied for context identification in a work environment.

3.5. Identifying and Adapting to User Context

On a typical work day, the knowledge worker tends to spend time in various contexts, such as at a personal desk, a pantry, a conference room, in serendipitous corridor talks, in spontaneous meetings, and in restrooms. Different locations are associated with different purposes and etiquette's for interaction. For initiating a proactive interaction in the office, it is essential to identify the user context and to deliver the intervention according to the etiquette of the context. In the following sections, we discuss various contexts and techniques that could be used to identify those contexts before proactively initiating an interaction.

- **Spontaneous Meetings:** Short spontaneous meetings are common in work settings. Generally, spontaneous meetings typically involve from 2 to 4 colleagues. Spontaneous meeting are often based at a desk or cabin. To identify a spontaneous meeting, (1) VAs can use voice detection for 3-5 seconds. If human voice is detected, the VA can assume that a spontaneous meeting is ongoing and switch the mode of intervention to other modalities as mentioned in the multi-level intervention table (Table 1).
- **Pantry:** Social settings such as pantries are common in office settings. In this setting, knowledge workers tend to be relaxed and have casual discussions. In this setting they would generally refrain from discussing intensive work topics. A VA could identify the pantry context by (1) Using Wi-Fi fingerprinting to identify the location as pantry; (2) using Bluetooth to identify the number people present at the pantry; and (3) using voice recognition if user is participating in the discussion.
- **Personal Desk:** The personal desk is the space in which a knowledge worker spends the most of the time. Generally the personal desk is placed in a cabin that is shared among 2-3 colleagues of the same hierarchy level. In some offices personal desks can also be placed in an open-plan office, which is generally shared by 10-20 persons. Although the desk is considered as a personal space, voice interaction here could be disturbing for adjacent colleagues. Hence, the VA should be careful while delivering direct interventions.

To recognize if the knowledge worker is present at the personal desk, the VA can use (1) Bluetooth signal strength of the laptop, (2) Wi-Fi fingerprinting to identify the location of the user, (3)

the calendar to check if the knowledge worker has a commitment at the moment, and (4) voice detection to detect if the user is in an online meeting.

- **Conference Rooms:** Conference rooms typically host important meetings and discussions among knowledge workers. In a conference room typically there are knowledge workers of multiple hierarchy levels. During voice interaction the VA has to take multiple factors into consideration. Knowledge workers may be completely reluctant to communicate with the VA in such a setting. In such situations the VA has to consider delivering health and well-being interventions using multi-modal (subsection 3.6) and multi-level (subsection 3.7) interventions.

Generally, conference rooms are big and dedicated spaces for seating 10-15 people. To identify the context of the conference room, (1) the location can be easily recognized using Wi-Fi fingerprinting. (2) As meetings in conference rooms are important, it is likely that knowledge workers enter such an event in the calendar. Hence, the location field in the calendar can be used to identify the location as a conference room. (3) Further, using Bluetooth the VA can recognize the number of employees present at the situation.

- **Restroom:** Although restroom visits are brief, knowledge workers tend to visit them regularly. However, they may not be inclined to engage in voice interactions in such settings. Therefore, the VA should rely on multimodal interventions (subsection 3.6). Restrooms can be effectively identified using the Wi-Fi fingerprinting method.

Moreover, VAs for the health and well-being of knowledge workers should adapt to the identified context. Adaptive VAs have been well-investigated in the domain of automotive voice interfaces [11]. Similarly for the office setting, the delivery of voice interventions should consider the user's environment. Adaptive VAs should consider multiple factors, including the context, the urgency of the intervention, the user's relationship with the person involved, and opportune or inopportune moments. The adaptation can be realized in multiple ways, such as changing the voice intervention structure (subsection 3.3) and switching modalities of interventions (subsection 3.6).

Each setting has distinct etiquette. Professional settings, such as offices, consist of highly formal environments (e.g., conferences and meetings) and informal ones (e.g., corridor talks and lunch/coffee breaks). In the case of health and well-being, especially in sensitive settings, harnessing sensors and applications available in smart devices – such as calendars, GPS, Wi-Fi, voice, gyroscopes, and Bluetooth – can enable VAs to adapt to their surroundings effectively.

3.6. Harnessing Multi-Modality for Proactive Interventions

In office settings, direct voice interventions without user control can be counterproductive. In our previous study [8], many participants preferred to use vibration, auditory icons, and GUIs as the primary mode and voice as the secondary mode for the intervention. Currently, multimodality among VAs is limited to GUIs, we propose utilizing audio notifications, such as auditory icons and rings, to deliver reminders. Instead of delivering voice statements as the first reminder, VAs could initially use auditory icons and rings as the primary mode of intervention to gain the attention of user, followed by voice statement interventions.

3.7. Developing Context-Aware Multi-Level Interventions

Current VAs primarily rely on direct voice interventions. However, we believe that implementing multi-level interventions could be more effective for settings like offices. For multi-level interventions, VAs can utilize modalities such as auditory cues (e.g., auditory icons or rings) and vibrations to deliver messages subtly. Depending on the context and urgency of the intervention, an appropriate modality can be selected to ensure timely and context-sensitive delivery.

The selection of these modalities should be context-dependent. For example, during spontaneous meetings, vibrations can be used for non-urgent interventions, while auditory cues can be used for

urgent ones. In a pantry setting, interventions could include a combination of GUI notifications, vibrations, or rings. If the user repeatedly ignores these cues, the VA could escalate to delivering a friendly voice intervention.

Building on our experience in developing VAs for office settings, we have designed multi-level interventions with up to three distinct levels (see Table 1).

Table 1

Context-dependent multi-level interventions. GUI: graphical, Vi: vibration, DV: direct voice, Ai: auditory icon, R: ring.

Level of Intervention	Urgent (Stress)		Non-urgent (Hydration, Sedentary, Stretching)		
	First	Second	First	Second	Third
Spontaneous Meetings	Vi, GUI, Ai, R	DV	Vi, GUI	Ai, R	DV
Conference room	Vi, GUI, Ai	DV, R	Vi, GUI	Ai, R	DV
Desk	Vi, GUI, Ai, R, DV	-	Vi, GUI	Ai, R, DV	-
Pantry	GUI, Vi, Ai	Ring, DV	Vi, GUI	Ai, R	DV
Restroom	GUI, Vi	Ai, R, DV	Vi, GUI	Ai, R	DV

3.8. Relationship with Colleagues

To identify the relationship with colleagues, a VA could utilize user-assigned relationship tags, which categorize the interactions with a person. These tags, manually assigned by the user, could range from professional to more personal designations, such as manager, colleague, team member, or supervisor. By associating specific tags with individuals, the VA can adjust its voice intervention structure accordingly—for example, using formal language with a supervisor, a more friendly tone with a colleague, and possibly even empathetic responses if a personal relationship is indicated. This user-driven approach ensures that the VA delivers responses tailored to the nature of the relationship, enhancing both personalization and usability.

To implement relationship tags and restructure voice interventions accordingly, the VA can utilize the Bluetooth functionality of a colleague’s mobile devices to identify the associated person. By detecting nearby Bluetooth signals, the VA can infer the identity of individuals in proximity and adapt its interventions based on the predefined relationship tags associated with them.

4. Conclusion

In conclusion, designing proactive VAs is a complex task, and we believe that the proposed factors can serve as valuable guidance for VA practitioners and researchers in creating more effective proactive VAs for office environments. Furthermore, these factors may also prove beneficial in designing VAs for other sensitive settings, such as automobile, hospitals, educational institutions, and shared spaces. Drawing on insights from our prior research, this paper identifies and presents eight key factors that should be considered when designing VAs for office settings. We appeal to all the researchers and practitioners in the domain of CUIs to consider these factors in their design and development process of proactive VAs.

References

- [1] Y. Gil, V. Ratnakar, Towards intelligent assistance for to-do lists, in: Proceedings of the 13th International Conference on Intelligent User Interfaces, IUI '08, Association for Computing Machinery, New York, NY, USA, 2008, p. 329–332. URL: <https://doi.org/10.1145/1378773.1378822>. doi:10.1145/1378773.1378822.

- [2] E. Kimani, K. Rowan, D. McDuff, M. Czerwinski, G. Mark, A conversational agent in support of productivity and wellbeing at work, in: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), 2019, pp. 1–7. doi:10.1109/ACII.2019.8925488.
- [3] R. Kocielnik, D. Avrahami, J. Marlow, D. Lu, G. Hsieh, Designing for workplace reflection: A chat and voice-based conversational agent, in: Proceedings of the 2018 Designing Interactive Systems Conference, DIS '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 881–894. URL: <https://doi.org/10.1145/3196709.3196784>. doi:10.1145/3196709.3196784.
- [4] T. Grover, K. Rowan, J. Suh, D. McDuff, M. Czerwinski, Design and evaluation of intelligent agent prototypes for assistance with focus and productivity at work, in: Proceedings of the 25th International Conference on Intelligent User Interfaces, IUI '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 390–400. URL: <https://doi.org/10.1145/3377325.3377507>. doi:10.1145/3377325.3377507.
- [5] V. W.-S. Tseng, M. L. Lee, L. Denoue, D. Avrahami, Overcoming Distractions during Transitions from Break to Work Using a Conversational Website-Blocking System, Association for Computing Machinery, New York, NY, USA, 2019, p. 1–13. URL: <https://doi.org/10.1145/3290605.3300697>.
- [6] J. Cambre, Y. Liu, R. E. Taylor, C. Kulkarni, Vitro: Designing a voice assistant for the scientific lab workplace, in: Proceedings of the 2019 on Designing Interactive Systems Conference, DIS '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 1531–1542. URL: <https://doi.org/10.1145/3322276.3322298>. doi:10.1145/3322276.3322298.
- [7] Y. Luo, B. Lee, D. Y. Wohn, A. L. Rebar, D. E. Conroy, E. K. Choe, Time for Break: Understanding Information Workers' Sedentary Behavior Through a Break Prompting System, Association for Computing Machinery, New York, NY, USA, 2018, p. 1–14. URL: <https://doi.org/10.1145/3173574.3173701>.
- [8] S. Ahire, S. Othman, M. Rohs, Dual-mode interventions: Giving agency to knowledge workers in proactive health interventions, in: Proceedings of the 6th ACM Conference on Conversational User Interfaces, CUI '24, Association for Computing Machinery, New York, NY, USA, 2024. URL: <https://doi.org/10.1145/3640794.3665578>. doi:10.1145/3640794.3665578.
- [9] S. Ahire, S. Benjamin, M. Rohs, Workfit: Designing proactive voice assistance for the health and well-being of knowledge workers, in: ACM Conversational User Interfaces 2024 (CUI '24), July 8–10, 2024, CUI '24, Association for Computing Machinery, New York, NY, USA, 2024. URL: <https://doi.org/10.1145/3640794.3665561>. doi:10.1145/3640794.3665561.
- [10] S. Ahire, M. Rohs, S. Benjamin, Ubiquitous work assistant: Synchronizing a stationary and a wearable conversational agent to assist knowledge work, in: Proceedings of the 1st Annual Meeting of the Symposium on Human-Computer Interaction for Work, CHIWORK '22, Association for Computing Machinery, New York, NY, USA, 2022. URL: <https://doi.org/10.1145/3533406.3533420>. doi:10.1145/3533406.3533420.
- [11] S. Wen, S. Ping, J. Wang, H.-N. Liang, X. Xu, Y. Yan, Adaptivevoice: Cognitively adaptive voice interface for driving assistance, in: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI '24, Association for Computing Machinery, New York, NY, USA, 2024. URL: <https://doi.org/10.1145/3613904.3642876>. doi:10.1145/3613904.3642876.
- [12] N. Cha, A. Kim, C. Y. Park, S. Kang, M. Park, J.-G. Lee, S. Lee, U. Lee, Hello there! is now a good time to talk? opportune moments for proactive interactions with smart speakers, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 4 (2020). URL: <https://doi.org/10.1145/3411810>. doi:10.1145/3411810.
- [13] S. Ahire, A. Abdalrazak, M. Rohs, H2office: A smartwatch and water-gauge system for facilitating hydration of knowledge workers, in: Pervasive Computing Technologies for Healthcare, Springer Nature Switzerland, 2024. URL: https://doi.org/10.1007/978-3-031-85575-7_2.
- [14] L. Reicherts, N. Zargham, M. Bonfert, Y. Rogers, R. Malaka, May i interrupt? diverging opinions on proactive smart speakers, in: Proceedings of the 3rd Conference on Conversational User Interfaces, CUI '21, Association for Computing Machinery, New York, NY, USA, 2021. URL: <https://doi.org/10.1145/3469595.3469629>. doi:10.1145/3469595.3469629.
- [15] J. Wei, B. Tag, J. R. Trippas, T. Dingler, V. Kostakos, What could possibly go wrong when

- interacting with proactive smart speakers? a case study using an esm application, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22, Association for Computing Machinery, New York, NY, USA, 2022. URL: <https://doi.org/10.1145/3491102.3517432>. doi:10.1145/3491102.3517432.
- [16] M. Schmidt, P. Braunger, A survey on different means of personalized dialog output for an adaptive personal assistant, in: Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization, UMAP '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 75–81. URL: <https://doi.org/10.1145/3213586.3226198>. doi:10.1145/3213586.3226198.
 - [17] S. T. Völkel, D. Buschek, M. Eiband, B. R. Cowan, H. Hussmann, Eliciting and analysing users' envisioned dialogues with perfect voice assistants, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21, Association for Computing Machinery, New York, NY, USA, 2021. URL: <https://doi.org/10.1145/3411764.3445536>. doi:10.1145/3411764.3445536.
 - [18] N. Zargham, L. Reicherts, M. Bonfert, S. T. Voelkel, J. Schoening, R. Malaka, Y. Rogers, Understanding circumstances for desirable proactive behaviour of voice assistants: The proactivity dilemma, in: Proceedings of the 4th Conference on Conversational User Interfaces, CUI '22, Association for Computing Machinery, New York, NY, USA, 2022. URL: <https://doi.org/10.1145/3543829.3543834>. doi:10.1145/3543829.3543834.
 - [19] N. Zargham, L. Reicherts, V. Avanesi, Y. Rogers, R. Malaka, Tickling proactivity: Exploring the use of humor in proactive voice assistants, in: Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia, MUM '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 294–320. URL: <https://doi.org/10.1145/3626705.3627777>. doi:10.1145/3626705.3627777.
 - [20] M. Komori, Y. Fujimoto, J. Xu, K. Tasaka, H. Yanagihara, K. Fujita, Experimental study on estimation of opportune moments for proactive voice information service based on activity transition for people living alone, in: M. Kurosu (Ed.), Human-Computer Interaction. Perspectives on Design. HCII 2019, volume 11566 of *Lecture Notes in Computer Science*, Springer, Cham, 2019. doi:10.1007/978-3-030-22646-6_39.
 - [21] J. Wei, T. Dingler, V. Kostakos, Understanding user perceptions of proactive smart speakers, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 5 (2022). URL: <https://doi.org/10.1145/3494965>. doi:10.1145/3494965.
 - [22] B. Shneiderman, Designing the User Interface: Strategies for Effective Human-Computer Interaction, 3rd ed., Addison-Wesley Longman Publishing Co., Inc., USA, 1997.
 - [23] J. Nielsen, 10 usability heuristics for user interface design, 2024. URL: <https://www.nngroup.com/articles/ten-usability-heuristics/>.
 - [24] S. Kang, H. Kim, Y. Noh, U. Lee, Poster: Toward context-aware proactive conversation for smart speakers, in: Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers, UbiComp '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 38–40. URL: <https://doi.org/10.1145/3460418.3479306>. doi:10.1145/3460418.3479306.
 - [25] S. Shang, L. Wang, Overview of WiFi fingerprinting-based indoor positioning, IET Communications 16 (2022) 725–733. URL: <https://doi.org/10.1049/cmu2.12386>. doi:10.1049/cmu2.12386.