Towards a Pattern Language for Smart Personal Assistants

ROBIN KNOTE, University of Kassel
MATTHIAS SÖLLNER, University of St.Gallen & University of Kassel
JAN MARCO LEIMEISTER, University of St.Gallen & University of Kassel

Supporting users in their daily activities, thus, making their lives more comfortable, has long been a goal for consumer-oriented systems development. With the rise of smart personal assistants (SPAs), however, we have reached a new milestone along the path towards this goal. These systems assist their owners by providing personalized and context-dependent information and service. Today's implementations reach from conversational agents, such as Siri, Cortana or Google Assistant, over chatbots, which are primarily text-based, to cognitive assistants, which assist according to a user's current cognitive or emotional state. However, although both research and practice proceed with full pace, recurring design elements of SPAs have not yet been investigated. We hence propose a pattern language for smart personal assistants to guide further empirical and design efforts. Therefore, we review existing information systems, computer science and human-computer interaction literature to find recurring design characteristics among 115 different assistants. The resulting pattern language contains 22 patterns that specify the interaction behavior and the intelligence of smart personal assistants.

Categories and Subject Descriptors: • **Human-centered computing~Personal digital assistants** • **Software and its engineering~Patterns** • **Software and its engineering~Design patterns**

Additional Key Words and Phrases: Smart Personal Assistants, Pattern Language

ACM Reference Format:

Knote, R., Söllner, M. and Leimeister, J.M. 2018. Towards a Pattern Language for Smart Personal Assistants. HILLSIDE Proc. of Conf. on Pattern Lang. of Prog. 25 (October 2018), 16 pages.

1. INTRODUCTION

Technology development rushes forward, thus offering unprecedented ways to reduce the complexity of our everyday lives. Right in the middle of this pace, recent advances in natural language processing and artificial intelligence have paved the way for smart personal assistants' (SPAs) success on the consumer market. Market forecasts predict the worldwide user count of SPAs such as Amazon Alexa, Apple's Siri or Microsoft Cortana to increase from 390 million in 2015 to 1.8 billion in 2021 which will result in 2.3 billion USD average sales growth per year (Tracitca, 2016). Thereby, an SPA is a system *"that uses input such as the user's voice [...] and contextual information to provide assistance by answering questions in natural language, making recommendations and performing actions"* (Baber, 1993, p. 223, p. 223). Thus, such kinds of digital assistants combine the comfort of intuitive natural language interaction with the utility of personalized and situation-dependent information and service provision. In practice, SPAs unfold their potential in various forms and contexts (Cowan *et al.*, 2017), such as on smartphones (Venkatesh *et al.*, 2017), in smart home environments (Fernando *et al.*, 2016), in cars (Bengler *et al.*, 2014), in service encounters (Xu *et al.*, 2017), or as support for elderly or impaired people (Fernando *et al.*, 2016).

As diverse as the range of application scenarios is, so are the terms used for SPAs and their interpretations in the existing literature. Terms like conversational agent, cognitive assistant, user assistance system, chatbot, virtual personal assistant, to name just a few, are widely used synonymously, although a closer look often reveals that the same name is used for different systems or that different names are used for the same kind of systems. The interchangeable use of terms has also been noticed in previous work (e.g., Cowan *et al.*, 2017). The prevalent variety of terms and meanings indicates that neither a common knowledge base nor a shared understanding for constituting design characteristics, i.e. patterns, of SPAs exist. However, such common base is important to

Author's address: R. Knote, M. Söllner, J.M. Leimeister, Research Center for Information System Design (ITeG), Pfannkuchstr. 1, 34121 Kassel; Germany; E-Mail: {robin.knote; soellner; leimeister}@uni-kassel.de

prevent divergence of future research and allow scholars to 'stand on the shoulders of giants' within an increasingly opaque and complex research domain (Gregor, 2006).

With this paper, we want to shed some light on patterns of SPAs. Based on an exhaustive literature review, we establish a pattern language of SPAs. We therefore focus on *interaction* and *intelligence* as abstract main characteristics, since the novel nature of these factors is salient for advanced user assistance systems (Maedche *et al.*, 2016). Considering an SPA to be a sociotechnical system, the design of assistive functionality (software) in conjunction with the device (hardware) is not only driven by technical potential but also highly dependent on social aspects, namely the user and the use contexts. Hence, in order to develop successful (i.e. useful and intensively used) SPAs, we must first understand which interaction and intelligence patterns are salient for this new class of systems. However, since our pattern language describes and classifies SPA characteristics, it does not yield for explaining causality or attempting predictive generalizations. Such descriptive contributions are especially needed when nothing or little is known about the subject of interest (Fawcett and Downs, 1986), as it is the case for SPA interaction and intelligence patterns. Future research can leverage our results to establish advanced types of theory that allow for explanation, prediction and design. Furthermore, the pattern language contributes to practitioners concerned with the analysis or design of SPAs. They receive guidance for existing implementations which will aid their own design decisions.

To set the scene for our paper, we first introduce the state-of-the-art of SPA research and practice as this will create a shared understanding of the unit of analysis. We then elaborate on the systematic literature review as the basis for our pattern language. Afterwards, we present the pattern language and containing patterns in detail. The paper closes with a conclusion, limitations and suggestions for future research.

2. BACKGROUND

Although SPAs have just recently gained success on the consumer market, personal assistance provided by information systems is not a novel topic at all. In the past, research around question answering systems like BASEBALL (Green Jr. *et al.*, 1961), ELIZA (Weizenbaum, 1966), and LUNAR (Woods and Kaplan, 1977) was mainly conducted in the field of artificial intelligence and focused on expert systems in relatively limited domains (Kincaid and Pollock, 2017). However, the advent of technical evolutions, such as cloud-based scalable infrastructure, natural language processing, semantic reasoning, voice recognition and voice synthesis paved the way for modern SPAs such as Apple's Siri, Microsoft's Cortana, Samsung's Bixby, Amazon's Alexa, Google's Google Assistant and also chatbots in the service encounter. These systems interact with the user via natural language and offer many opportunities of service and information provision to reduce effort and complexity of users' everyday tasks (Cowan *et al.*, 2017).

However, a unified definition approach for SPAs (or respective synonyms) is still missing. A broad definition approach has already been conducted in the early 1990s by Baber (1993, p. 223) who considered an SPA to be *"an application that uses input such as the user's voice… and contextual information to provide assistance by answering questions in natural language, making recommendations and performing actions"*. More technical definitions draw on the term 'agent' to describe SPAs. For example, Fuckner *et al.* (2014, p. 89, p. 89) describes an SPA as a "*specialized intelligent artificial agent that helps users to do their activities*" as an "*intermediary between humans and other agents in a multiagent environment.*" The term 'agent' aims to point out that the SPA as an autonomous entity is capable of perceiving and taking actions within its environment to achieve a certain goal (Russell and Norvig, 2003), namely to assist the user conducting a specific task. Further, the SPA as an agent (e.g., Alexa) is able to interact with other agents, such as technical agents (e.g., a smart fridge) and human agents (users). The multi-agent concept also encompasses a layer view. Therein, an SPA consists of different layers, each conducting a specific sub-task (e.g., interface agent, interaction agent, transaction agent). For example, the user interacts with the interface agent which delegates more specific tasks to other types of agents (Fuckner *et al.*, 2014).

The main purpose of SPAs is to enhance the user's perception, cognition and/or action abilities (Jalaliniya and Pederson, 2015). From a sociotechnical perspective and compared to other classes of information systems, the novelty of SPAs lies in two major aspects: the way how users *interact* with the device as well as the assistant's knowledgeability and human-like behavior, often summarized as *intelligence* (Maedche *et al.*, 2016; Russell and Norvig, 2003). Due to natural language interfaces and context-awareness, typical utilization barriers are heavily reduced. As opposite to other classes of systems where users must learn commands specified in a given syntax to instruct the system, SPA users may communicate in ways which feel more natural to them, like writing or talking in natural language and pointing at things which appear interesting. Moreover, modern SPAs like Amazon Alexa or the Google Assistant serve as single, ubiquitous and easy-to-access entry point to a user's digital infrastructure, from online profiles to smart home appliances, just like a web browser used to be the 'gate to the world' at the beginning of the internet era. While the interface design leads to increased comfort and ease of interaction, the SPA's behavior is mainly determined by intelligence aspects. SPAs are often based on semantic models (such as ontologies) of some general domain knowledge (Kincaid and Pollock, 2017). However, the intelligence of SPAs does not only refer to the semantic understanding of 'the world'. Rather, SPAs are supposed to be adaptive to users' needs, wishes and prior interaction to improve their assistance behavior accordingly (Maedche *et al.*, 2016). To realize adaptive behavior and foster evolution, some SPAs rely on machine learning algorithms. For example, Alexa's natural language processing abilities are based on the analysis of a massive amount of user-generated utterances (Maas *et al.*, 2017). Consequently, in contrast to basic systems, SPAs are characterized by more sophisticated features of interaction and intelligence. They allow users to decide whether to follow the assistance, provide a high extent of context-aware and proactive assistance, include adaptation capabilities and detect users' needs (Maedche *et al.*, 2016).

While most of the human-computer-interaction (HCI), information systems (IS) and computer science (CS) literature focus on the development and evaluation of SPAs (e.g., Griol *et al.*, 2013; Kanaoka and Mutlu, 2015; Derrick *et al.*, 2011) and SPA components (e.g., Cassell, 2000; Armentano *et al.*, 2006), recent work shifts the focus to the user and the use context. For example, Purington *et al.* (2017) examined how users personify Amazon's Alexa and which social roles the SPA may take. Their results indicate that personification of SPAs, such as 'natural' voice output, predicts user satisfaction. However, since personification makes users assign human-like traits to their SPAs this may lead to unrealistic expectations and, consequently, to dissatisfactions when expectations are not met (Luger and Sellen, 2016; Cowan *et al.*, 2017). Other studies have investigated emotional responses towards SPAs (Sandbank *et al.*, 2017; Yang *et al.*, 2017), the triumvirate of security, privacy and trust (Campagna *et al.*, 2017; Nasirian *et al.*, 2017) and willingness-to-pay for trustful SPAs (Mihale-Wilson *et al.*, 2017).

Since our goal is to develop a pattern language for SPAs, prior structuration and characterization efforts are especially valuable as foundation to build up on. Inspired by driver assistance systems in the automotive context (Bengler *et al.*, 2014), Maedche *et al.* (2016) suggest a classification of user assistance systems based on two dimensions: (1) the degree of intelligence of the system and (2) the degree of interaction implemented by the system. According to these characteristics, they distinguish four types of user assistance systems. Basic user assistance systems are characterized by a low degree of interaction and low degree of intelligence. Interactive user assistance systems offer a higher degree of interaction as they provide assistance in cooperation with the user. Intelligent user assistance systems extend basic systems by features that make them adaptive to their users and given context situations. Last, anticipating user assistance systems combine both intelligence and interaction to anticipate future situations and proactively adapt their assistance. From a technical perspective, the latter class of systems is suggested to be the most sophisticated one.

Purington *et al.* (2017) uses four dimensions to delimit SPAs, including degree of personification, degree of sociability, integration and technical qualities and issues. Degree of personification refers to the extent to which the technology is personified by the user. Degree of sociability means the extent of interaction. Integration is the system's ability to connect with other entities. Finally, technical qualities and issues refer to the performance of the system's tasks, e. g., the extent to which the system gives intelligent responses to humans' voice requests.

Finally, Jalaliniya and Pederson (2015) suggest four types of information exchange between SPAs and users. According to them, explicit input is action intentionally and consciously directed towards the assistant, such as when a user navigates through the menu of an SPA. Implicit input is action performed by the user without the conscious intention of communicating with an SPA, e.g., by acting with the physical world and other humans. Explicit output is an assistant's action that the user cannot avoid consciously perceiving, such as a check for the correct understanding of the latest user request. Implicit output occurs when the SPA creates a change in the user's perception space to address the unconscious part of cognition, such as adapting ambient light to the user's mood.

However, although these classification approaches provide valuable starting points for our endeavor, further steps are needed for specifying recurring SPA design characteristics.

3. METHOD

To identify papers relevant for the development of an SPA pattern language, we conducted a systematic literature review (Vom Brocke *et al.*, 2015; Webster and Watson, 2002). Prior study of literature (Purington *et al.*, 2017; Maedche *et al.*, 2016; Venkatesh *et al.*, 2017; Nunamaker *et al.*, 2011; Wang and Benbasat, 2005) revealed the following keywords for an open database search: *"smart assistant" OR "conversational agent" OR "virtual assistant" OR "assistance system" OR "personal assistant"*. The search phrase was adapted to fit the databases' syntactic requirements. The open database search was constrained to title, abstract, keywords and a publication period from 2000 to date. Databases include *AISeL, EBSCO Business Source Premier, ScienceDirect, IEEE Xplore, ACM DL* and *ProQuest*. The open database search resulted in 2802 hits. Titles, abstracts and keywords were screened to fit the purpose of the study. Selected papers either provide a conceptual view on SPAs in general or describe the design of a specific SPA.

Table 1. Literature Survey Results

	Databases and Amount of Papers						
Process Steps	ACM DL	AISeL	EBSCO	IEEE XPlore	ProQuest	Science Direct	Total
Search	800	26	136	1074	94	672	2802
Screening	123	20	27	110	11	63	354
Relevant	53	14	14	62	1	41	185

Screening resulted in 354 articles that appeared relevant. By thoroughly reading the full texts, the number of contributions has further been reduced to a manageable amount. A backward and forward search as well as an open search on Google Scholar revealed three further papers. Table 1 lists the number of search results after each review phase. We further investigated product websites of commercial SPAs, such as the Amazon Echo or Google Assistant product lines. Among all sources, a total amount of 115 SPAs were examined.

4. RESULTS

4.1 Literature Findings

Results of our literature review reveal that a delimitation of SPAs by intelligence and interaction characteristics seems appropriate. Therefore, we inductively established a categorization system for SPAs from the body of literature. In detail, we started with the high-level concepts interaction and intelligence as defined by Maedche *et al.* (2016) and allocated design attributes of SPAs from the literature according our experience in this domain (e.g., Knote *et al.*, 2016; Knote and Söllner, 2017; Knote *et al.*, 2018). We then grouped attributes (e.g. text communication and voice communication) into more abstract dimensions (e.g. communication mode) so that each dimension comprises at least two characteristics. The final category system comprises 10 dimensions, each with up to six distinct characteristics. For the interaction of SPAs, we found 17 distinctive characteristics in five dimensions which describe the exchange between an SPA and its users. In the following, we explain each dimension and respective characteristics in detail, providing justificatory references for each characteristic and the percentage of SPAs from our review obtaining this feature.

Communication mode refers to the primary way(s) a user communicates with an SPA and vice-versa. Communication is either based on user-entered and/or SPA generated text (15,65%; Sansonnet *et al.*, 2012), user's and/or synthesized voice (20%; Weeratunga *et al.*, 2015), optical sensors, cameras and generated animations (2,61%; Jalaliniya and Pederson, 2015), a combination of voice and vision or text (49,57%; Hauswald *et al.*, 2016), or observational sensing and/or unconscious acting (i.e., assistance is not inevitably augmentable for the user; 6,96%; Chen *et al.*, 2014).

Direction of explicit interaction comprises user-to-system interaction (3,48%; Campagna *et al.*, 2017), system-to-user interaction (15,65%; Sato *et al.*, 2014) and bidirectional interaction (80,87%; Tsujino *et al.*, 2013). User-to-system interaction means that the user provides input which is intentionally and consciously directed towards the SPA whereas the system's response may be unconscious for the user, e.g., by executing a service without responding to the request. System-to-user interaction refers to output which occurs when an SPA addresses the conscious mind to create a change in the environment that the user cannot avoid consciously perceiving (Jalaliniya and Pederson, 2015). In this case, however, the user does not put an explicit request upfront but rather receives the result of the SPA's ability to passively observe and make sense of the user's context. Interaction is bidirectional if the SPA is designed to deliver its service in communicational exchange, as it is the case for most SPAs commonly referred to as chatbots or conversational agents.

Query input describes the way in which the user can direct requests towards the SPA. Requests can either be predefined formal prompts that users must know to trigger a desired action (10,43%; Tsujino *et al.*, 2013), natural language requests (72,17%; Sugawara *et al.*, 2011) or accumulations of context data which, from a user perspective, is often collected unconsciously via sensors (17,39%; Czibula *et al.*, 2009).

Response output means the way that an SPA is technically able to formulate responses to user requests. An SPA provides visual output if it responds via text, images, videos, an avatar or any combinations of the aforementioned (30,43%; Onorati *et al.*, 2012). Voice output refers to responses via synthesized speech as it is common for most commercial SPAs currently available (17,39%; Schmeil and Broll, 2007). SPAs that combine visual and verbal responses, such as smart speakers with an integrated screen, are classified as voice and vision (52,17%; Kincaid and Pollock, 2017).

Action refers to an SPA's capabilities to execute services based on the query input. One can broadly distinguish between the general ability to, for example, play music, set alarms or control smart household objects (56,52%; Hauswald *et al.*, 2016) and 'simple' functionality of question answering and information retrieval (43,48%; Sugawara *et al.*, 2011).

The meta-characteristic intelligence comprises 14 characteristics in five dimensions that specify the knowledgeability and human-like information processing and response behavior of SPAs.

Assistance domain refers to the purpose of assistance. It determines both the functionalities and the knowledge models (i.e., semantic models like ontologies) that must be implemented to provide appropriate assistance for a context. An SPA may either provide general assistance like retrieving information, searching on the web or playing one's favorite music (39,13%; Sansonnet *et al.*, 2012), or specific assistance for certain complex tasks or to a dedicated user group (60,87%; Sugawara *et al.*, 2011; Kincaid and Pollock, 2017).

Accepted commands differ according to the SPA's ability to dissolve and process user requests of different complexity. The simplest form is the manual entry of specific data for the assistant to fulfil its purpose (14,78%; Chen *et al.*, 2014), followed by simple commands such as "send email to Jeff" (31,3%; Weeratunga *et al.*, 2015) and compound commands such as "every day at 6am get the latest weather and send it via email to Jeff" (10,43%; Campagna *et al.*, 2017). However, some SPAs do not offer the user the ability to control system behavior (43,48%; Venkatesh *et al.*, 2017). Often this correlates with a passive or observational communication mode.

Adaptivity represents the system's ability to learn from (usually a rich amount of) data and adapt the assistance based on interpretations of this data (44,35%). Examples are the improvement of speech recognition (Arsikere and Garimella, 2017) or tailored interaction for different users in the same context (Armentano *et al.*, 2006). An SPAs is characterized to have either static behavior (55,65%), if assistance is not reflected and revised against context (Grujic *et al.*, 2009), or adaptive behavior if assistance is a function of context or prior assistance (Campagna *et al.*, 2017).

Collective intelligence is defined as the ability to learn, to understand, and to adapt to an environment by using the knowledge of the user crowd (Leimeister, 2010). SPAs may leverage the potentials of collective intelligence to improve machine learning algorithms and, thus, increase the quality of their assistance. For example, the analysis of many users' natural language utterances may lead to a steeper learning curve for speech recognition algorithms since adaptivity is based on a large and heterogenous data set. Hence, individual SPA users may benefit from the 'wisdom of the crowd' (19,13%; Campagna *et al.*, 2017). However, most SPAs do not leverage the potentials of crowd engagement (80,87%; Schmeil and Broll, 2007).

Embodiment refers to the aspiration to present the user a clearly identifiable counterpart who provides personal assistance. In SPAs, this is mostly accomplished through anthropomorphism, *"a conscious mechanism wherein people infer that a non-human entity has human-like characteristics and warrants human-like treatment"* (Purington *et al.*, 2017, p. 2854). Embodied or anthropomorphic design is usually applied to provide a shared common ground, represent an authentic entity, combine verbal and non-verbal communication and align minds by being interesting, creative and humorous (McKeown, 2015). In practice, embodiment is accomplished by virtual characters, i.e., avatars (12,17%; Ochs *et al.*, 2017), a (often human-like) computer voice (24,35%; Trovato *et al.*, 2015) or a combination of both (37,39%; Zoric *et al.*, 2005). However, some SPAs do not use embodiment at all (26,09%; Venkatesh *et al.*, 2017).

4.2 Pattern Language Map

Based on the findings from the literature, we established a pattern language map in the shape of a tree with two major branches, i.e. intelligence and interaction. Figure 1 shows the tree-shaped pattern language map. In the following sections, we describe the patterns in more detail. It is important to note that the patterns are by far not complete. Future research, however, may find our work a good basis to identify further patterns and investigate relations among them.

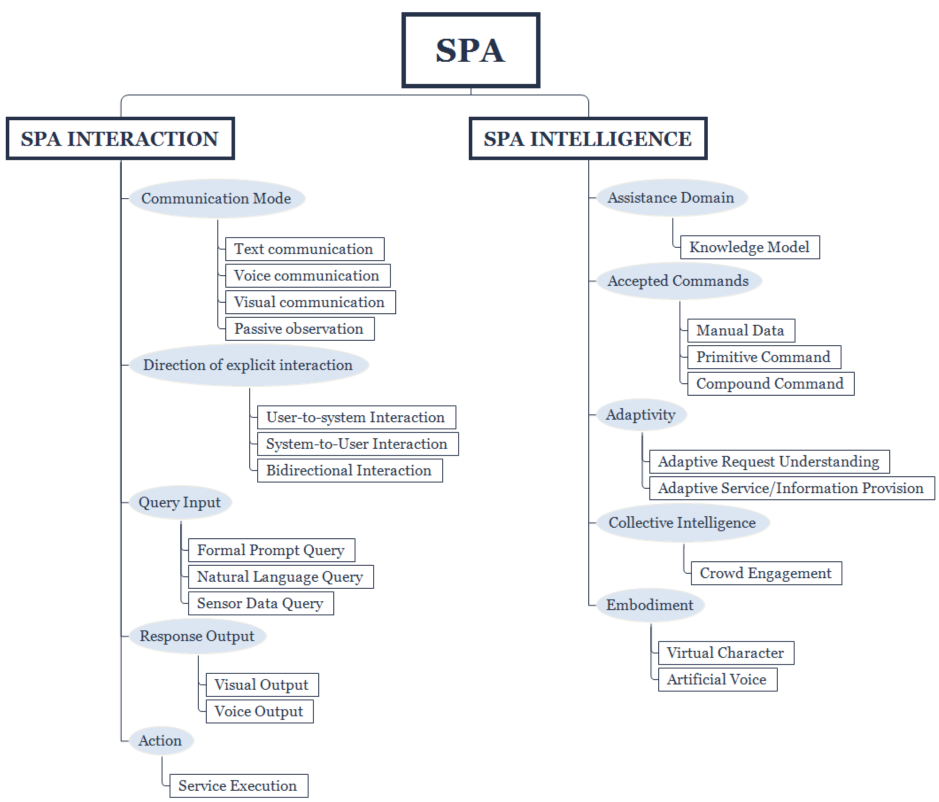

Fig. 1. Pattern Language for SPA interaction and SPA intelligence

4.3 SPA Interaction Patterns

Communication Mode

Context: The way how an SPA delivers information and services varies according to its purpose, sensors, actuators and information processing abilities.

Problem: Users want to easily communicate with their assistants. In turn, the system should provide output which is easy to understand. The interaction should feel natural to the user. Furthermore, not every interaction mode is appropriate for any situation.

Table 2. Communication Mode Patterns

Pattern Name	Forces	Solution	Consequences
Text Communication	*Situation-dependence:* Users often face situations where they cannot or do not want to speak freely (e.g., in public or in noisy work environments). *Datacenter Resources:* The automated analysis of visual or audio data requires a massive amount of hardware and sophisticated algorithms. Some SPA providers cannot raise these resources. *Inclusion and Accessibility:* People with hearing or speech impairments should also be able to communicate with the assistant.	For screen and keyboard-based SPAs, such as Chatbots, text communication is the predominant mode of interaction. Therefore, the SPA reacts to users' text requests, enhances peer-to-peer text communication (e.g., with context-specific information) or leads text conversations towards a direction which is relevant to its assistance domain.	*Benefits* Users have control over how the interaction takes place, for example by correcting the text before submitting it. Chat history serves as protocol for SPA interactions. Sensitive or personal communication can be better controlled by the user compared to speech assistance. *Liabilities* Users must have their hands free, which in certain contexts cannot be ensured (e.g., physical work). Visually impaired or illiterate users need additional systems, such as screen readers and speech-to-text interfaces.
Voice Communication	*Situation-dependence:* Some situations require to use both hands (e.g., physical work) and/or to be visually focused (e.g., driving) *Natural interaction:* Users are more likely to accept and continuously use the SPA when they can interact with it in a way which feels natural. *Inclusion and Accessibility:* Visually impaired or illiterate users should also be able to use the SPA.	For microphone and speaker based systems, such as smart speakers, voice input and output may increase users' acceptance of the system and the information and service it provides. This, however, requires computer-generated voice to be human-like.	*Benefits* Most familiar communication mode Users may personify and even establish social relationships with the SPA which drives continuous use. SPA can be used hands-free. *Liabilities* Voice processing requires many computer resources and sophisticated voice partitioning, natural language understanding and speech generation functionality. People with hearing or speech impairments are mainly excluded if voice is the only communication mode. Always-on microphone raises privacy concerns.
Visual Communication	*Situation-dependence:* Some tasks require to use both hands. *Inclusion and Accessibility:* People with hearing or speech impairments should also be able to use the assistant.	Camera and screen based systems, such as smart glasses, offer the potential to recognize and evaluate users' gestures and mimics. They provide information and services via screen, e.g., by displaying the height of a famous building the user is pointing to.	*Benefits* Assistance becomes increasingly ubiquitous. SPA can be used hands-free. Natural interaction. *Liabilities* Image and video processing requires many computer resources and sophisticated algorithms. Visually impaired people are mainly excluded if vision is the only communication mode. Always-on camera raises privacy concerns.

Passive Observation	*Situation-dependence:* In some situations, such as sports or cognitively demanding tasks, users need assistance which does not distract their attention away from the task.	Passively observing SPAs, such as smart health gadgets, usually collect data via sensors without the user being conscious about it. They further provide information or services only, when the user requests them (e.g., after a workout).	*Benefits* Assistance becomes passive and pervasive. Users have total control over when and where they receive assistance. *Liabilities* Sophisticated algorithms are needed to interpret and make sense of mixed sensor data.

Direction of explicit interaction

Context: Depending on the assistance purpose, the direction of explicit interaction, i.e. whether the system, the user or both are actively involved in the interaction, may differ.

Problem: The user wants the SPA to perform a certain task or to provide context-specific information.

Table 3. Direction of Explicit Interaction Patterns

Pattern Name	Forces	Solution	Consequences
User-to-System Interaction	*Task-dependence:* Some tasks such as turning on the lights or playing music require the user to simply formulate an instruction.	A primarily sensing SPA collects user input which is intentionally and consciously directed towards the SPA. The system, however, responses in a way that is mainly unconscious for the user, e.g., by executing a service without responding to the request.	*Benefits* The user has control over mainly simple tasks in which the SPA may interact on behalf of the user. *Liabilities* System behavior may not always be understandable if it does not explicitly respond to user requests.
System-to-User Interaction	*Task-dependence:* Some tasks such as providing health information during sports require the SPA to act without the user formulating an instruction.	A primarily acting SPA generates output which addresses the conscious mind to create a change in the environment that the user cannot avoid consciously perceiving. Explicit user requests are not needed for the SPA to act. The system usually leverages passive observation abilities to makes sense of the user's context and anticipate needs.	*Benefits* The SPA may perform appropriate actions without explicit user instructions. *Liabilities* System behavior may not always be understandable if it does not explain which input leads to SPA actions.
Bidirectional Interaction	*Task-dependence:* Some tasks such as such as using the SPA to send an e-mail follow a multi-step process and thus require conversational exchange.	Sensing and acting SPAs deliver services in communicational exchange (e.g., by laddering through question-answering dialogues).	*Benefits* The user has total control over SPA behavior. Complex requests can be managed in user-SPA cooperation. The user may correct unwanted requests. *Liabilities* Sophisticated dialogue management functionalities must be implemented. SPAs should be controlled via natural language to simplify bidirectional interaction.

Query Input

Context: SPAs make sense of some kind of input to deliver appropriate information and services. User input may be of different varying forms and complexity.

Problem: Users must be able to communicate in a way an SPA understands.

Table 4. Query Input Patterns

Pattern Name	Forces	Solution	Consequences
Formal Prompt Query	*Implementation complexity:* Natural language processing functionality is complex and requires a massive amount of datacenter resources.	The SPA provides control via predefined formal prompts.	*Benefits* As they are close to imperative programming, formal prompt control is the least complex to implement input option. *Liabilities* The user must learn the syntax and semantics of prompts to receive appropriate assistance. An exhaustive documentation should be provided.
Natural Language Query	*Natural interaction:* Users want to formulate instructions in ways that feel natural to them.	An SPA which is able to process natural language can dissemble requests into executable directives. It also can handle uncertainties and inaccuracies of the human language.	*Benefits* The SPA allows for more natural user experiences which may increase acceptance. *Liabilities* Compared to formal prompt queries, natural language requests are exponentially more complex.
Sensor Data Query	*Unobtrusiveness:* In some situations, information required for assistance can be inferred from sensor data without disturbing the user.	An SPA is usually equipped with various sensors. The combination of sensor data, also called sensor fusion, may form a precise picture of the user's context and assistance needs.	*Benefits* The user receives assistance without explicitly formulating instructions. *Liabilities* A concept should be developed to determine which data is collected and combined for which purpose and what can be inferred to provide appropriate assistance in a given situation.

Response Output

Context: Depending on the assistance purpose and the context, in which users usually receive an SPA's assistance, information and/or services must be delivered in a suitable form.

Problem: SPAs must be able to communicate in a way the user understands.

Table 5. Response Output Patterns

Pattern Name	Forces	Solution	Consequences
Visual Output	see visual communication	Screen-based SPAs deliver assistance via text, pictures or (animated) video sequences. They thus influence the user's visual perception space.	see visual communication
Voice Output	see voice communication	Speaker-based SPAs deliver assistance via speech or acoustical signals. They thus influence the user's acoustical perception space.	see voice communication

Action

Context: SPAs are usually integrated in user's virtual life (e.g., social media, e-commerce, online banking) or life-enhancing hardware (e.g., smart home, e-health gadgets, driver assistance).

Problem: Users want the SPA to be the single point of control over both their digital and their physical environment.

Table 6. Action Pattern

Pattern Name	Forces	Solution	Consequences
Service Execution	*Digital Outreach:* The SPA must be able to control the user's existing online profiles. *Physical Outreach:* The SPA must be able to control smart gadgets.	The SPA offers control over connected (third-party) services and gadgets via unified, easy to use interface (e.g., speech). To implement service execution abilities, some SPAs use closed, 'hardwired' approaches while others rely on an open, modular strategy. While the first requires the developer to anticipate and implement control opportunities, the latter allows for continuous enhancement of an SPA's service execution abilities (e.g., Alexa Skills).	*Benefits* The user has a single point of control over his digital and smart physical environment. *Liabilities* Possible compatibility issues. Third-party modules should be quality-checked by the SPA provider. Network structure requires extended security and privacy considerations.

4.4 SPA Intelligence Patterns

Knowledge Model

Context: Appropriateness of assistance is dependent on the task and user context. Assistance context determines which parts of the world an SPA must be aware of and how these parts are represented.

Problem: Depending on the assistance purpose and context, SPAs need different understandings (or representations) of 'the world'.

Table 7. Assistance Domain Pattern

Pattern Name	Forces	Solution	Consequences
Knowledge Model	*Purpose*: While specific purpose assistants require specific domain knowledge, general SPAs usually need a wide-spread understanding of different fields. *User Group*: SPAs may either provide information and services for a broad user group or for a limited or specialized user group (e.g., surgeons) which determines the knowledge the system must encompass.	The knowledge of an SPA is usually based on semantic knowledge models, such as ontologies or neural networks. Most specific SPAs that do not need in-time adaptivity rely on exhaustive (but rather static) ontologies that represent the relevant parts of the world.	*Benefits* Only knowledge relevant for the assistance domain is implemented. Decisions are basically traceable to a certain degree. *Liabilities* Knowledge must be represented in a comprehensive manner. Knowledge should be regularly updated to ensure consistency. Traceability of decisions can be limited when probabilistic models and/or machine learning algorithms are used.

Accepted Commands

Context: Besides question-answering, SPAs usually offer various opportunities for users to control system behavior.

Problem: Complexity of user requests vary. The more complex a user request is to understand for the system, the more sophisticated technology is needed.

Table 8. Accepted Commands Patterns

Pattern Name	Forces	Solution	Consequences
Manual Data Entry	*Implementation effort*: Resolving primitive or compound user requests requires sophisticated algorithms and thus higher implementation effort. *Task adequacy*: Simple assistance tasks which do not need extensive user control may be conducted with fewer or less complex commands.	Manual data entry includes entering specific values when asked and control via common elements, such as checkboxes or bars. These simple command structures are often found in text-based SPAs, such as Chatbots or virtual tutors.	*Benefits* Manual data entry is the least complex form of control over SPA behavior. *Liabilities* Using manual data entry requires the user to leanr how the system works. Control may be non-intuitive. Accepted commands should at least be documented. Users may not have full control over system behavior.
Primitive Command	*Implementation effort*: Resolving compound user requests requires sophisticated algorithms and thus higher implementation effort. *Learning effort*: Users want the SPA to understand their language and act accordingly. They do not want to first learn how the system works until they can learn how to conduct tasks with it.	Primitive commands can either be stated via text or voice directly addressed to the SPA. They are usually followed by information output or service execution (e.g., sending a mail).	*Benefits* Users have a more natural experience because they can instruct SPAs more easily. They do not need to learn how the system works (e.g., input format, reaction to certain input). *Liabilities* Implementation effort is higher because input queries must be decomposed and knowledge models must represent appropriate actions. This requires developers to anticipate user behavior and establish fallbacks. If only primitive commands are accepted, users must formulate each part of a compound command in a separate query.
Compound Command	*Learning effort*: Users want the SPA to understand their language and act accordingly. They do not want to first learn how the system works until they can learn how to conduct tasks with it. *Convenient input*: Users want to put multiple requests in a single input query.	When accepting compound commands, the SPA is able to decompose queries, such as "ring the alarm bell, turn on the lights and start the coffee machine every day at 7 a.m.".	*Benefits* Users can instruct SPAs in a way that feels most natural for them. *Liabilities* Decomposing compound commands is complex and requires more datacenter resources and more sophisticated algorithms. It must be ensured, that, if one part of the query cannot be interpreted, the rest of the query remains valid or the invalid parts can be corrected/repeated.

Adaptivity

Context: Like human personal assistants, some SPAs are able to learn from users' requests over a period of time. This may increase assistance quality since the SPA is able to improve request handling and information/service provision based on prior interactions.

Problem: SPA should improve assistance quality by autonomously adapting to user peculiarities and learn from prior interactions.

Table 9. Adaptivity Patterns

Pattern Name	Forces	Solution	Consequences
Adaptive Request Understanding	*Individualism*: Natural language processing is a complex task. The way how users articulate in written or spoken natural language heavily varies. An SPA should correctly understand all user peculiarities, such as dialects or speech impediments.	Some SPAs rely on flexible speech models that allow for improvements over time. SPAs thereby learn from correctly and incorrectly performed actions or responses to user requests.	*Benefits* User requests can be better understood after some interactions. *Liabilities* Adaptive request understanding requires the (semi-)automated analysis of spoken language and, consequently, hardware resources and sophisticated machine learning algorithms. Utterances may contain personal data that are subject to privacy and legal regulations.
Adaptive Information/ Service Provision	*Behavioral Patterns*: Since humans often follow behavioral patterns, SPAs should (proactively) infer appropriate actions from data gathered in previous interactions.	Some SPA's knowledge relies on models and algorithms that can be trained by repetitively performing an action in similar contexts or situations.	*Benefits* Users receive better assistance after some interactions. *Liabilities* Adaptive information and service provision requires sophisticated machine learning algorithms. Behavioral patterns may allow for inference to personal data and is subject to privacy and legal regulations.

Collective Intelligence

Context: Especially commercial SPAs collect and store data from user interaction in the cloud. Hence, there is a potential to make sense of other users with similar behavioral patterns and, thus, increase the quality of assistance in a given situation.

Problem: Assistance quality should be independent from individual user's utilization frequency.

Table 10. Collective Intelligence Pattern

Pattern Name	Forces	Solution	Consequences
Crowd Engagement	*First-time or irregular use:* First-time and irregular users should receive a similar high-quality assistance as regular users. *User Crowd*: Especially commercial and/or general-purpose SPAs often have a larger user crowd in which behavioral patterns and/or natural language characteristics of individual users are alike.	(Anonymously) stored usage data from the user crowd is aggregated to increase the assistance quality for individual users with similar behavioral patterns.	*Benefits* All users, be it regular, irregular or first-time, profit from each other's' interactions. Algorithms can be trained faster. *Liabilities* Usage data is subject to privacy and legal regulations. Anonymity must be ensured, so that no user is identifiable for another user. Selecting, merging and analyzing requires great computational and, if conducted semi-autonomously, human power.

Embodiment

Context: Research has found that SPAs are more likely accepted when they are ascribed human-like traits by their users (also called anthropomorphism; see e.g., Purington *et al.*, 2017).

Problem: Users want to personify their SPA just like they would do with a human personal assistant.

Table 11. Embodiment Patterns

Pattern Name	Forces	Solution	Consequences
Virtual Character	*Visual Appearance*: Users feel more comfortable when the assistant is identifiable as a virtual character.	Screen-based SPAs can offer virtual characters. Assistance is usually assigned to these characters, e.g., by speech boxes or animated mimics and gestures. Avatars are either human-like or abstract.	*Benefits* The user can communicate with an embodied virtual assistant which is suggested to increase system acceptance. *Liabilities* Visually impaired people are mainly excluded and cannot enjoy the full user experience if a virtual character is the only form of embodiment.
Artificial Voice	*Human-like voice*: Users feel more comfortable when the assistant communicates with a human-like voice.	Speaker-based SPAs often provide virtual voices. This is either accomplished by native voice functionality or text-to-speech. Artificial voices are usually human-like.	*Benefits* The user can communicate with a human-like artificial voice in natural language which is suggested to increase system acceptance. *Liabilities* People with hearing or speech impairments are mainly excluded and cannot enjoy the full user experience if voice is the only form of embodiment.

5. CONCLUSION AND FUTURE WORK

This paper presents a pattern language for smart personal assistants, especially regarding the two major characteristics interaction and intelligence. Due to the overall amount of 115 SPAs from which we derived the patterns, we believe we have built a suitable foundation. Future work, however, should overcome the following limitations when it builds up on our work. First, as pattern languages are supposed to be extendible, next steps should include finding more patterns that extend or specify existing ones. Second, relations, dependencies and conditions should be introduced in order to move from the language's hierarchical nature to a full-fledged grammar that can be used in systems development. Third, the patterns and the pattern language should be reshaped with regard to related interaction, privacy, security and other relevant patterns. All of these steps hopefully contribute to the emergence of a useful pattern language for SPAs which guides system development.

6. ACKNOWLEDGEMENTS

This paper presents research that was conducted in the context of project "AnEkA" (project number: 348084924), funded by the German Research Foundation (DFG). The responsibility for the content of this publication remains with the authors.

We want to thank our shepherd Y C Cheng for the valuable suggestions which helped to improve our paper in the PLoP shepherding process. We would also like to thank all participants of our PLoP writers' group for their very helpful comments.

7. REFERENCES

Armentano, M., Godoy, D. and Amandi, A. (2006), “Personal assistants. Direct manipulation vs. mixed initiative interfaces”, *International Journal of Human-Computer Studies*, Vol. 64 No. 1, pp. 27–35.

Arsikere, H. and Garimella, S. (2017), “Robust Online i-Vectors for Unsupervised Adaptation of DNN Acoustic Models. A Study in the Context of Digital Voice Assistants”, in *Interspeech 2017, 20-24 August 2017*, ISCA, ISCA, pp. 2401–2405.

Baber, C. (1993), *Developing interactive speech technology*, Taylor & Francis, Inc.

Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C. and Winner, H. (2014), “Three Decades of Driver Assistance Systems. Review and Future Perspectives”, *IEEE Intelligent Transportation Systems Magazine*, Vol. 6 No. 4, pp. 6–22.

Campagna, G., Ramesh, R., Xu, S., Fischer, M. and Lam, M.S. (2017), *Almond: The Architecture of an Open, Crowdsourced, Privacy-Preserving, Programmable Virtual Assistant*, International World Wide Web Conferences Steering Committee.

Cassell, J. (2000), “Embodied conversational interface agents”, *Communications of the ACM*, Vol. 43 No. 4, pp. 70–78.

Chen, C.-C., Huang, T.-C., Park, J.J., Tseng, H.-H. and Yen, N.Y. (2014), “A smart assistant toward product-awareness shopping”, *Personal and Ubiquitous Computing*, Vol. 18 No. 2, pp. 339–349.

Cowan, B.R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., Earley, D. and Bandeira, N. (2017), “"What can i help you with?"”, in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '17, Vienna, Austria, 04.09.2017 - 07.09.2017*, ACM Press, New York, New York, USA, pp. 1–12.

Czibula, G., Guran, A.-M., Czibula, I.G. and Cojocar, G.S. (2009), “IPA - An Intelligent Personal Assistant Agent for Task Performance Support”.

Derrick, D.C., Jenkins, J.L., Nunamaker, J.F. and Jr. (2011), “Design Principles for Special Purpose, Embodied, Conversational Intelligence with Environmental Sensors (SPECIES) Agents”, *Transactions on Human-Computer Interaction*, Vol. 3 No. 2, pp. 62–81.

Fawcett, J. and Downs, F.S. (1986), *The relationship of theory and research*, ACC, Norwalk.

Fernando, N., Tan, F.T.C., Vasa, R., Mouzaki, K. and Aitken, I. (2016), “Examining Digital Assisted Living: Towards a Case Study of Smart Homes for the Elderly”, in *Proceedings of the 24th European Conference on Information Systems (ECIS)*, Istanbul, Turkey.

Fuckner, M., Barthes, J.-P. and Scalabrin, E.E. (2014), “Using a personal assistant for exploiting service interfaces”, in Hou, J.-L. (Ed.), *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD): May 21-23, 2014, National Tsing Hua University, Hsinchu, Taiwan, Hsinchu, Taiwan*, IEEE, [Piscataway, N.J.], pp. 89–94.

Green Jr., B.F., Wolf, A.K., Chomsky, C. and Laughery, K. (1961), “Baseball. An automatic question-answerer”, *Proceedings of the Western Joint Computer Conference*.

Gregor, S. (2006), “The Nature of Theory in Information Systems”, *MIS Quarterly*, Vol. 30 No. 3, pp. 611–642.

Griol, D., Carbo, J. and Molina, J.M. (2013), “A statistical simulation technique to develop and evaluate conversational agents”, *AI Communications*, Vol. 26 No. 4, pp. 355–371.

Grujic, Z., Kovaeic, B. and Pandzic, I.S. (2009), “Building Victor-A virtual affective tutor”.

Hauswald, J., Mudge, T., Petrucci, V., Tang, L., Mars, J., Laurenzano, M.A., Zhang, Y., Yang, H., Kang, Y., Li, C., Rovinski, A., Khurana, A. and Dreslinski, R.G. (2016), “Designing Future Warehouse-Scale Computers for Sirius, an End-to-End Voice and Vision Personal Assistant”, *ACM Transactions on Computer Systems*, Vol. 34 No. 1, pp. 1–32.

Jalaliniya, S. and Pederson, T. (2015), “Designing Wearable Personal Assistants for Surgeons. An Egocentric Approach”, *IEEE Pervasive Computing*, Vol. 14 No. 3, pp. 22–31.

Kanaoka, T. and Mutlu, B. (2015), “Designing a Motivational Agent for Behavior Change in Physical Activity”, in Begole, B., Kim, J., Inkpen, K. and Woo, W. (Eds.), *CHI 2015: Extended abstracts publication of the 33rd Annual CHI Conference on Human Factors in Computing Systems April 18-23, 2015, Seoul, Republic of Korea, Seoul, Republic of Korea*, Association for Computing Machinery, New York, New York, pp. 1445–1450.

Kincaid, R. and Pollock, G. (2017), “Nicky. Toward a Virtual Assistant for Test and Measurement Instrument Recommendations”.

Knote, R., Baraki, H., Söllner, M., Geihs, K. and Leimeister, J.M. (2016), "From Requirement to Design Patterns for Ubiquitous Computing Applications".

Knote, R., Janson, A., Eigenbrod, L. and Söllner, M. (2018), *The What and How of Smart Personal Assistants: Principles and Application Domains for IS Research*, Multikonferenz Wirtschaftsinformatik (MKWI), Lüneburg, Germany.

Knote, R. and Söllner, M. (2017), "Towards Design Excellence for Context-Aware Services - The Case of Mobile Navigation Apps".

Leimeister, J.M. (2010), "Collective Intelligence", *Business & Information Systems Engineering*, Vol. 2 No. 4, pp. 245–248.

Luger, E. and Sellen, A. (2016), ""Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents", in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16, Santa Clara, California, USA, 07.05.2016 - 12.05.2016*, ACM Press, New York, New York, USA, pp. 5286–5297.

Maas, R., Rastrow, A., Goehner, K., Tiwari, G., Joseph, S. and Hoffmeister, B. (2017), "Domain-Specific Utterance End-Point Detection for Speech Recognition", in *Interspeech 2017, 20-24 August 2017*, ISCA, ISCA, pp. 1943–1947.

Maedche, A., Morana, S., Schacht, S., Werth, D. and Krumeich, J. (2016), "Advanced User Assistance Systems", *Business & Information Systems Engineering*, Vol. 58 No. 5, pp. 367–370.

McKeown, G. (2015), "Turing's menagerie. Talking lions, virtual bats, electric sheep and analogical peacocks: Common ground and common interest are necessary components of engagement", in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII): 21 - 24 Sept. 2015, Xi'an, Xi'an, China*, IEEE, Piscataway, NJ, Piscataway, NJ, pp. 950–955.

Mihale-Wilson, C., Zibuschka, J. and Hinz, O. (2017), "About User Preferences and Willingness to Pay for a Secure and Privacy Protective Ubiquitous Personal Assistant", in *Proceedings of the 25th European Conference on Information Systems (ECIS)*, Guimarães, Portugal.

Nasirian, F., Ahmadian, M. and Lee, O.-K. (2017), "AI-Based Voice Assistant Systems: Evaluating from the Interaction and Trust Perspectives", *AMCIS 2017 Proceedings*.

Nunamaker, J.F., Derrick, D.C., Elkins, A.C., Burgoon, J.K. and Patton, M.W. (2011), "Embodied Conversational Agent-Based Kiosk for Automated Interviewing", *Journal of Management Information Systems*, Vol. 28 No. 1, pp. 17–48.

Ochs, M., Pelachaud, C. and Mckeown, G. (2017), "A User Perception--Based Approach to Create Smiling Embodied Conversational Agents", *ACM Transactions on Interactive Intelligent Systems*, Vol. 7 No. 1, pp. 1–33.

Onorati, T., Malizia, A., Olsen, K.A., Diaz, P. and Aedo, I. (2012), "I feel lucky. An automated personal assistant for smartphones", *AVI '12 Proceedings of the International Working Conference on Advanced Visual Interfaces, Capri Island, Italy*, pp. 328–331.

Purington, A., Taft, J.G., Sannon, S., Bazarova, N.N. and Taylor, S.H. (2017), ""Alexa is my new BFF"", in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17, Denver, Colorado, USA, 06.05.2017 - 11.05.2017*, ACM Press, New York, New York, USA, pp. 2853–2859.

Russell, S.J. and Norvig, P. (2003), *Artificial intelligence: A modern approach,* 2nd ed., Prentice Hall, New Jersey.

Sandbank, T., Shmueli-Scheuer, M., Herzig, J., Konopnicki, D. and Shaul, R. (2017), "EHCTool", in Papadopoulos, G.A., Kuflik, T., Chen, F., Duarte, C. and Fu, W.-T. (Eds.), *IUI'17: Companion of the 22nd International Conference on Intelligent User Interfaces March 13-16, 2017, Limassol, Cyprus, Limassol, Cyprus*, The Association for Computing Machinery, New York, New York, pp. 125–128.

Sansonnet, J.P., Correa, D.W., Jaques, P., Braffort, A. and Verrecchia, C. (2012), "Developing web fully-integrated conversational assistant agents", *Proceedings of the 2012 ACM Research in Applied Computation Symposium*, pp. 14–19.

Sato, A., Watanabe, K. and Rekimoto, J. (2014), "MimiCook. A cooking assistant system with situated guidance", *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, pp. 121–124.

Schmeil, A. and Broll, W. (2007), "MARA - A Mobile Augmented Reality-Based Virtual Assistant", in Sherman, W. (Ed.), *IEEE Virtual Reality Conference, 2007: VR '07 ; Charlotte, North Carolina, USA, 10 - 14 March 2007 ; proceedings, Charlotte, NC, USA*, IEEE Service Center, Piscataway, NJ, Piscataway, NJ, pp. 267–270.

Sugawara, K., Manabe, Y., Shiratori, N., Yaala, S.B., Moulin, C. and Barthes, J.-P.A. (2011), "Conversation-based support for requirement definition by a Personal Design Assistant", in Wang, Y. (Ed.), *Proceedings of the 10th IEEE International Conference on Cognitive Informatics & Cognitive Computing: ICCI*CC 2011 August 18 -20, 2011, Banff, Canada, Banff, AB, Canada*, IEEE Computer Society, Los Alamitos, Calif., pp. 262–267.

Tracitca (2016), "The Virtual Digital Assistant Market Will Reach $15.8 Billion Worldwide by 2021", available at: https://www.tractica.com/newsroom/press-releases/the-virtual-digital-assistant-market-will-reach-15-8-billion-worldwide-by-2021/ (accessed 20 August 2017).

Trovato, G., Ramos, J.G., Azevedo, H., Moroni, A., Magossi, S., Ishii, H., Simmons, R. and Takanishi, A. (2015), "Designing a receptionist robot. Effect of voice and appearance on anthropomorphism", in *24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 235–240.

Tsujino, K., Iizuka, S., Nakashima, Y. and Isoda, Y. (2013), "Speech Recognition and Spoken Language Understanding for Mobile Personal Assistants. A Case Study of "Shabette Concier"", in *IEEE 14th International Conference on Mobile Data Management (MDM), 2013: 3 - 6 June 2013, Milan, Italy ; proceedings ; [including workshops], Milan, Italy, 3/6/2013 - 6/6/2013*, IEEE, Piscataway, NJ, Piscataway, NJ, pp. 225–228.

Venkatesh, V., Aloysius, J.A., Hoehle, H. and Burton, S. (2017), "Design and Evaluation of Auto-ID Enabled Shopping Assistance Artifacts in Customer's Mobile Phones: Two Retail Store Laboratory Experiments", *MIS Quarterly*, Vol. 41 No. 1, pp. 83–113.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R. and Cleven, A. (2015), "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research", *Communications of the Association for Information Systems*, Vol. 37 No. 1.

Wang, W. and Benbasat, I. (2005), "Trust In and Adoption of Online Recommendation Agents", *Journal of the Association for Information Systems*, Vol. 6 No. 3, pp. 72–101.

Webster, J. and Watson, R.T. (2002), "Analyzing the Past to Prepare for the Future: Writing a Literature Review", *MIS Quarterly*, Vol. 26 No. 2, pp. xiii–xxiii.

Weeratunga, A.M., Jayawardana, S.A.U., Hasindu, P.M.A.K., Prashan, W.P.M. and Thelijjagoda, S. (2015), "Project Nethra - an intelligent assistant for the visually disabled to interact with internet services", in *2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS): Conference proceedings 17th-20th December, 2015, Peradeniya, Sri Lanka*, IEEE, [Piscataway, New Jersey], pp. 55–59.

Weizenbaum, J. (1966), "ELIZA—a computer program for the study of natural language communication between man and machine", *Communications of the ACM*, Vol. 9 No. 1, pp. 36–45.

Woods, W.A. and Kaplan, R. (1977), "Lunar rocks in natural English. Explorations in natural language question answering", *Linguistic structures processing*, Vol. 5, pp. 521–569.

Xu, A., Liu, Z., Guo, Y., Sinha, V. and Akkiraju, R. (2017), "A New Chatbot for Customer Service on Social Media", in *CHI'17 Proceedings, Denver, Colorado, USA, 6/5/2017 - 11/5/2017*, ACM, New York, NY, New York, NY, pp. 3506–3510.

Yang, Y., Ma, X. and Fung, P. (2017), "Perceived Emotional Intelligence in Virtual Agents", in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17, Denver, Colorado, USA, 06.05.2017 - 11.05.2017*, ACM Press, New York, New York, USA, pp. 2255–2262.

Zoric, G., Smid, K. and Pandzic, I.S. (2005), "Automated gesturing for virtual characters. Speech-driven and text-driven approaches".