PUBLIZISTIK

Neither friend, nor device

The role of personal epistemologies in communication with smart speakers

Marlene Wassmer · Christian Schwarzenegger

Abstract With smart speakers diffusing into society, artificial intelligence is moving from the imaginative reservoirs of dystopian storytelling into vernacular living. How do users perceive communication with it? Are Alexa and Siri considered simple devices, sentient assistants, or even artificial friends? Based on nine qualitative interviews with former smart speaker users in Germany, this study analyzes smart speaker use and related personal epistemologies within a media repertoire perspective. By presenting six interrelated action-guiding principles explaining smart speaker use and people's ambivalent sensemaking, we argue that smart speakers appear neither as friends nor as mere neutral devices to their users. The identified principles explain the peripheral role of smart speakers within media repertoires as handy but suspicious gimmicks. For future smart speaker adoption, whether smart assistants are interpreted as simple-minded, exploitative gimmicks or relevant, reliable, and trustworthy companions will be crucial.

Keywords Media beliefs · Media ideologies · Smart assistants · Media repertoire · AI

1 Introduction

"Alexa, tell me a joke!" Without having to lift a finger, we now ask an artificial intelligence (AI)—or at least devices marketed as such—to do the shopping, read out

✉ Marlene Wassmer, M. A. · Dr. habil. Christian Schwarzenegger
Institut für Medien, Wissen und Kommunikation, Universität Augsburg,
Universitätsstraße 10, 86159 Augsburg, Germany
E-Mail: marlene@wassmer-bw.de

Dr. habil. Christian Schwarzenegger
E-Mail: christian.schwarzenegger@phil.uni-augsburg.de

the weather forecast, adjust the lighting, or play ambient music. The manufacturers of smart speakers advertise their speakers with integrated artificially intelligent voice-assistants as charming everyday companions that can reliably assist their users and even put smiles on their faces from time to time. Smart speakers appear to be manifestations of a social development in which AI is increasingly becoming a part of people's everyday lives. Thus, AI is moving from the imaginative reservoirs of dystopian storytelling into the center of society and vernacular living. The most popular smart speakers on the global market are currently Amazon Echo, with the virtual assistant Alexa, and Google Home, with Google Assistant (Tenzer 2022a), with Amazon taking the largest market share, at 26.4% (Tenzer 2022b). Amazon introduced its devices in September 2016 in Germany (Trenholm 2016), and Google Home smart speakers were available from August 2017 (Bager 2017). In 2021, 20% of the German population owned at least one smart speaker. 76% of them owned an Amazon Echo, 13% owned Google Home, and 13% owned an Apple HomePod[1] (Tenzer 2022b). How do users perceive communication with an AI? What kind of conversation partners are Alexa and Siri considered to be? Are they regarded as just handy devices, simple-minded aids, or household members—or do they even qualify as artificial friends (Brandtzaeg et al. 2022)? With smart assistants as an AI-driven innovation diffusing into society, it is essential to research how people use these new devices and how they make sense of them in their daily lives. Research has shown that people do not have a complete understanding of how algorithms work (Alizadeh et al. 2020) and that they find it difficult to trust algorithms, AI in general, and smart speakers in particular (Bucher 2017; Eslami et al. 2016; Ferrario et al. 2020; Glikson and Woolley 2020; Lomborg and Kapsch 2020; Pridmore and Mols 2020; Rader and Gray 2015; Siles et al. 2020; Toff and Nielsen 2018; Ytre-Arne and Moe 2021). In this paper, we thus investigate the epistemological reasoning related to smart agents and how these sense-making processes affect use and play out in patterns of interaction. Individual sense-making does not occur in a void but is, rather, embedded in social contexts and informed by discourses in society. Within this trajectory, we use the combined framework of folk theories (Ytre-Arne and Moe 2021), media ideologies (Gershon 2010), and personal epistemologies of the media (Schwarzenegger 2020) to investigate how they shape thought and action toward smart speaker use. In particular, we ask whether users approach their devices merely with regard to functionality or as if they are engaging with an actual personality. We aim to identify key principles explaining users' individual sensemaking regarding smart speakers and, relatedly, their communication with these devices. According to the concept of folk theories and personal epistemologies, the attitudes and expectations of would-be users regarding particular media technologies are closely interwoven with patterns and ways of usage. This means that the everyday use of smart speakers unveils the role users assign to the devices. Thus, we must analyze actual smart speaker use in conjunction with users' sensemaking processes and folk theories regarding media technologies, beyond smart speakers alone. In contemporary high-choice media environments and media-saturated societies, no one device is used in isolation; rather, each device is always part of a larger media

---

[1] Multiple answers were possible.

ecology. In line with established media repertoire research (e.g., Hasebrink and Hepp 2017), the concept of personal epistemologies (Schwarzenegger 2020) thus indicates that particular media, as part of a personalized repertoire, are made sense of and shaped by the totality of available technologies and devices and the particular blend of media used by an individual. Accordingly, use and related interpretations become particularly clear when smart speaker use is considered within a media repertoire perspective. This is how we approach the use of and sensemaking for smart assistants in this paper.

2 Literature review: smart speakers in media repertoires and personal epistemologies of their use

2.1 Media repertoires

Today, media have become an integral part of the human lifeworld, and the sheer number of possibilities for media selection and use makes it more difficult and, at the same time, even more necessary to grasp media repertoires (Hasebrink and Hepp 2017, p. 364). A person's media repertoire is the "entirety of media he or she regularly uses" (Hasebrink and Domeyer 2012, p. 758). The media repertoire approach examines the use of media as a social, meaningful practice and specific behavioral patterns of users (Hasebrink and Domeyer 2012, p. 759). Users attribute a specific meaning to each medium within their individual repertoires. Thus, the use of a medium is related to the use of other media, is influenced by attributions of meaning and perceptions, and appears to be independent of the actual uses that a medium offers (Boczkowski et al. 2018). Therefore, media technologies can be understood as an "environment of practice" (Madianou and Miller 2012, p. 173), and these practices include normative, emotional attributions of meaning and conceptions of users.

> People's perceptions of a medium influence their use of that medium, whereas the perception of one specific medium is shaped against the totality of media at an individual's disposal at certain times. The individual's beliefs about media will shape how they utilize particular media, while they decide to use other media differently or not at all. (Schwarzenegger 2020, p. 366)

Along these lines, the role of smart speakers for users only becomes clear when the smart agents are elaborated in the context of the entire media repertoire. For instance, when a user is very prone to using the newest technologies and familiar with AI-supported media, the expectations regarding a smart speaker's capabilities will likely differ from those of less experienced users, and folk theories regarding use and impact of smart agents may be more elaborate or accurate. In contrast, users who have less experience with AI may base their personal reasoning on public debate about the perils and potentials of AI and may be more easily disappointed or overwhelmed by the actual performance of the technology. To date, there are no studies that locate smart speakers in the individual media repertoire. Also, folk theories have been studied with an emphasis on particular technologies, without

treating them as an integral part of media repertoires. The concept of personal epistemologies has not previously been used to investigate the sensemaking of AI and interaction with smart agents. The integrated perspective adopted in this paper can thus help us to better understand the relationship between the use of a particular technology and its location in media repertoires, and the epistemological ground shaping the use and being conversely shaped by the use.

In this paper, we explore how people use smart speakers alongside other media (technologies), i.e., how smart speakers find their position in established media repertoires. We further ask how users' personal epistemologies shape this smart speaker use and the emergence, maintenance and (re)configuration (Peters and Schrøder 2018) of media repertoires over time. Therefore, this study addresses the following research question:

RQ1 How do users communicate with their smart speakers, and what role do smart speakers play within their individual media repertoire?

2.2 Research on smart speaker use

Previous studies on the adoption, use, and non-use of smart speakers allow initial conclusions about the smart speaker's potential role in media repertoires. The dominant factor impacting the user acceptance of smart speakers is their great perceived usefulness in daily life (Kowalczuk 2018, p. 425; McCloskey and Bennett 2020, p. 51; Lau et al. 2018, p. 7). Furthermore, Lau et al. (2018, p. 7) show that respondents' perceived identity as an early adopter or desire to have such an image are also driving factors that influence smart speaker adoption. McLean and Osei-Frimpong show that *utilitarian benefit*, i.e., perceived benefits derived from using the devices; *symbolic benefit*, i.e., the perception of the smart speaker as a status symbol; and *social benefit*, i.e., the perception of social presence or even friendship through interaction with the smart agent, significantly influence the intention to use the device (McLean and Osei-Frimpong 2019, p. 28). Perceived risks regarding data security and privacy exerted a negative effect on the use of voice assistants as a moderator (McLean and Osei-Frimpong 2019, p. 28; see also Kowalczuk 2018, p. 426). Lee and Cho, on the other hand, find that escape from everyday life emerged as the strongest factor explaining smart speaker use (Lee and Cho 2020, p. 1150). In a study by Pridmore and Mols regarding smart speaker use and perception, users praised the practicality of the devices due to their hands-free nature, enjoyed using them, were curious about their features, and integrated their smart speakers into everyday routines, which further increased device adoption (Pridmore and Mols 2020, pp. 7–9; see also McCloskey and Bennett 2020, p. 51).

In their study of the domestication of smart speakers, Brause and Blank identify eight patterns of use: in addition to the widespread pattern of *convenience*, in which the smart speaker performs simple tasks such as setting an alarm, and *entertainment*, in which, for example, the voice assistant tells a joke, the smart speaker also offers *companionship, healthcare, a sleep aid*, and *peace of mind*. In addition, it helps users to achieve *self-control and productivity*, and it offers *increased access* to other smart devices (Brause and Blank 2020, p. 7). Of course, none of these functions are

exclusive features of smart speaker devices; rather, they could be obtained through other technologies. Although other media, e.g., TV or radio, can offer companionship, smart speakers allow for a different degree of reciprocity and something like a conversation. Following Natale, the novel interaction between humans and AI challenges the very concept of medium because the machine is the channel and the producer of communication messages (Natale 2020a, p. 1; see also Natale 2020b). This places the kind of companionship that can be found in the form of smart assistants on an entirely new level. Reciprocity and availability were identified as key features of establishing friendship-like bonds with an AI, whereas not being able to trust and find similarity are hindering this bonding (Brandtzaeg et al. 2022).

Among German smart speaker owners in 2021, people between 35 and 54 years of age were most strongly represented (26%), followed by the 18–34 age group, at 22% (Beyto 2021, p. 16). Smart speaker use is lower in older cohorts (Beyto 2021, p. 16). In 2020, the devices were mostly located in the living room (75%), the kitchen (52%), and the bedroom (47%) (Beyto 2020, p. 44). According to Beyto, in 2020, smart speakers in Germany were primarily used to activate streaming services, control smart homes, and ask questions, as well as for everyday organization (Beyto 2020, p. 46).

In summary, existing research suggests that smart assistants are generally used to handle rather simple tasks. More complex tasks and applications are currently beyond the capabilities of the AI that is available for domestic use.

2.3 Folk theories, media ideologies, and personal epistemologies of the media

To understand how users appropriate smart devices into their everyday lives, comprehending how they make sense of media technologies is a requirement. In communication studies, laypersons are considered to hold "specific ideas about how media work" (Naab 2013, p. 48) and their effects. These so-called folk theories are neither objective nor experimentally tested, but "they embody cognitive biases that influence thought and action" (Gelman and Legare 2011, p. 380). Within this trajectory, following Ytre-Arne and Moe (2021), the notion of "folk theories" bundles the understandings that people draw on in everyday life. Accordingly, "a folk theory approach centers on revealing the conceptions people hold of how the media works—that is, their theories" (Ytre-Arne and Moe 2021, p. 810). These conceptions do not necessarily include actual comprehension of how a technology works and do not need to be based on correct assumptions. Even if a folk theory, e.g., on the functioning of AI, is not based on evidence, users will be guided by the assumptions they have about AI's functionality and capabilities. Folk theories are typically seen as rooted in experience (Nielsen 2016; Toff and Nielsen 2018), but because they can generally be regarded as interpretations of what something (e.g., a social phenomenon, a particular technology, or device) is, what it does, and what it ought to do (Nielsen 2016, p. 840) they can also go beyond personal experience. Accordingly, the question of the provenance and constitution and also the transformation of folk theories becomes a germane concern for researchers.

Within a similar trajectory, Gershon's related concept of *media ideologies* can be defined as "beliefs, attitudes, and strategies about a single medium" (Gershon

2010, p. 389). Gershon studied how, over time, different media were deemed suited, acceptable, or out of touch regarding the practice of breaking up (Gershon 2010, p. 390). She shows how interpretations of what particular media are for and how they are meant to be used shapes which media were used in particular social contexts. Compared to the notion of folk theories, Gershon stresses the importance of a media repertoire perspective when analyzing perceptions of a medium. Folk theories are, in short, subjective theories from laypeople based on experience and assumptions about how a medium works, how it should be used, and what it is not good for. Thus, folk theories focus on knowledge and assumptions, whereas the notion of media ideologies adds the individual meaning users attribute to a medium, following a normative approach and analyzing media use as a cultural practice.

Schwarzenegger fundamentally connects with the concept of media ideologies but expands it to adopt a more holistic view of individual "sensemaking of media in the world" (Schwarzenegger 2020, p. 365). Ytre-Arne and Moe (2021) have concluded that many studies examining folk theories or similar concepts—especially in the context of HCI studies—have focused on certain technologies or contexts of use and have not attempted to situate the user experience and sense-making strategies within broader media contexts. However, following Schwarzenegger (2020, p. 365), we argue that, to fully understand how and why users act on media, it is necessary to analyze the underlying beliefs and attitudes related to media—the personal epistemologies of the media (Schwarzenegger 2020, p. 365)—in an integrated fashion, beyond media or technologies alone. Thus, by analyzing personal epistemologies of the media, we can inquire into the folk theories smart speaker users hold about their devices but approach smart speaker use as individually meaningful practice, one influenced by factors beyond technology- and media-related experiences alone. The concept of personal epistemologies allows us to include the entire process of sensemaking about media, i.e., perceptions and beliefs about manufacturers, which goes beyond the technology itself but still affects its use. Personal epistemologies consist of a variety of dimensions (e.g., attitudes, knowledge, assumptions, folk theories, previous experiences, and worldviews) and therefore cannot be grasped entirely. Rather, we aim to identify key principles with which to characterize smart speaker users' individual and complex sensemaking of media in the world, as Schwarzenegger (2020) showed in his analysis of media news repertoires.

2.4 Research on users' perception of and sensemaking regarding smart speakers

It is plausible that average users have only vague ideas about how smart voice assistants work. Indeed, the exact functioning of AI remains barely comprehensible behind the speaking interface (Burrell 2016, p. 1). This is why Alizadeh et al. (2020) investigated folk theories about AI. Their results show that laypeople have varied expectations: some associate AI with machine learning, others only associate it with automatization (Alizadeh et al. 2020, p. 5). These users "do not have the appropriate knowledge when it comes to the implementation, which can lead to unrealistic expectations from the technology and disappointment when these expectations are not met" (Alizadeh et al. 2020, p. 5). A larger body of studies exists on folk theories

of algorithms, showing that users have many assumptions and ambivalent feelings about algorithms (e.g., Bucher 2017; Eslami et al. 2016; Lomborg and Kapsch 2020; Rader and Gray 2015; Siles et al. 2020; Toff and Nielsen 2018; Ytre-Arne and Moe 2021). However, these studies have not contextualized folk theories of AI or algorithms within broader patterns of media use and repertoires.

Allowing a smart assistant in your home requires a certain level of trust that no harm is likely to result from living with it. However, research has found that it can be difficult to trust an AI, because it is constantly changing and evolving through ongoing learning processes (Ferrario et al. 2020, p. 527; see also Glikson and Woolley 2020). Also, smart speakers have repeatedly attracted attention in the german public discourse on AI, frequently in relation to data protection issues (see e.g. Adorján 2021; Fuest 2019; Reichelt and Hegemann 2019). Smart speakers are thus also subject to normative evaluations: in 2020, almost 60% of Germans agreed that smart speakers may pose as a risk to their privacy (Beyto 2020, p. 7), and the devices had a poor image as compared to other technologies (Beyto 2020, p. 29). However, a follow-up study from Beyto, in 2021, implied that people now tend to think in a more positive way about smart agents and their AI. In studies from Malkin et al. (2019, p. 257) and Lau et al. (2018, p. 11), users were not greatly concerned about their own privacy; rather, they expressed concerns about the generation of data from secondary users, such as children (Malkin et al. 2019, p. 264). In an analysis of the end-user agreement for Amazon Echo, Neville has shown that the terms and conditions are not understandable to a layperson (Neville 2020, p. 343). Consequently, users can scarcely know what opportunities and risks they are taking by using a smart speaker and, thus, must act based on what they think to know and believe to be true. In a study by Lau et al., users show a diffuse trust toward companies regarding data generation and "an incomplete understanding of the privacy risks" (Lau et al. 2018, p. 21). Given this uncertainty about privacy, users develop "protective routines" (Pridmore and Mols 2020, p. 9), such as using the mute function. However, some users seem to experience a lack of controllability on the part of smart speakers regarding privacy protection (Lau et al. 2018, p. 17; Pridmore et al. 2019, p. 128). Overall, users are resigned to what they see as the inevitable generation of their data in the digital environment, explain they have nothing to hide, believe their data are uninteresting, or agree to a privacy-convenience trade-off because, for them, the benefits of the smart speaker outweigh privacy concerns (Lau et al. 2018, pp. 13 and 19; Pridmore et al. 2019, p. 129).

In summary, current research shows that users prioritize the practicality of the devices over privacy concerns, while non-users stress infrastructural or privacy issues (Lau et al., 2018, p. 18; Liao et al. 2019, pp. 107–108; Pridmore et al. 2019, p. 130; Pridmore and Mols 2020, pp. 5–7). Also, potential users and non-users interpret "the convenience that household IPAs [Intelligent Personal Assistants] can offer in everyday life as a risk to lose autonomy and to become dependent on technology platforms" (Pridmore and Mols 2020, p. 8). Pridmore and Mols note that European respondents are more aware of the generation of their data by smart speakers and other Internet-of-Things devices than US respondents (Pridmore and Mols 2020, p. 9). This heightened reported awareness can also be indicative of how critical public discourse regarding privacy concerns and data protection issues resonates

in folk theories and personal epistemologies of European (potential) users. The personal epistemology framework allows the incorporating of data protection and data security attitudes and traditions into cultural contexts and asks how general attitudes toward big tech, dataveillance or data capitalism, even though it is not directly about smart speakers, affects the (non-)use of such devices.

The works cited demonstrate how closely intertwined the (non)use of smart speakers and the users' sensemaking can be, especially regarding privacy. To understand the personal sensemaking of (former) smart speaker users and investigate the *connection* between communication with smart speakers, i.e., smart speaker use and users' interpretations and sensemaking processes, this study pursues the following research question:

RQ2 Based on what epistemological grounds do people make sense of their smart speakers, and how do these affect smart speaker use?

By addressing these two guiding research questions, this study illustrates why and, relatedly, how people (do not) use smart speakers and on which epistemological grounds they make sense of that use. By analyzing users' complex sensemaking process, this study can provide insights into what kind of interlocutors smart speakers are for users and what role a new AI-driven media technology assumes in people's media repertoires. Our study integrates previous findings on smart speaker use and perceptions but also extends it by contextualizing smart speaker use within a media repertoire perspective and individual sensemaking processes.

3 Method

In this qualitative study, we analyze the use of and interaction with smart speakers and related personal epistemologies exploratively. In-depth-interviews with nine (former) smart speaker users in Germany were conducted by the first author. The participants all had experience in smart speaker usage because they owned or had owned a smart speaker themselves or had lived in a household with at least one smart speaker. The informants all come from the circle of acquaintances of the first author and were recruited via theoretical sampling (Strauss 1998, p. 70) because this study is based on Grounded Theory (Strauss 1998). Having this in common, the interviewees were recruited via maximum variation sampling and snowball sampling (Draucker et al. 2007, p. 1142). The main characteristics of the sample can be found in Table 1. At the time of the interviews, seven participants used their smart speakers, two had stopped using them, and one of those two had even sold their device. Thus, experience with smart speakers on the part of nine individuals could be analyzed while including the perceptions of two former users. To ensure a high level of variation within the sample, participants additionally varied regarding the following criteria: how they had come to own a smart speaker, the number of smart speakers owned, the smart speaker model and its manufacturer, housing situation, daily routines and everyday life, occupation, educational background, age, and gender. Informants were between 24 and 69 years old, four of them were female,

Table 1 Main characteristics of the sample

Pseudonym, Gender, Age	Occupational Status	Living Situation	Smart Speaker	Special Information on Diffusion
Marie, f, 24	M. A. student communication, works in social media marketing	Family house with father and brother	1 Amazon Echo	Received smart speaker as a present from her partner, Emre
Emre, m, 24	High school diploma, apprenticeship, works in retail	Single household	1 Amazon Echo Show	Wished for a smart speaker after having used the one at Marie's place
Lea, f, 25	M. A. student geoinformation, writing her thesis at a company	2-room-flat with partner	1 Apple HomePod (sold)	Moved in with her partner, who owned a smart speaker; they sold the device after a few months
David, m, 27	M. A. economics, works in digital consulting	3-room-flat with partner	1 Amazon Echo	Received smart speaker as a present in a contest
Max, m, 29	M. Sc. industrial engineer, works in marketing and sales	3-room-flat with partner	3 Amazon Echoes	Bought 3 smart speakers after having used one in a holiday apartment, recommended device to parents (mother: Andrea)
Nadine, f, 44	Secondary school certificate, housewife	Family house with husband and 3 children	5 Amazon Echoes	Her husband bought the devices
Andrea, f, 59	PhD Mathematician, works in IT	Family house with husband	2 Amazon Echoes	Son Max recommended device to them; they asked for it as a present from him
Klaus, m, 63	PhD, retired mechanical engineer	Family house with wife	1 Amazon Echo	Bought the device but is not using it anymore
Jens, m, 69	Secondary school certificate, retired salesperson	Two-party-house with wife	3 Amazon Echoes, one of which is an Amazon Echo Show	Bought the devices

and five were male. Additionally, one interviewee provided insights into the smart speaker usage of her children and elderly father.

The interviews took place between June 15th and July 28th, 2021, most of them in an online setting, and lasted about an hour. Because media use is embedded in people's daily lives, some general information about the individual's living situation was collected at the beginning of the interview. Then, the interview focused on the participant's media use over time, especially their smart speaker use, within a media repertoire perspective and examined related beliefs, perceptions, and expectations. Finally, the participants were asked to share some general thoughts about AI. All

participants received compensation in the form of a small gift worth about five Euros.

With permission, each interview was digitally recorded and then transcribed. To protect the privacy of participants, pseudonyms have been used. In line with the principles of Grounded Theory, we began transcribing and analyzing the interviews while we were still recruiting more informants. Thus, we were able to reflect on the collected data and include new questions throughout an iterative research process. We ultimately had nine informants.

The analysis involved open, axial, and selective coding until theoretical saturation was reached (Glaser and Strauss 1967; Tracy 2010). After reading all transcripts carefully, we coded each transcript line by line in the open coding phase to "crack the data open" and identify concepts. In the axial coding and selective coding phases, we developed the codes into larger concepts and categories, returning to all interviews repeatedly, and finally analyzed the categories' relationships to one another. We arrived at the point of theoretical saturation when no new codes emerged from the material, and the data were condensed into the six theoretical core categories that we present below. These explain the epistemological grounds based on which individuals navigate their media repertoires and make sense of their smart speakers. The core categories can be divided into two groups: first, there are four sensemaking principles showing that smart speakers are not reliable assistants or companions to users, despite the way in which the manufacturers market them. These four categories are (1) *comfortable insignificance*, (2) *forced simplicity,* the interpretation of smart speakers as (3) *controllable infants*, and a *desired limitedness* (4) . On the other hand, we found two further sensemaking principles showing that, at the same time, smart speakers are also not merely any device for the users; rather, they are attributed with interpretations that differ from the sensemaking for other media within the users' repertoires. These two core categories we found are (5) *occasional humanization* and (6) *exploitative presence.* The six core categories describing the personal sensemaking of smart speaker use, along with general contextual findings, are presented below.

4 Findings

We identified six core categories as key action-guiding principles that illustrate the relationship between the informants' everyday smart speaker use and the interpretations they make. These categories help to explain how and based on what epistemological grounds people use smart speakers and navigate their media repertoires. The epistemological principles are interrelated and activated situationally, with some being dominant and others rather dormant in individual blends. For this reason, the way the sensemaking of smart speakers translates into use is ambivalent and sometimes even contradictory. To us, this represents a strength on the part of the insight gained because it highlights the fact that human beings are often not rational, logical, or consistent in their reasoning and actions. These contradictory actions can still be explained by the folk theories and epistemologies people have. As explained in the literature review, individual sensemaking and media use are closely interwoven

and dependent on one another. Following this view, smart speaker use and related sensemaking must be discussed together. We present the core categories of personal epistemologies (RQ 2) integrated with findings on the use of smart speakers (RQ 1) to illustrate people's varied, sometimes paradoxical, realities of media use.

In order to contextualize the interviewees' smart speaker use and sensemaking, respondents' media use is presented below within the *general findings*. Also, the participants' media repertoires and the smart speaker's role within these are briefly described, referring to RQ 1 of this paper. The reasons for the smart speakers' role within media repertoires become clear in the second section, concerning *users' communication with smart speakers and related sensemaking*. This section points out when and how smart speakers are used, referring to RQ 1, and what identified guiding principles of personal epistemologies are related to that usage, referring to RQ 2.

4.1 General findings

High interest in media technology. The respondents show a medium to very high interest in media technologies and software. In addition, most are enthusiastic about media technology innovations, observe the market, and follow trends. The interviewees describe that it is important to them to keep up with the times and be informed about the latest software and hardware. Consequently, the sample of the present study can be described as quite tech-savvy overall. Exceptions did not choose the smart agent themselves but, rather, came in contact with it through a proxy, suggesting that smart speakers have yet to diffuse into average households. Within the sample, Jens is the most innovative because he expresses a high level of interest in media technologies, lives in a highly connected smart home environment, and was an early adopter of innovations in the past. In contrast, Nadine is the least tech- and innovation-driven user within the sample. She only encounters new media technologies through her husband, her general media and smart speaker use is limited to a few simple tasks, and she seems more comfortable with analog media.

High centrality of digital media within repertoires, but smart speakers usually have a peripheral role. The relatively strong interest in media technologies within the sample is accompanied by the fact that the participants' media repertoires consist almost exclusively of digital media. Analog media, e.g., newspapers or books, may still be present in printed form, but they have a low level of relevance in repertoires and are increasingly being replaced by digital alternatives. The most central technology in the media repertoire of all respondents is the smartphone. Other key media in the informants' repertoires are laptops and tablets, followed by TVs. Three respondents own a smart watch[2], two have a gaming station, and two use a Kindle. For most interviewees, the smart speaker does not play an important role within their media repertoire. When, during the interviews, they were asked to rank their media

---

[2] Marie also says that she would like to buy a smartwatch as soon as she has the financial means to do so. David owns an outdoor watch, which differs from a smartwatch in some of its functions. For reasons of simplicity, it is also referred to as a smartwatch in this paper.

according to their importance, the smart speaker typically scored low. This was the same regardless of how the respondents initially obtained their smart speakers. Two users eventually stopped using the devices because they did not fulfill a relevant function. Additionally, respondents recurrently expressed privacy concerns, which lead to restricted use, e.g., not having a smart speaker in the bedroom or, in Klaus's case, ceasing the use of the device entirely. Overall, smart speakers are not central but, rather, peripheral in media repertoires. The reasons for this can be found in the users' sensemaking of smart speakers.

4.2 Users' communication with smart speakers and related sensemaking

The relationship between users and smart personal assistants is complex, and users have very ambivalent, sometimes contradictory, perceptions and interpretations regarding AI and smart speakers in particular. Users appear to perceive smart speakers as an entity between object and subject (see Pradhan et al. 2019, p. 15). In the following, we present the four core theoretical categories showing that, to the users, smart speakers do not appear to be friends, reliable companions, or beings with identity and agency.

(1) Comfortable insignificance: smart speakers are mainly just speakers.
Firstly, in daily life, Alexa and Siri are not used for pleasant conversations. Smart speakers do not appear as a "buddy" or a daily companion who is almost a friend. Rather, users mainly perceive smart speakers as devices that can play music and have voice control. Thus, they use them for the basic function of listening to music. For example, Emre gave his girlfriend, Marie, a smart speaker because she did not own any other speakers. Lea and her partner sold their device because they were not satisfied with the sound quality. After losing its primary utility, the smart speaker could not maintain a relevant or desirable function in their media repertoire. Apart from listening to music, the respondents mainly use smart speakers to hear summaries of the weather forecast or the news or for setting timers and alarms. The smart speakers' voice-control is handy because users are often busy with another activity, e.g., cleaning or cooking, when giving their smart speaker commands (skip or repeat songs or adjust volume). Still, regarding RQ 1, smart speakers are used mostly for tasks informants consider simple. They could easily be substituted for by other technologies or a slight change in practice. Regarding RQ 2, smart speakers are perceived as handy devices for simple needs and small routines. Interaction with smart assistants is convenient (Brause and Blank 2020, p. 7) but not considered very meaningful by the participants.

(2) Forced simplicity: smart speakers are not smart enough. Voice control provides comfort; however, informants reported several problems with their devices, mainly due the smart assistant's lack of intelligence. Technical problems such as the smart agent not reacting when being called go hand in hand with the general perception of smart assistants as often being "too stupid." That is why Andrea states, "I reckon that, when I'm doing some research on the Internet, I'm a lot more successful than any smart assistant." Nadine describes her young children asking Alexa

questions such as "Why is water blue?" which the device was not able to answer properly. Also, smart assistants are neither empathetic, nor do they have social intelligence. For example, during several interviews used in this study, the participants' smart assistant began to speak because it was activated by its name, not understanding that the user was only talking about the device and was not intending to use it. In this situation, Max appeared to be annoyed and spoke to Alexa in an unfriendly tone. Other informants also reported being annoyed, e.g., when their smart assistant talked for too long. The lack of intelligence can make communication with them "stressful" (Marie). Max explained that communication with smart assistants would be more enjoyable for him if they appeared less like a device (with technical problems) and more like an intelligent assistant:

> I think you wouldn't do that [speaking in an aggressive tone] if it was more like a smart buddy next to you that is making sense of things [i.e., commands]. [...] But when you realize that two devices are somehow making sounds, you'll become more decisive.

Here, it becomes clear that the smart speaker is not an intelligent companion for Max, despite how corporations present these devices. Users therefore do not seem to see a smart speaker as an intelligent everyday helper, an assistant, or even a friend. The smart speaker is unmistakably a technical device that, like all technologies, can malfunction. In summary, referring to RQ 1, it can be said that, when experiencing a malfunction or a lack of intelligence, users become frustrated and switch from smart speakers to other technology within their media repertoire, e.g., performing a Web search on another device or controlling audio players manually. However, the frustration and somewhat strident tone adopted when telling the device to stop a failed effort to perform a task signals two tendencies regarding RQ 2 and users' epistemological grounds: first, users expect more smartness and opportunities for interaction, and frustration stems from this deflation of expectations and experience. Users are forced to limit themselves to a very simple use. Second, users' interactions with the smart devices are not merely functionalist but emotionally colored. The unsatisfying experience of communication with the AI for smart speakers may explain why the respondents think smart assistants are quite harmless, although they know the potential AI offers for the future.

(3) Controllable infants: smart assistants are still harmless. Although all participants own or have owned a smart speaker and widely describe themselves as technology lovers, they see AI critically and express ambivalent feelings in this regard: "I see AI in a positive light, first of all. It can help a lot. But it can also do a lot of harm" (Klaus). Lea even articulated a dystopian future with regard to AI, in which it might control humankind. Only Andrea and Marie show pure optimism about AI and a potential future with it. These positions are less indicative of an actual insight into the workings and current level of AI performativity; rather, interpretations are shaped by a variety of factors. However, although smart speakers make use of an AI, the respondents are quite relaxed about the devices. Compared to the possibilities that AI in general seems to offer, they experience the AI of smart assistants as harmless: "When I interact with Alexa [...], it makes you realize how far away you

still are [from a strong AI] and how stupid it actually is—still (laughs)" (Marie). Emre therefore describes the smart agent as sort of a "newborn artificial intelligence being." As a result, regarding users' personal epistemologies of the media, it can be stated that the respondents' attitudes and worldviews about AI in general are ambivalent. Some participants can articulate well-thought-out opinions about AI and its potential and seem quite well informed. Concerning smart speakers' integrated AI, however, participants do not see any potential for harm in their intelligence, because they are perceived as quite stupid. If an AI was to rise to world domination, it would clearly not be the one they own now.

(4) Desired limitedness: users do not want their smart assistants to have agency or control. However, even if smart speakers were reliable and more intelligent, we found that users do not want their virtual assistants to have too much control or agency. On the contrary, in regard to RQ 1, informants often switched to another device when doing something more important and meaningful. The reason for media switching in this case lies in the users' sensemaking, as we aimed to reveal in RQ 2. Because interaction with smart assistants is voice based, smart speakers appear insufficient and/or unsafe for specific tasks. For example, when making a purchase on the Internet, the informants reported having the wish to also see[3] what they are buying. They would rather complete the purchase with a click, not with a spoken word.

> I would never do online shopping via Alexa, never say [...], "Alexa, put this and that on my shopping list," and somehow, I also find it [...] a bit scary to make such a transaction [...], without [...] pressing the button myself with my hand [...], like, "buy now." (Marie)

In addition, the informants criticized smart assistants for their selection of sources, for example, when reporting the news[4] or googling something.

> Alexa would give me a preselection of news media [...]. Also, I'm much faster and more effective when I [...] search for [...] it by myself, and yes, some things [...] are also more fun to look up [...] by myself [...]. (Max)

In summary, the situations described here are moments in which the artificial voice assistant takes away its user's agency. They are activities that require conscious engagement with content and therefore are more important to the user than, for example, switching off the light or setting an alarm. If activities such as making a purchase or researching the news are performed via smart speakers, the smart assistant takes considerations, decisions, and actions away from the user. This is perceived as disturbing by the informants. We thus conclude that the user's need for autonomy conflicts with the purpose and characteristics of intelligent voice control, e.g., delegation of control over a task, which can motivate media switching. This

---

[3] Only two interviewees owned an Amazon Echo Show, i.e., a smart speaker with an integrated screen; however, one of those respondents felt the screen was without purpose.

[4] Most participants liked their smart assistants to provide a short summary of the news, but when they wanted more information, they switched to another medium.

limits the range of use of smart speakers to simple, small tasks that would require little cognitive effort if performed by the users themselves.

These findings suggest that users do not to want Alexa and Siri to be their friendly daily companions. Rather, they expect smart assistants to be unobtrusive and discreet. Analyzing smart speaker use patterns, media switching, and personal epistemologies in an integrated fashion, it becomes clear that smart speakers are mostly not deemed an important element within media repertoires. They appear as a nice-to-have device, a fancy gimmick that makes some things easier and, therefore, simply must function properly. However, our study shows that users do not see smart speakers as devices like computers, smartphones[5], or TVs. A smart speaker is not a friend or sentient being and is not perceived as a distinct personality, but it is also more than any other device. This underscores how emotionally driven and sometimes paradoxical the users' epistemologies regarding smart speakers are. In what follows, we present two core categories of personal epistemologies showing that, in some situations, smart speakers are more than just any device to their users.

(5) Occasional humanization: smart assistants as interlocutors. Firstly, users do not refer to the devices by calling them "smart speakers." Rather, they use the smart assistants' names, i.e., "Alexa" or "Siri." When talking about their smart speakers, they sometimes use adjectives such as "stupid," "clever," and "funny," which are generally used to describe subjects but not objects. This shows that, to some extent, smart speakers and their integrated AI are being humanized. Users reported that the smart assistant doing something completely wrong or delivering an absurd answer may be a source of amusement. Also, testing the limits and boundaries of the smart assistant, for example, by asking impossible questions, attempting to embarrass the assistant, or provoking reactions through inappropriate language are amusing practices that have been reported (e.g., from Emre) and show that AI is not treated purely functionally. Consequently, within a limited scope, smart assistants can become occasional social interlocutors, at least for elderly, lonely, and/or physically disabled people, such as one informant's father. She gave her father a smart speaker so that he could ask Alexa about the weather or the news, which he could not easily find out by himself due to his ocular disorder. The tasks assigned to smart speakers—which are mundane for most users in the sample—may become more important for people with disabilities because they can make an otherwise tedious task easier for them to perform. In addition to the smart speaker being handy for people such as the informant's father, he was also now able to ask the smart assistant to tell him a joke when he felt lonely, his daughter reported. Even if the father's experience is reported secondhand, it is indicative of the daughter's sensemaking, who sees it as a help and support for her visually impaired father.

---

[5] Although smartphones do have smart assistants as well, most interviewees reported not using them. It appears that smart speakers are more associated with voice-control than smartphones (or tablets or notebooks) are. This could be due to the fact that smart speakers are advertised especially because of the intelligent voice-control, while smartphones and other devices with smart assistants are not.

(6) Exploitative presence: selective criticality and pragmatic trust in data protection. The fact that smart speakers are not seen in the same way as any other device is also evident when analyzing participants' general perceptions of smart speakers' existence: users perceive smart speakers as more present in the room than other devices. Marie compares the presence of a smart speaker in a room to the presence of a pet. Along with the folk theory that smart speakers record everything their users say and may sell the collected personal data, some interviewees reported not wanting smart speakers in their bedroom, although they owned several other smart speakers, e.g., in the living and dining rooms. A personal epistemology that is not constantly or generally critical but, rather, is so only in specific situations is what Schwarzenegger describes as *selective criticality* (Schwarzenegger 2020, p. 370). In addition, many interviewees are aware of potential privacy risks when using a smart speaker, but their wish for carefree everyday use leads to a *pragmatic trust* (Schwarzenegger 2020, p. 370) in the devices and their manufacturers; i.e., users tend to trust technologies on at least a pragmatic level because, otherwise, maintaining use would cause a permanent cognitive struggle: "I want to believe that I am not being tapped; therefore, I believe that I am not being tapped" (Emre). In line with other research, users claim they have nothing to hide, think their data is not interesting to anyone, or show *digital resignation* (Draper and Turow 2019) and *surveillance realism* (Dencik and Cable 2017), meaning they realize their data are collected anytime they are being digitally active and react with resignation and a sense of powerlessness. Users recognize the privacy-convenience trade-off (Lau et al. 2018, p. 13): "At the end of the day, there are also people behind this (i.e., smart speaker companies) who earn their living with it" (Jens). All this is very close to the complicated trust relationship between users and smart speaker companies described by Lau et al. (2018, p. 1). It shows that smart speakers are more than just a simple device with voice-control, because they have a special presence in a room. Users think about what smart speakers may hear, leading to the question of what the smart speaker is allowed to know, as well as the question of how to manage the use and placement of the device.

To sum up, with regard to RQ 1, we find that smart speakers are mainly used for simple tasks and basic functions, such as that of an audio amplifier. The voice control is considered handy; however, voice interaction also reveals the limited smartness of the AI. Smart speakers play a rather peripheral role in most users' media repertoires, and to perform more complex or meaningful tasks, users switch to other media. Regarding personal epistemologies, as queried in RQ 2, we note that users see smart speakers as a handy but shady gimmick, rather than as a reliable and trustworthy companion. They perceive interaction with them as not very meaningful but, rather easy to substitute for. With this being said, we found several indications that users perceive smart speakers not as just any device. Rather, they engage in complex, ambivalent, and sometimes contradictory sensemaking with regard to AI in general and smart assistants in particular. Although most of our informants declared that their smart speakers were not important to them, their use was still emotionally and morally charged. We therefore argue that smart speakers are neither friends nor simple devices for users. This is in line with a study by Pradhan et al., which shows

that the perceptions of the devices "fluidly move between human and object like perceptions" (Pradhan et al. 2019, p. 15).

5 Discussion

In this study, we analyzed how people use smart speakers and based on which epistemological principles they make sense of that use. We identified six theoretical core categories that help explain smart speaker users' practices and perceptions of living and communicating with an AI-driven media innovation. The media repertoire perspective, in conjunction with the personal epistemologies framework, allowed for an integrated perspective on user practices, folk theories, media ideologies, and sensemaking processes regarding the smart speaker in comparison to other media (see Gershon 2010; Nielsen 2016; Schwarzenegger 2020; Ytre-Arne and Moe 2021). Also, this perspective helps combine experience and practice with knowledge or beliefs nourished from various sources, as well as how they translate into action. Our results support previous research cited in this paper but also extend current studies by integrating factors that have previously been researched in mutual isolation. We are positive that research along this trajectory can increase our understanding of the epistemological grounds on which media innovations are adopted into repertoires, when and how they are maintained within them, and how they can impact the (re)configuration of the repertoire over time.

All in all, we found that smart speakers, in their current form, are typically considered too stupid or at least not smart and reliable enough to be a friend or dependable assistant, but nonetheless, their special presence makes them different from other media within the repertoire. Users exhibit ambivalent interpretations, which were presented in the form of six interrelated action-guiding epistemological principles that are activated in different blends depending on the context. They point to a complex, situational, and sometimes even paradoxical sense-making on the part of smart speaker users, which adds up to the complicated trust-relationship that Lau et al. (2018, p. 1) found smart speaker users have with their devices and manufacturers regarding data security. The analyzed smart speaker use is relatively similar to that Brause and Blank (2020) and Lau et al. (2018) describe, but with this study being the first to identify the role of smart speakers within media repertoires, our perspective highlights the fact that the smart assistant is treated differently than other technologies, which are neither held accountable for malfunctions nor elicit emotional responses, as AI does. The users' sensemaking explains the peripheral role smart speakers assume in media repertoires. As Ytre-Arne and Moe found for folk theories of algorithms, smart speaker users also experience their smart agents as confining, practical, reductive, intangible, and exploitative (see Ytre-Arne and Moe 2021, pp. 814–819). Smart speakers are considered practical due to intelligent voice control, but at the same time, they are confining and reductive because, for example, they pre-select information for their users and are unable to offer a diversity of content. As with Ytre-Arne and Moe (2021, pp. 814–819), participants experienced the devices and their use of generated data as intangible, which strained their trust in smart speakers and their manufacturing companies. Accordingly, respondents' folk

theories partially paint the devices and the companies behind them as exploitative. Regarding data protection, users are insecure, show digital resignation (Draper and Turow 2019) and surveillance realism (Dencik and Cable 2017), and express various folk theories regarding what happens with their data. However, this distrust and caution were not present all the time. Users must at least pragmatically trust that nothing bad will happen because of maintaining smart speakers in their repertoires. We therefore conclude that the individual sensemaking and use or non-use of smart speakers because of data security can partly be explained by pragmatic trust and selective criticality (see Schwarzenegger 2020).

Of course, our study has several limitations. Firstly, due to the very high level of education and the low average age of 40.4 years, the admittedly small sample of the present study is indicative of the core group of smart assistant usership but not exhaustive of all user segments. For instance, more elderly or impaired people may have to rely on AI to complete tasks because they cannot substitute and switch so easily. The companionship provided by Siri or Alexa could have a different meaning for people experiencing loneliness. Children may more easily suspend disbelief and imagine the AI to be an actual being. All this remains to be explored. A greater variety of personal epistemologies and folk theories could have been found by including a larger number of non-users in the sample. However, non-users may not yet have elaborated folk theories or any other epistemological grounds regarding communicating and living with an AI. Thus, there is a difference between non-users actively choosing not to use (based on personal epistemologies) and non-users who simply have not thought about using. As with any other self-report data, there could be biases. For instance, informants may over- or underestimate how humanely they treat the AI and how they talk to it in a natural setting. Thus, it would be an important addition to observe an actual interaction with the speaker over some time. Also, users may overemphasize the peculiarities of their smart speakers in contrast to other devices when prompted to and elaborate on epistemologies they would not otherwise report. Still, even if some sensemaking is only reported for the interview, these reports help to unveil the underlying perceptions, ideas, and attitudes linked to the use of smart speakers and how these are gradually shaping patterns of usage.

It is striking how morally charged the use of smart speakers is for many. Most informants associated the devices with a privacy risk, raised the issue on their own during the interviews, and seemed to feel the need to justify their smart speaker use. In part, this can be explained by the critical media discourse that some respondents referred to. Future research could investigate the public discourse on smart assistants and reveal the extent to which the devices are portrayed as a privacy risk. Comparisons between folk theories and personal epistemologies regarding smart assistants, in conjunction with public discourse about the potentials and perils of AI, should also be informed by transnational and transcultural research collaborations.

Because the technology is likely to improve over time, it is also likely that folk theories regarding smart assistants will evolve. However, we deduce that, for future smart speaker adoption, whether smart assistants will be successfully interpreted as relevant aids or even considered suitable and trustworthy companions, beyond basic functions and simple tasks, will be crucial. This will determine whether we imagine

ourselves using a suspicious device in the future or interacting and communicating with a machine assistant, a confidant, or maybe even a digital friend.

References

Adorján, J. (2021). Der Feind hört mit. Süddeutsche Zeitung Magazin, June 3rd. https://sz-magazin.sueddeutsche.de/gute-frage/smart-speaker-alexa-bedenken-ueberwachung-90267. Accessed 14 Sept 2021.

Alizadeh, F., Esau, M., Stevens, G., & Cassens, L. (2020). eXplainable AI: Take one step back, move two steps forward. In C. Hansen, A. Nürnberger & B. Preim (Eds.), *Mensch und Computer 2020 – Workshopband* (pp. 1–6). Gesellschaft für Informatik e. V.. https://doi.org/10.18420/muc2020-ws111-369.

Bager, J. (2017). Google Home in Deutschland: Eine ernstzunehmende Konkurrenz für Amazons Echo. heise online, August 8th. https://www.heise.de/newsticker/meldung/Google-Home-in-Deutschland-Eine-ernstzunehmende-Konkurrenz-fuer-Amazons-Echo-3794806.html. Accessed 11 Sept 2021.

Beyto (2020). Beyto smart speaker Studie 2020 | Deutschland. Digitalagentur Beyto. https://www.beyto.com/smart-speaker-studie-2020/. Accessed 11 Sept 2021.

Beyto (2021). Beyto smart speaker & voice studie 2021–2022 | Deutschland. Digitalagentur Beyto. https://www.beyto.com/smart-speaker-voice-studie-2021-2022/. Accessed 11 Sept 2022.

Boczkowski, P. J., Matassi, M., & Mitchelstein, E. (2018). How young users deal with multiple platforms: The role of meaning-making in social media repertoires. *Journal of Computer-Mediated Communication*, *23*(5), 245–259. https://doi.org/10.1093/jcmc/zmy012.

Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI friend: How users of a social chatbot understand their human-AI friendship. *Human Communication Research*, *48*(3), 404–429. https://doi.org/10.1093/hcr/hqac008.

Brause, S. R., & Blank, G. (2020). Externalized domestication: Smart speaker assistants, networks and domestication theory. *Information Communication and Society*, *23*(5), 1–13. https://doi.org/10.1080/1369118X.2020.1713845.

Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44. https://doi.org/10.1080/1369118X.2016.1154086.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 1–12. https://doi.org/10.1177/2053951715622512.

Dencik, L., & Cable, J. (2017). The advent of surveillance realism: public opinion and activist responses to the Snowden leaks. *International Journal of Communication*, *11*, 763–781.

Draper, N. A., & Turow, J. (2019). The corporate cultivation of digital resignation. *New Media & Society*, *21*(4), 1824–1839. https://doi.org/10.1177/1461444819833331.

Draucker, C. B., Martsolf, D. S., Ross, R., & Rusk, T. B. (2007). Theoretical sampling and category development in grounded theory. *Qualitative Health Research*, *17*(8), 1137–1148. https://doi.org/10.1177/1049732307308450.

Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kirlik, A. (2016). First I "like" it, then I hide it: Folk Theories of Social Feeds. *CHI'16: Proceedings of the 2016 CHI*

Conference on Human Factors in Computing Systems, 2371–2382. https://doi.org/10.1145/2858036.2858494.

Ferrario, A., Loi, M., & Vigano, E. (2020). In AI we trust incrementally: A multi-layer model of trust to analyze human-artificial intelligence interactions. *Philosophy & Technology*, *33*, 523–539. https://doi.org/10.1007/s13347-019-00378-3.

Fuest, B. (2019). Wie Amazon-Mitarbeiter Ihre Alexa-Aufnahmen mithören.WELT, April 11th. https://www.welt.de/wirtschaft/webwelt/article191741517/Amazons-Echo-Mitarbeiter-hoeren-Alexa-Aufnahmen-mit.html. Accessed 14 Sept 2021.

Gelman, S. A., & Legare, C. H. (2011). Concepts and folk theories. *Annual Review of Anthropology*, *40*, 379–398. https://doi.org/10.1146/annurev-anthro-081309-145822.

Gershon, I. (2010). Breaking up is hard to do: Media switching and media ideologies. *Journal of Linguistic Anthropology*, *20*(2), 389–405. https://doi.org/10.1111/j.1548-1395.2010.01076.x.

Glaser, B. G., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. London: Weidenfeld & Nicolson.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 1–92. https://doi.org/10.5465/annals.2018.0057.

Hasebrink, U., & Domeyer, H. (2012). Media repertoires as patterns of behaviour and as meaningful practices: A multimethod approach to media use in converging media environments. *Participations. Journal of Audience & Reception Studies*, *9*(2), 757–779.

Hasebrink, U., & Hepp, A. (2017). How to research cross-media practices? Investigating media repertoires and media ensembles. *Convergence: The International Journal of Research into New Media Technologies*, *23*(4), 362–377. https://doi.org/10.1177/1354856517700384.

Kowalczuk, P. (2018). Consumer acceptance of smart speakers: A mixed methods approach. *Journal of Research in Interactive Marketing*, *12*(4), 418–431. https://doi.org/10.1108/JRIM-01-2018-0022.

Lau, J., Zimmerman, B., & Schaub, F. (2018). Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1–31. https://doi.org/10.1145/3274371.

Lee, H., & Cho, C.-H. (2020). Uses and gratifications of smart speakers: Modelling the effectiveness of smart speaker advertising. *International Journal of Advertising*, *39*(7), 1150–1171. https://doi.org/10.1080/02650487.2020.1765657.

Liao, Y., Vitak, J., Kumar, P., Zimmer, M., & Kritikos, K. (2019). Understanding the role of privacy and trust in intelligent personal assistant adoption. In N. G. Taylor, C. Christian-Lamb, M. H. Martin & B. Nardi (Eds.), *Information in contemporary society* (Vol. 11420, pp. 102–113). Berlin: Springer.

Lomborg, S., & Kapsch, P. H. (2020). Decoding algorithms. *Media, Culture & Society*, *42*(5), 745–761. https://doi.org/10.1177/0163443719855301.

Madianou, M., & Miller, D. (2012). Polymedia: Towards a new theory of digital media in interpersonal communication. *International Journal of Cultural Studies*, *16*(2), 167–187. https://doi.org/10.1177/1367877912452486.

Malkin, N., Deatrick, J., Tong, A., Wijesekera, P., Egelman, S., & Wagner, D. (2019). Privacy attitudes of smart speaker users. *Proceedings on Privacy Enhancing Technologies*, *4*(4), 250–271. https://doi.org/10.2478/popets-2019-0068.

McCloskey, D., & Bennett, G. (2020). Examining artificial intelligence acceptance: Exploring smart speaker adoption with the unified theory of acceptance and use of technology (UTAUT2). *Pennsylvania Economic Review*, *27*(2), 44–70.

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa... examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, *99*, 28–37. https://doi.org/10.1016/j.chb.2019.05.009.

Naab, T. (2013). *Naive Medientheorien und Third-Person Perception. Eine Untersuchung zur Integrierbarkeit beider Konzepte.* Baden-Baden: Nomos.

Natale, S. (2020a). To believe in Siri: A critical analysis of AI voice assistants. Communicative Figurations. Working Paper No. 32, 3–17. https://www.kommunikative-figurationen.de/fileadmin/user_upload/Arbeitspapiere/CoFi_EWP_No-32_Simone-Natale.pdf. Accessed 14 Sept 2021.

Natale, S. (2020b). Communicating through or communicating with: Approaching artificial intelligence from a communication and media studies perspective. *Communication Theory*. https://doi.org/10.1093/ct/qtaa022.

Neville, S. J. (2020). Eavesmining: A critical audit of the Amazon Echo and Alexa conditions of use. *Surveillance & Society*, *18*(3), 343–356. https://doi.org/10.24908/ss.v18i3.13426.

Nielsen, R. K. (2016). Folk theories of journalism. The many faces of a local newspaper. *Journalism Studies*, *17*(7), 840–848. https://doi.org/10.1080/1461670X.2016.1165140.

Peters, C., & Schrøder, K. C. (2018). Beyond the here and now of news audiences: A process-based framework for investigating news repertoires. *Journal of Communication*, *68*(6), 1079–1103. https://doi.org/10.1093/joc/jqy060.

Pradhan, A., Findlater, L., & Lazar, A. (2019). "Phantom friend" or "just a box with information": Personification and ontological categorization of smart speaker-based voice assistants by older adults. *PACM on Human-Computer Interaction*, *3*(CSCW), 1–21. https://doi.org/10.1145/3359316.

Pridmore, J., & Mols, A. (2020). Personal choices and situated data: privacy negotiations and the acceptance of household intelligent personal assistants. *Big Data and Society*. https://doi.org/10.1177/2053951719891748

Pridmore, J., Zimmer, M., Vitak, J., Mols, A., Trottier, D., Kumar, P. C., & Liao, Y. (2019). Intelligent personal assistants and the intercultural negotiations of dataveillance in platformed households. *Surveillance & Society*, *17*(1/2), 125–131.

Rader, E., & Gray, R. (2015). Understanding user beliefs about algorithmic curation in the Facebook news feed. In *CHI'15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 173–182). https://doi.org/10.1145/2702123.2702174.

Reichelt, P., & Hegemann, L. (2019). Alexa, hörst du grad zu? ZEIT ONLINE, October 18th. https://www.zeit.de/digital/datenschutz/2019-09/smart-speaker-amazon-echo-google-home-daten. Accessed 14 Sept 2021.

Schwarzenegger, C. (2020). Personal epistemologies of the media: Selective criticality, pragmatic trust, and competence-confidence in navigating media repertoires in the digital age. *New Media & Society*, *22*(2), 361–377. https://doi.org/10.1177/1461444819856919.

Siles, I., Segura-Castillo, A., Solís, R., & Sancho, M. (2020). Folk theories of algorithmic recommendations on Spotify: Enacting data assemblages in the global South. *Big Data & Society*, *7*(1), 1–15. https://doi.org/10.1177/2053951720923377.

Strauss, A. L. (1998). *Grundlagen qualitativer Sozialforschung* (2nd edn.). München: W. Fink.

Tenzer, F. (2022a). Marktanteile der Hersteller am Absatz von intelligenten Lautsprechern weltweit vom 3. Quartal 2016 bis zum 3. Quartal 2021. Statista, 26. Januar. https://de.statista.com/statistik/daten/studie/818999/umfrage/marktanteile-der-hersteller-am-absatz-von-intelligenten-lautsprechern-weltweit/. Accessed 18 Apr 2022.

Tenzer, F. (2022b). Statistiken zum Thema Smart Speakers. Statista, 25. Januar. https://de.statista.com/themen/4662/smart-speakers/#topicHeader__wrapper. Accessed 18 Apr 2022.

Toff, B., & Nielsen, R. K. (2018). "I just Google it": Folk theories of distributed discovery. *Journal of Communication*, *68*(3), 636–657. https://doi.org/10.1093/joc/jqy009.

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837–851. https://doi.org/10.1177/1077800410383121.

Trenholm, R. (2016). Amazon Echo (and Alexa) arrive in Europa, and Echo comes in white now too. c|net, 14. September. https://www.cnet.com/tech/home-entertainment/amazon-echo-and-alexa-arrives-in-europe/. Accessed 11 Sept 2021.

Ytre-Arne, B., & Moe, H. (2021). Folk theories of algorithms: Understanding digital irritation. *Media, Culture & Society*, *43*(5), 807–824. https://doi.org/10.1177/0163443720972314.

Marlene Wassmer holds a master's degree in Media and Communication Studies from the University of Augsburg. She currently works outside of academia.

Dr. habil. Christian Schwarzenegger holds a PhD and habilitation in communication from the University of Augsburg, Germany. He is a lecturer "Akademischer Rat" at the Department of Media, Knowledge and Communication at the University of Augsburg.