HUMAN-MACHINE COMMUNICATION

Human-Machine Communication
Volume 6, 2023
https://doi.org/10.30658/hmc.6.10

Boundary Regulation Processes and Privacy Concerns With (Non-)Use of Voice-Based Assistants

Jessica Vitak[1], **Priya C. Kumar**[2], **Yuting Liao**[3], **and Michael Zimmer**[4]

1 University of Maryland, College Park, Mayland, USA
2 Pennsylvania State University, State College, Pennsylvania, USA
3 Intuit Inc.
4 Marquette University, Milwaukee, Wisconsin, USA

Abstract

An exemplar of human-machine communication, voice-based assistants (VBAs) embedded in smartphones and smart speakers simplify everyday tasks while collecting significant data about users and their environment. In recent years, devices using VBAs have continued to add new features and collect more data—in potentially invasive ways. Using Communication Privacy Management theory as a guiding framework, we analyze data from 11 focus groups with 65 US adult VBA users and nonusers. Findings highlight differences in attitudes and concerns toward VBAs broadly and provide insights into how attitudes are influenced by device features. We conclude with considerations for how to address boundary regulation challenges inherent in human-machine interactions.

Keywords: privacy, internet of things, voice-based assistants, communication privacy management theory

Acknowledgments

The authors thank Katie Chamberlain Kritikos for her contributions to the project. This study was funded by National Science Foundation under Grant Nos. 1640640 and 1640697.

CONTACT Jessica Vitak • jvitak@umd.edu • 4130 Campus Drive, College Park, MD, 20742, USA

ISSN 2638-602X (print)/ISSN 2638-6038 (online)
www.hmcjournal.com

Introduction

Recent years have seen an explosion in the Internet of Things (IoT) and smart technologies designed to simplify people's lives. IoT refers to a network of interconnected computing components, digital and mechanical objects, and living organisms; each *thing* is given a unique identifier enabling data transfer over the network (Alhammadi et al., 2019). These smart technologies are everywhere—from workplaces to homes, cars, and schools—and include smart light switches, appliances, thermostats, digital assistants, door locks, and more. They enable people to remotely complete routine tasks, such as turning lights on/off, checking refrigerator contents, or adjusting a home's temperature. Importantly, in order to offer such functionality, IoT devices collect and transmit significant amounts of data about people and their environment.

Although IoT devices provide significant utility and convenience, they also raise concerns about what data is being collected, how that data is stored, to whom that data is transmitted, what control users have managing that data, and how that data might be used in the future. These devices contain a variety of sensors that collect audio, location, movement, and other trace data. Analysis of such data can reveal information about people's likes and dislikes, eating and exercise habits, location, and more (Boeckl et al., 2019). Privacy threats include platforms misusing data collected from IoT devices (Lynskey, 2019), law enforcement unexpectedly accessing IoT devices (Díaz, 2020), and harm from intimate partners (Levy & Schneier, 2020).

In this paper, we focus on one of the most popular IoT interfaces—voice-based assistants (VBAs) found in smartphones and smart speakers in millions of homes. Because of their popularity, VBAs provide an important exemplar of human-machine communication; users interact with a human-like conversational user interface to achieve tasks (Guzman, 2019; Weidmüller, 2022). As companies design newer versions of these devices, they introduce new features that collect a wider range of data through more channels, especially once users start linking multiple smart devices together. While marketed as increasing convenience, the influx of audio, video, and other sensor data creates new privacy risks for people deciding which smart devices to use and how to interact with them.

Using data collected from 11 focus groups ($N = 65$) with both VBA users and nonusers in the US, we evaluate how those who regularly use these devices—as well as those who have chosen not to use them—feel about these advances in device features, as well as the wider implications of the growth of IoT technology. We interpret our findings using Petronio's (2002) Communication Privacy Management theory (CPM), which considers the tensions individuals experience when sharing private information and the turbulence that arises when privacy rules are broken. While this theory has largely focused on interpersonal communication, we extend it to human-machine communication to explore factors people consider when deciding whether to use VBAs. We argue CPM provides a useful framework for considering what *ownership* and *control* mean when data is shared with a company rather than an individual, and we reflect on how companies may address challenges with boundary regulation in their features and policies.

Literature Review

One of the most common applications of IoT is in home automation, providing users with convenient ways to manage home appliances, lights, power outlets, door locks, and other smart devices (Zeng et al., 2017). Smart home devices can be managed through a mobile or web interface, or through voice commands to smart speakers via voice-based assistants (VBAs). Use of VBA-powered smart speakers has grown: 33% of the US population (age 12+) owned a smart speaker in 2022, up from 27% in 2020 and 18% in 2018 (Edison Research, 2022). People use VBAs most frequently to access music, conduct hands-free searches, and control other devices connected to smart hubs (Ammari et al., 2019). VBAs can be customized to make routines more efficient, such as lowering lights and playing soothing music at bedtime, and they can support caregiving and accessibility for older adults and people with disabilities (Pradhan et al., 2018).

Privacy Concerns With Smart Home Devices

Although useful in many scenarios, smart devices blur boundaries between public and private spaces, and scholars have started exploring how users and nonusers understand the privacy implications of integrating "always listening" VBAs in home environments. In evaluating the public's understanding of privacy issues related to social robots, Lutz & Tamò-Larrieux (2020) found respondents were most concerned about personal information shared with device manufacturers. This study was largely based on nonusers of smart devices; similarly, Lau et al. (2018) found that nonusers saw little utility in VBAs/smart speakers and were less trusting of service providers, while McLean and Osei-Frimpong (2019) found that perceived privacy risks of smart speakers significantly dampened perceptions of device benefits.

Conversely, researchers have found that VBA users generally have low privacy concerns regarding their smart devices (Lutz & Newlands, 2021). Compared to nonusers, users report higher confidence that companies will ensure the privacy, safety, and security of their data (Liao et al., 2019). VBA users often have a limited understanding of how the systems collect, store, and analyze their data (Lau et al., 2018; Zeng et al., 2017), and news articles have highlighted how data sharing, access, and use by these companies may be surprising or problematic to users (e.g., Day et al., 2019; Fowler, 2018). Likewise, Ammari et al. (2019) found that respondents frequently could not articulate specific privacy concerns; when they did, concerns centered on uncertainty about when the device was listening and third parties accessing VBA data. Zeng et al. found that respondents rationalized this lack of concern as not feeling personally targeted, trusting potentially adversarial actors, and believing their existing mitigation strategies were sufficient.

These studies highlight that smart device users often express few privacy concerns, but their rationalizations suggest an incomplete understanding of privacy risks, a complicated trust relationship with VBA companies, and reliance on the sociotechnical context in which VBAs reside. Building on this, Easwara Moorthy and Vu (2015) found that when users understand privacy risks associated with VBAs, they attempt to mitigate concerns by using simple strategies (e.g., only using in private spaces). While their research calls for better design of VBAs to account for such user practices, subsequent work suggests that privacy controls are infrequently used and not aligned with user needs (Lau et al., 2018; Malkin et al., 2019).

Framing the Study: Communication Privacy Management Theory

Numerous privacy frameworks provide a means for evaluating how users navigate privacy concerns related to technology. For example, privacy calculus (Culnan, 1993; Laufer & Wolfe, 1977) describes the cost-benefit analysis individuals make when deciding whether to share personal information. In the case of VBAs, privacy calculus would argue that consumers engage in a rational analysis of the risks and benefits of using a smart speaker; if the benefits outweigh the risks, they are more likely to use it. Alternatively, Nissenbaum's (2009) theory of privacy as contextual integrity (CI) asserts that interactions occur in particular contexts, and norms govern people's expectations of how personal information should flow within any given context. If a technology or practice disrupts those norms, it could pose a privacy concern, irrespective of whether the information was public or private. VBAs, which record and transmit audio data to a third party, might represent a disruption of existing informational norms. In the current study, we rely on Petronio's (2002) Communication Privacy Management Theory (CPM) to evaluate how users' privacy calculus is impacted by the types of contextual disruptions CI highlights.

Building on work by Altman (1975), Petronio (2002) argues that people engage in a "mental calculus" when making information disclosure decisions. CPM provides insights into how people navigate tensions between revealing and concealing information—tensions that might erupt with changes in pre-existing contextual norms. While CPM is an interpersonal communication theory, we can extend its principles to human-machine communication and interactions between users and smart devices. We argue that such an extension is useful, given that the anthropomorphization of VBAs leads many users to perceive them as social beings (Guzman, 2019).

CPM provides five core assumptions regarding the relationship between individuals and their private information (Petronio, 2002; Petronio et al., 2021). First, people believe they own and have the right to control access to their private information. A smart speaker user would therefore believe they own and control any data collected by their device, including voice commands. Second, people employ privacy rules to control their private information. Privacy rules are generally organized into three categories related to boundary permeability, ownership, and linkage (Xu et al., 2022). When considering interactions between a user and a smart speaker, less control is possible than in interpersonal communication. How privacy rules are enforced is unclear and relies heavily on whether a user trusts the company with whom they share their data.

The third and fourth assumptions of CPM note that private information, once shared, becomes co-owned, and co-owners negotiate rules regarding if, when, and how information can be further shared. Companies' privacy policies provide a legal framework for how they manage that co-ownership; however, a variety of scenarios may cause misunderstandings and rule breakdowns. Fifth, when rules are violated, boundary turbulence arises and may cause relational tensions and a breach of trust. Such turbulence may be challenging to navigate when a company breaks a privacy rule. While the easiest way to resolve turbulence would be to stop using a device, that may not be a feasible solution.

Two recent survey studies have explored how CPM may apply to human-machine communication, and specifically to smart speaker use. Xu and colleagues (2022) found that smart speaker users employed two types of privacy rules when interacting with devices: privacy settings review (ownership rule) and limiting access (permeability rule). However, linkage rules were not observed, likely because users can rarely negotiate with companies regarding data sharing. Kang and Oh (2021) also explored the role that perceived benefits and risks played in the use of privacy management strategies. They found that privacy self-efficacy had a moderating effect on the employment of these strategies; those with high self-efficacy were more likely to engage in higher disclosure and higher boundary control.

Taken together, this prior work on VBAs illustrates how privacy concerns might influence people's adoption and use of smart home devices and VBAs. Specifically, nonusers might be more sensitive to privacy issues, while users might value social and utilitarian benefits over privacy (McLean & Osei-Frimpong, 2019; Zheng et al., 2018) or trust the company to mitigate lingering privacy concerns (Liao et al., 2019). In this paper, we extend this prior work to consider the role that perceived privacy risks play in VBA (non-)adoption and how the addition of advanced features in VBA-embedded devices affects perceptions of privacy risks. Specifically, we ask:

> **RQ1:** How do users and nonusers navigate privacy concerns related to VBAs?
>
> **RQ2:** How do users' and nonusers' attitudes toward VBAs shift as new features are added that collect more types of data?

Method

This study was conducted at two US public universities: one located on a suburban campus in the eastern US with 41,000 students, the other on an urban campus in the midwestern US with 24,000 students. In January 2018, the authors obtained a random sample of approximately 3,000 university staff at each university and invited them to complete a survey about their VBA use. To help ensure a diverse pool of adult participants, the sample population included all university staff levels, but excluded faculty and undergraduate student employees. Participants could enter their email addresses if they were interested in joining a follow-up focus group. We received survey responses from 1,160 people, and 705 expressed interest in a follow-up study.[1]

We chose focus groups because they are especially useful for exploring perceptions and generating ideas (Straus, 2019). They also provide a natural setting for participants to interact, respond to, and build on others' comments (Krueger, 2014). To maximize the diversity of perspectives, we used criterion sampling (Patton, 2002). We first divided prospective participants into groups based on whether they used home-based VBAs, phone-based VBAs, both, or neither. We then created three types of sessions: (1) users only, (2) nonusers only, and (3) a mix of users and nonusers. We conducted 11 focus groups (2–8 participants per group) with 65 people across the two universities. See Table 1 for session details.

[1] See Liao et al. (2019) for an analysis of the survey data.

TABLE 1 Descriptive Data for Focus Group Sessions

Focus Group #	Group Type	Number of Participants	Gender (% male)	User Type (% VBA user)	Age Mean (SD)
Group 1	User Only	4	25%	100%	41.75 (11.84)
Group 2	User Only	7	14%	100%	39.14 (11.28)
Group 3	Mix	6	50%	50%	39.67 (14.15)
Group 4	Mix	6	50%	33%	36.00 (12.08)
Group 5	Mix	8	63%	38%	38.13 (14.23)
Group 6	User Only	6	50%	100%	39.50 (15.15)
Group 7	Nonuser Only	4	25%	0%	35.25 (16.68)
Group 8	User Only	8	25%	100%	35.13 (11.49)
Group 9	Mix	8	63%	75%	31.38 (9.16)
Group 10	Mix	6	17%	67%	37.33 (8.94)
Group 11	Nonuser Only	2	50%	0%	50 (12.76)
Totals	**4 User, 5 Mixed, 2 Nonuser**	**65**	**40%**	**66%**	**37.45 (11.23)**

Each session lasted 1 hour and included a semi-structured protocol, starting with questions about participants' general attitudes toward new technologies, followed by a discussion about their use (or non-use) of VBAs. For each session, a moderator from the research team guided the participants through the prepared questions, while a second team member observed and took notes. Participants viewed a commercial for the newly released Amazon Echo Show—which includes a screen, camera, and additional integrations with other smart devices—and shared their reactions. We chose the Echo Show because it encapsulated broader trends in IoT development, including advanced audio and visual features and deeper links into ecosystems of devices and accounts. In some sessions, participants also discussed the Echo Look, a recently released device at the time of data collection that included a camera and was marketed as a tool to upload pictures of outfits and get fashion advice from peers. At the conclusion of each session, participants received a US$15 Amazon gift card.

TABLE 2 Subset of First-Round Codes From Qualitative Analysis of Focus Groups

Code Name	Code Description
Compare VBAs	Explicit statements comparing features of or attitudes toward two or more versions of VBAs (e.g., Siri, Home, Echo Show).
Privacy-Security	Talking broadly about how technology affects privacy, security, surveillance, and related topics. Strategies used to attain desired level of privacy/security. Comments about corporations using/accessing their data.
VBA Listening	Responses to question, "Do you have a sense of when these devices are listening for your voice or if they're always listening?" General comments about VBA microphones and their capabilities, as well as concerns about when VBAs are capturing audio data or what happens to that data.
Nothing to Hide	Comments that there are minimal risks to using VBAs (e.g., "life is boring").
Privacy Apathy	Comments reflecting belief that privacy is dead, we're already tracked in many ways, etc.
Echo Show	Comments and discussion after watching the Echo Show commercial.

Sessions were audio recorded, and files were transcribed and imported into Dedoose for qualitative analysis. Data analysis included two cycles of coding (Miles et al., 2014). First, the research team developed an initial codebook based on the interview protocol and researcher notes from the sessions (*provisional* or *protocol* coding). Each team member coded a transcript separately, noting where new codes could be added or existing codes collapsed. The team met to refine and finalize the codebook. Two team members then coded each transcript, with the team meeting regularly to resolve coding differences by consensus. For the second cycle, the team identified six codes relevant to this study's research question (listed in Table 2). Excerpts were exported into Excel, and each team member selected specific codes and analyzed the excerpts for patterns (*pattern coding*). For instance, one pattern in the *VBA listening* code was perceptions of home VBAs as more invasive than phone-based VBAs. The team discussed these patterns and linked them to the research questions to identify key themes related to the research questions. All participant names reported below are pseudonyms to protect participant identities.

Findings

RQ1: Rationalizing Privacy Concerns in VBA (Non-)Adoption

We observed notable differences in how VBA users and nonusers talked about privacy concerns. Aligning with and extending prior work, we found that users focused more on the benefits of the technology—often downplaying privacy risks because they felt the data was not sensitive, or felt they lacked any meaningful ability to control data collection in the first place—while nonusers described privacy concerns as one of the reasons they avoided VBAs.

VBA users lacked a sense of data sensitivity and felt little ability to control their data, leading to lower privacy concerns and a focus on utility. While data ownership and control are key components of CPM, many VBA users expressed little interest in managing their voice data due to a perception that the data is not sensitive—and thus posed no risks. For example, James said, "There's nothing I would share that Alexa would hear that would embarrass me at any point in time." Likewise, Emma said she doesn't worry about potential security risks from these devices because she is not doing anything to warrant attention: "I'm boring. I don't have my ballistic missiles sitting in my living room." Others described their lives as "uninteresting" and unworthy of government focus, as when Jackie said, "I live a very boring and average life. I would probably never be tagged by the FBI or anything like that because I don't do anything." These comments align with the nothing to hide trope (Madden & Rainie, 2015), which argues that only "bad" people have things they want to keep private. For example, John said, "If you're gonna be that concerned about a device listening in, chances are you're probably doing something you really don't want people overhearing."

Others' comments referenced a bigger challenge with data ownership: as VBAs are merely the latest in an ongoing expansion of data-hungry technologies, some felt they no longer own their data—and thus lack ways to meaningfully control it. For example, Charlotte said,

> I think there are video cameras on every street. They are watching us everywhere; they are listening to our every peep and move . . . I guess I don't know how to prevent that or what to think about it. It just doesn't seem like there's a lot of privacy anymore.

Some users framed potential privacy risks in relation to other privacy/security threats, rationalizing their VBA use in ways that reflected broader attitudes toward privacy that go beyond data shared through device interaction. As Anthony noted, "The bigger security concern is if I use Alexa to purchase something. Is that machine any more vulnerable when I put my credit card into a dozen different websites? That level of security is what I'd be most worried about."

In light of their perceived lack of control, these participants may have instead prioritized the perceived benefits of VBAs as part of their privacy calculus. Brian reflected broadly on this when he said,

> no matter what technology you use, I feel like if they want to find something, they can find out . . . your phone is tracked wherever you go, so they can tell you your whole life story if they wanted to.

Participants also shared examples that highlighted their lack of control. Kyle noted that data breaches at major corporations suggest that our data is already "out there," while Anne spoke about searching for something on Google only to see ads for that product on other sites.

The belief that data collection and surveillance are omnipresent—and that individuals have little control over what data is collected and who has access to it—led to a sense of apathy and resignation toward data collection among many people we spoke with. Jackie said "it's useless to fight" to protect personal data, and that the increasing reliance on technology will lead future generations to "be even more used to technology . . . People are just going to accept this information." Veronica echoed this sentiment, saying, "I don't think there's running away from technology that we can do efficiently in this age, and I don't mind."

Veronica's comment that she "doesn't mind" technological advances was reflected in several comments that align with the privacy calculus people engage in when deciding if and how to use technologies. For example, Adam said, "I feel like a lot of these companies are collecting these data anyways. I don't like that they do, but if they're going to collect it, I'd rather get the most utility out of it as possible." In that same session, Jay added, "I realized if I'm gonna have a modern smartphone, I'm always gonna have that technology and I can't guarantee it's turned off, so I might as well use it. I mean, it's built in—there's no escaping it."

Nonusers stressed the need for trustworthy providers and control over access to information before they would consider adopting VBAs. While many VBA users shared feelings of resignation toward data collection, those who had *not* adopted VBAs expressed a range of privacy concerns when describing their decision not to use them. Participants' comments referenced trust-related concerns, as well as a desire to control access to their data, reflecting the need to mitigate potential boundary turbulence before adopting VBAs.

Nonusers referenced their use of other Google or Amazon services and data they already shared with these companies. Unlike VBA users, who rationalized their use by saying the company already had their data, nonusers wanted to minimize the data these tech giants had about them, so their privacy calculus was somewhat different. Jada said, "I have a Google phone and Google accounts. I feel like Google knows everything about my life. But I still worry about setting myself up to use a device that would know more information about me." Trust also played an important role, which Gwen noted:

> I think there's a bit of a trust factor for me. I don't really trust the corporations, so I'm only willing to let them into parts of my life where I'm like, "Okay, this is really useful." And I also think as we get more smart devices around our home, it's just easier for them to be hacked, and I think that's going to happen more and more.

Likewise, Leah expressed concerns about trading personal information for minimal benefits, like using VBAs to play music:

> It's one more thing that is used to collect data on you; I assume it's one more thing that can be hacked. I'm old-fashioned. I'm happy with the radio and CDs [compact discs]. I can take those extra four steps to the radio or CD player and turn it on.

At the time of data collection, several media reports had identified bugs with Amazon's Echo devices, including a heavily covered story of Alexa laughing without being prompted (Chokshi, 2018). From a CPM perspective, such accounts can be viewed as instances of turbulence, as they violate people's expectations of how the device works, what data it collects, and how it uses that data. In interpersonal relationships, individuals may re-negotiate rules following such turbulence; in the case of VBA nonusers, such stories may reify their choice. Cliff shared:

> When the review units of the . . . Google Home Mini went out, the button was constantly pushed to listen by manufacturing defect. So here's a device that's constantly listening and they get updates continuously from the server. Let's say somebody wanted to change it; how hard would that be to get it to change?

Walter stopped using Google Assistant after hearing concerning news stories "of people just mentioning certain words and suddenly, boom, the phone's responding." He also worried about weak security protocols in IoT devices making everything more vulnerable: "I don't want to have the ability to turn on and off a light and someone can come in and steal what's on my hard drive."

Other participants worried about unknowns associated with these devices, including how their data could be used in the future and security risks posed by the wider IoT ecosystem. Wade pointed to the newness of these technologies and the lack of existing legislation to protect consumers:

> Probably the biggest drawback for me in terms of not wanting to get one is there's a lot of unknowns, it's all pretty new. Until there's legal precedent, or more history behind it, I don't really want to jump into it.

Likewise, Nina felt the lack of clarity in data collection processes was unnerving, saying, "I don't want a corporation listening to what's going on in my household. I don't know what it's recording. I don't know what's being done with that information."

RQ2: Shifts in Privacy Attitudes Across Types of VBA Devices

Our second research question considered how participants responded to advances in VBAs' features. Initially only available on smartphones, VBAs have expanded to a variety of home devices, including versions with cameras and screens. Features in newer versions of smart speakers aim to reduce friction between users and the task they want to accomplish, which requires greater access to user data and complicates communication processes. Participants discussed their (dis)comfort with these features, and across both users and nonusers, they described newer VBAs—and smart technologies more broadly—as increasingly "creepy," which echoes previous research looking at user perceptions of data collection by mobile apps (Shklovski et al., 2014).

As devices move from phones to homes, friction decreases and privacy concerns increase. During each focus group, we began by discussing phone-based VBAs, including Apple's Siri and Google's Assistant. Most participants reported using phone-based VBAs at some point, although they often described technical issues that limited device utility. For example, participants described having a hard time accomplishing tasks, like when Jordan said he didn't use Siri much because "she didn't really accomplish [requests I gave her] well." Jordan used both the Amazon Echo and Google Home and was much more favorable toward home-based VBAs.

Some participants referenced specific VBA features when describing their concerns. For example, Jin said, "I don't feel like Siri is listening [all the time], because she doesn't turn on unless I press my home button and say 'Hi, Siri.'" Erika echoed this, saying, "I don't have an Alexa or Google Home. But I have Google [Assistant] on my phone . . . and I really like that I have to trigger it." Renee suggested that explicit triggering features kept VBAs from entering "creepy" territory: "If you have to trigger it, it's not creepy. . . . I don't mind saying 'Okay, Google,' but if it's still listening and I don't want it to be listening anymore, that's creepy." Importantly, different VBAs have different activation features, but home devices are typically activated by voice alone, whereas the original versions of Siri and Google Assistant required users to hold down a button to activate the feature. Home VBAs may have a "mute" button, but this significantly reduces the utility of the device, and prior research suggests they are not widely used (Lau et al., 2018).

Many participants expressed concern that their speakers were always listening—not just when they spoke the activation phrase—based on personal experiences. For example, Marilyn said, "She's [Alexa] definitely always listening because randomly she thinks she hears 'Alexa' but we never said that and she will start talking. In that aspect, it's clear that they are always listening and who knows if they are saving [it]." Relatedly, some users expressed concerns that anyone could trigger the device, like when Faith described a movie setting off her Echo device: "It's kind of creepy because we'd be watching in the living room and the dad would shout the daughter's name [Alexis] and all of a sudden you'd hear, 'I'm sorry, I didn't quite catch that.'"

Addressing these perceived risks requires trust between users and the companies providing these devices, especially given that it is often unclear what data is being collected and how it is used. But this also raises questions of whether the companies *should* be trusted. This sentiment was highlighted by Huong, who said, "We're trusting Google that what they show me . . . is what they kept. For the most part, I trust Google on that, and Amazon. But there's that open concern; it's like, what are you opening yourself up to?" Building on this, participants expressed concerns about not knowing *when* these devices were listening and *how much* they captured. Jackie said:

> . . . it's always listening for you to say "Alexa." Do I really know it's not listening to other things? What if it's listening to a conversation about my religious or political beliefs and it's tagging things? I don't want to sound paranoid, but I really don't trust corporations and I don't trust the government to not do those things just because they say it's wrong.

Because of these concerns, several participants said they refused to put home-based VBAs in particularly private places like bedrooms. James said he wouldn't even put a TV in his room because of privacy concerns. Likewise, Chen described why she removed her Echo device from her bedroom:

> I'm really concerned about privacy . . . I remember at first when I put it in my bedroom, and we talked about my son whose name is Max. I don't know what the similarity was, maybe Alexa and Max. And it started to work and joined the conversation. So it made me mad.

From listening to seeing, newest VBAs are perceived as creepy and invasive. In each session, participants viewed an Amazon-produced Echo Show commercial and discussed their reactions. In several sessions, a related product (the Echo Look) came up because it shared camera features with the Show. While some participants noted the benefits of more advanced VBAs (e.g., Huong described the convenience of having a screen so she can see how much time is left after setting a timer), the word "creepy" emerged repeatedly, without prompting, by users and nonusers in nearly all focus groups.

The main Echo Show feature that provoked strong responses from participants was the "Drop-In" feature, which Amazon describes as a "two-way intercom." For this feature to work, users create a list of approved contacts they can connect with. Once a contact approves this privilege, they can instantly connect via audio (on Echo devices) or video (on the Echo Show). One participant, Sun-Joo, shared her experiences trying out Drop-In on her Echo Dot, describing tensions between feeling connected to her family and being *too* connected:

> I don't need them to call me every minute of the day. If it tells them I'm active, they know I'm home, so if I don't answer, I get a text message, "Hey, where are you? I just tried to call you." . . . I'm trying to find a balance.

No other participants had direct experience with the feature.

Immediate reactions after watching the commercial reflected wariness toward features like Drop-In, with participants describing them as creepy and invasive. For example, Liz said, "I'm the kind of person that has a piece of tape over my computer camera because I don't trust it. So the Drop-In thing, that's creepy." Likewise, Walter described the stress of having to be more aware of what he did in private spaces. Speaking about the Echo Look—described as a "Hands-Free Camera and Style Assistant with Alexa"—he asked, "What happens when you come out of the shower and it takes a picture of your body and tells you you need to diet, you need to exercise more?" Multiple participants expressed concerns about Echo devices equipped with cameras, especially since the Look is marketed for bedroom use (to provide feedback on outfits). Olivia said:

> I feel a little uncomfortable with the idea of a camera that could always be on because they always say cover your laptop camera . . . if you had something that had a camera that was looking into your bedroom or an intimate space, I feel like that's really creepy. If somebody were to hack that or hack a Drop-In and just like, actively watch you . . . I don't like that.

While participants' initial reaction to the Echo Show captured its general "creepiness," their comments also reflected feelings of weariness toward and being overwhelmed by more invasive technologies that collected more data, both in terms of quantity and quality. These devices led them to think about more things that could go wrong (e.g., camera positioning, being careful about what you say near the device)—such as when Huong said, "I don't have a problem with pointing cameras outside, but I'm not too comfortable with the cameras inside always on"—or to voice displeasure with technology making them always accessible, as when Sun-Joo described her experience with the Drop-In feature (detailed above).

Managing devices could also get overwhelming, as when John described conflicting feelings about his devices:

> There are times when I very much love having everything connected and hooked up. But then, after awhile, it just gets a little bit where I'm like this is too much. And trying to find that balance is definitely an interesting tightrope to walk because I definitely see the advantages and benefits of it, but at the same time, I'm like, you know, is it too much?

Moving beyond VBAs to consider the wider ecosystem of smart devices in homes—as well as improvements in machine learning that enable devices to make better predictions—these themes of wariness and weariness were exacerbated further. Some participants expressed discomfort with widespread data collection and sharing between companies, while others expressed concerns related to the increasing reliance on technology to accomplish basic tasks. For example, Rebecca asked, "Where do we draw the line? To the point where we're 100% dependent upon devices doing certain things for us?" Zack also pushed back against extreme customization, sharing how he tried to sabotage the underlying algorithm in his VBA: "I've been trying to feed it specific information and it fails in so many ways to get any type of personalized response."

Discussion

In this study, we have explored the role that privacy considerations play in (non-)use of voice-based assistants (VBAs), as well as how privacy concerns are shifting as smart technologies add new features, collect more data, and become better equipped to make inferences and recommendations based on user data. VBAs help us better understand human-machine communication, as users vocally interact with smart speakers to accomplish a variety of tasks (Guzman, 2020). Researchers have described VBAs, and the smart speakers that house them, as hybrids between humans and machines (Weidmüller, 2022) and have found that users attribute human-like characteristics to them (Etzrodt & Engesser, 2021; Garcia et al., 2018; Guzman, 2020). VBAs also provide an important case study for evaluating privacy risks of broader IoT ecosystems because of how they are perceived by users, where they are used (private spaces), the types of data they collect (audio/video), and their function as a hub for a range of smart home devices.

CPM (Petronio, 2002; Petronio et al., 2021) provides a useful framework for considering how people balance the benefits and risks of technologies like VBAs. Recent studies have extended this theory—which was developed to address interpersonal relationships—to human-machine interactions (e.g., Kang & Oh, 2021; Xu et al., 2022). In this paper, we build on these studies to consider how both users and nonusers rationalize decisions related to these devices, using data from focus groups to unpack the complex set of factors that influence these decisions.

CPM is guided by a set of assumptions that helps explain why so many users we spoke to expressed cynicism and apathy toward data privacy. In interpersonal relationships, people negotiate rules related to ownership and control of personal information—and re-negotiate those rules when they experience privacy breakdowns (Petronio et al., 2021). One's relationship with a VBA—and by extension the company that manages that VBA—is much more one-sided, with users often having to agree to certain rules and restrictions via terms of use. It is unsurprising, then, that participants felt less agency and described their data as already being "out there" when news stories regularly highlight data breaches, scandals, and other uses of their data that go beyond expectations (Sheshadri et al., 2017).

CPM helps us move beyond simple explanations that people "just don't care" about their privacy anymore, a sentiment suggested in many studies of technology use. For example, work evaluating privacy attitudes toward IoT found that perceived benefits and organizational trust positively influenced willingness to share personal information, but perceived risks and information sensitivity had no effect on use (Kim et al., 2019). The authors suggest consumers place higher value on the benefits of these technologies and "do not pay much attention" (p. 278) to IoT-based privacy risks. Such an interpretation may apply to active VBA users, but it does not address the privacy concerns raised by nonusers—many of whom noted their concerns were a major factor in the decision to not use VBA devices. For nonusers, organizational trust may be lower—a factor prior work has associated with VBA nonusers (Lau et al., 2018)—and their desire for data control likely supersedes perceived benefits when making purchasing decisions.

More than a decade ago, boyd (2010) noted that networked publics like social media were blurring boundaries between public and private spaces. We argue that smart devices further complicate this blurring due to their widespread popularity, the passive nature of most data collection, and the limited ability to view and edit that data. This limited access to data makes it exceedingly challenging to identify rule violations. Rather, users must trust companies are abiding by the rules they've laid out in their terms of use; even when a rule is violated, there is often little recourse outside of unplugging or removing the device.

CPM focuses on boundary regulation—individuals negotiate how thick or thin a boundary should be for a given piece of private information (Petronio et al., 2021). More sensitive information tends to have thicker boundaries to better protect it from undesired access, while thinner boundaries enable easier flow of information. Our contemporary technological ecosystem increasingly relies on thin boundaries to facilitate the flow of data from individuals to other people (e.g., through social media posts) and companies (e.g., through automated data collection). Researchers have sometimes framed this focus on increasing boundary permeability in terms of "information friction" (e.g., Floridi, 2005), which describes the amount of work required for an entity to access another's information. VBAs provide an example of how this concept works in practice: by default, devices are always listening for a "wake word"—this reduces friction for a person interacting with the device, but increases risks related to inappropriate or unintended data flows. To increase friction, one could use the mute button; however, by removing hands-free interaction, a primary benefit of smart devices is lost. Friction can also be embodied in privacy settings and features that help verify users' intentional interaction with a smart device, and can be useful in verifying things are operating as they should; however, research suggests that users rarely employ available privacy settings (Lau et al., 2018; Malkin et al., 2019).

Individuals' ability to engage in boundary regulation is further challenged when the "rules" constantly change, as can be seen in both updates to terms of use and in the frequent release of new or updated technologies. Participants discussed concerns about feature creep—the ongoing expansion of device features that facilitate additional data collection and monitoring (Surowiecki, 2007). Participants worried that devices were always listening, and their concerns increased for newer devices with cameras. Concerns also emerged from uncertainty around *when* devices collected data and *what* happened to collected data. In 2019, Amazon responded to these concerns by adding new features to allow users to repeat their last command and to explain why it made a recommendation (Ellis, 2019); however, such features focus on transparency rather than providing opportunities to regulate boundaries and control information flows.

What would boundary regulation of VBA data look like? One example is IoT Inspector (Huang et al., 2020), which allows users to capture, analyze, and visualize network traffic generated within their smart homes. Researchers continue to develop ways to increase users' awareness of data flows generated from smart devices (see, for example, Thakkar et al., 2022)—something many of our participants expressed a desire to see in new devices. We hope that future work continues to explore options for helping users negotiate their interactions with machines as they are increasingly confronted with challenges to privacy-based decision-making.

Conclusion

Research suggests IoT technologies will continue to expand over the next decade, as will the push toward creating smart home ecosystems that provide instant access to and control over one's home environment. With such expansion comes greater privacy risks, and we must continue to evaluate how users assess and respond to these risks. By extending CPM to human-machine interactions, we can further explore how communication behaviors—including the disclosure of private information—are shaped by the features and affordances of these technologies.

Such evaluations can also inform future designs of sociotechnical systems to empower users through flexible and intuitive interfaces that provide more transparency about what data is collected and more control over how data is used. Given that our participants expressed concerns regarding AI and devices that collect a greater variety and quantity of data, it becomes even more important to provide users with ways to increase friction and restrict data flows. Skeba and Baumer (2020) provide a useful initial consideration of how the use of AI, algorithms, and big data reduce friction and impact privacy, but more research is needed in this space. Finally, future research in this space must consider how to effectively communicate information about data collection and use so people can make fully informed decisions before sharing their data.

Author Biographies

Jessica Vitak, PhD, is an associate professor in the College of Information Studies and director of the Human-Computer Interaction Lab (HCIL) at the University of Maryland. Her research evaluates the privacy and ethical implications of big data, the internet of things, and other "smart" technologies. She seeks to understand how privacy concerns play a role in technology adoption and use, and she develops tools and resources to help children and adults make more informed decisions when using technology and sharing sensitive data.

http://orcid.org/0000-0001-9362-9032

Priya Kumar, PhD, is an assistant professor at Pennsylvania State University's College of Information Sciences and Technology. Her research on the datafication of family life aims to shift digital technology discourse and design away from a focus on individual control and toward more networked understandings of privacy and agency. For more information, visit https://priyakumar.org/.

http://orcid.org/0000-0001-9244-7915

Yuting Liao, PhD, is a senior UX researcher at Intuit. Her research focuses on trust, privacy, and data ethics to guide technology design and build harmonious sociotechnical interactions in various contexts, including conversational AI, Fintech, and health technology.

http://orcid.org/0000-0002-5008-2097

Michael Zimmer, PhD, is a privacy and data ethics scholar whose work focuses on digital privacy and surveillance, the ethics of big data, internet research ethics, and the broader social and ethical dimensions of emerging digital technologies. Dr. Zimmer is an Associate Professor in the Department of Computer Science at Marquette University in Milwaukee, Wisconsin, where he also serves as Director of Marquette's Center for Data, Ethics, and Society.

http://orcid.org/0000-0003-4229-4847

References

Alhammadi, A., AlZaabi, A., AlMarzooqi, B., AlNeyadi, S., AlHashmi, Z., & Shatnawi, M. (2019). Survey of IoT-based smart home approaches. *2019 Advances in Science and Engineering Technology International Conferences* (pp. 1–6). IEEE. https://doi.org/10.1109/ICASET.2019.8714572

Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, crowding*. Brooks/Cole Publishing Company.

Ammari, T., Kaye, J., Tsai, J. Y., & Bentley, F. (2019). Music, search, and IoT: How people (really) use voice assistants. *ACM Transactions on Computer-Human Interaction*, *26*(3), 1–28. https://doi.org/10.1145/3311956

Boeckl, K., Fagan, M., Fisher, W., Lefkovitz, N., Megas, K., Nadeau, E., Piccarreta, B., Gabel O'Rourke, D., & Scarfone, K. (2019). *Considerations for managing internet of things (IoT) cybersecurity and privacy risks* (NIST Internal or Interagency Report 8228). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8228 boyd, danah. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In Z. Papacharissi (Ed.), *A networked self* (pp. 39–58). Routledge. https://doi.org/10.4324/9780203876527-8

Chokshi, N. (2018, March 8). Amazon knows why Alexa was laughing at its customers. *The New York Times*. https://web.archive.org/web/20180309093135/https://www.nytimes.com/2018/03/08/business/alexa-laugh-amazon-echo.html

Culnan, M. J. (1993). "How did they get my name?": An exploratory investigation of consumer attitudes toward secondary information use. *MIS Quarterly*, *17*(3), 341–363. https://doi.org/10.2307/249775

Day, M., Turner, G., & Drozdiak, N. (2019, April 24). Amazon's Alexa team can access users' home addresses. *Bloomberg*. https://web.archive.org/web/20190426112016/https://www.bloomberg.com/news/articles/2019-04-24/amazon-s-alexa-reviewers-can-access-customers-home-addresses

Díaz, Á. (2020). *When police surveillance meets the 'internet of things'*. Brennan Center for Justice. https://web.archive.org/web/20201218170200/https://www.brennancenter.org/our-work/research-reports/when-police-surveillance-meets-internet-things

Easwara Moorthy, A., & Vu, K.-P. L. (2015). Privacy concerns for use of voice activated personal assistant in the public space. *International Journal of Human-Computer Interaction*, *31*(4), 307–335. https://doi.org/10.1080/10447318.2014.986642

Edison Research. (2022, March 23). The infinite dial 2022. https://web.archive.org/web/20220325084543/https://www.edisonresearch.com/the-infinite-dial-202

Ellis, C. (2019, September 25). Amazon Alexa gets new privacy controls, and will tell you what it's hearing. *TechRadar*. https://web.archive.org/web/20190928170929/https://www.techradar.com/news/amazon-alexa-gets-new-privacy-controls-and-will-tell-you-what-its-hearing

Etzrodt, K., & Engesser, S. (2021). Voice-based agents as personified things: Assimilation and accommodation as equilibration of doubt. *Human-Machine Communication*, *2*, 57–79. https://doi.org/10.30658/hmc.2.3

Floridi, L. (2005). The ontological interpretation of informational privacy. *Ethics and Information Technology*, *7*(4), 185–200. https://doi.org/10.1007/s10676-006-0001-7

Fowler, G. A. (2018, May 24). Hey Alexa, come clean about how much you're really recording us. *Washington Post*. https://web.archive.org/web/20201114141514/https://www.washingtonpost.com/news/the-switch/wp/2018/05/24/hey-alexa-come-clean-about-how-much-youre-really-recording-us/

Garcia, D. M. P., Lopez, S. S., & Donis, H. (2018). Voice activated virtual assistants personality perceptions and desires: Comparing personality evaluation frameworks. *Proceedings of British HCI 2018* (pp. 1–10). BCS Learning and Development Ltd. https://doi.org/10.14236/ewic/HCI2018.40

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior*, *90*, 343–350. https://doi.org/10.1016/j.chb.2018.08.009

Guzman, A. L. (2020). Ontological boundaries between humans and computers and the implications for human-machine communication. *Human-Machine Communication*, *1*, 37–54. https://doi.org/10.30658/hmc.1.3

Huang, D. Y., Apthorpe, N., Li, F., Acar, G., & Feamster, N. (2020). IoT Inspector: Crowdsourcing labeled network traffic from smart home devices at scale. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(2), Article 46. https://doi.org/10.1145/3397333

Kang, H., & Oh, J. (2021). Communication privacy management for smart speaker use: Integrating the role of privacy self-efficacy and the multidimensional view. *New Media & Society*, 146144482110266. https://doi.org/10.1177/14614448211026611

Kim, D., Park, K., Park, Y., & Ahn, J.-H. (2019). Willingness to provide personal information: Perspective of privacy calculus in IoT services. *Computers in Human Behavior*, *92*, 273–281. https://doi.org/10.1016/j.chb.2018.11.022

Krueger, R. A. (2014). *Focus groups: A practical guide for applied research*. SAGE.

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), Article 102. https://doi.org/10.1145/3274371

Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues*, *33*(3), 22–42. https://doi.org/10.1111/j.1540-4560.1977.tb01880.x

Levy, K., & Schneier, B. (2020). Privacy threats in intimate relationships. *Journal of Cybersecurity*, *6*(1), tyaa006. https://doi.org/10.1093/cybsec/tyaa006

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* (pp. 102–113). Springer International Publishing. https://doi.org/10.1007/978-3-030-15742-5_9

Lutz, C., & Newlands, G. (2021). Privacy and smart speakers: A multi-dimensional approach. *The Information Society*, *37*(3), 147–162. https://doi.org/10.1080/01972243.2021.1897914

Lutz, C., & Tamó-Larrieux, A. (2020). The robot privacy paradox: Understanding how privacy concerns shape intentions to use social robots. *Human-Machine Communication*, *1*, 87–111. https://doi.org/10.30658/hmc.1.6

Lynskey, D. (2019, October 9). "Alexa, are you invading my privacy?" The dark side of our voice assistants. *The Guardian*. https://web.archive.org/web/20191010025233/https://www.theguardian.com/technology/2019/oct/09/alexa-are-you-invading-my-privacy-the-dark-side-of-our-voice-assistants

Madden, M., & Rainie, L. (2015). *Americans' attitudes about privacy, security and surveillance*. Pew Research Center. https://web.archive.org/web/20191104064056/https://www.pewresearch.org/internet/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

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

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

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford Law Books.

Patton, M. Q. (2002). *Qualitative research & evaluation methods*. SAGE.

Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure*. SUNY Press.

Petronio, S., Child, J. T., & Hall, R. D. (2021). Communication privacy management theory: Significance for interpersonal communication. In D. O. Braithwaite & P. Schrodt (Eds.), *Engaging theories in interpersonal communication* (3rd ed., pp. 314–327). Routledge. https://doi.org/10.4324/9781003195511

Pradhan, A., Mehta, K., & Findlater, L. (2018). "Accessibility came by accident": Use of voice-controlled intelligent personal assistants by people with disabilities. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Article 459). New York: ACM. https://doi.org/10.1145/3173574.3174033

Sheshadri, K., Ajmeri, N., & Staddon, J. (2017). No (privacy) news is good news: An analysis of *New York Times* and *Guardian* privacy news from 2010–2016. *Proceedings of the 15th Annual Conference on Privacy, Security and Trust* (pp. 159–168). IEEE. https://doi.org/10.1109/PST.2017.00027

Shklovski, I., Mainwaring, S. D., Skúladóttir, H. H., & Borgthorsson, H. (2014). Leakiness and creepiness in app space: Perceptions of privacy and mobile app use. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2347–2356). New York: ACM. https://doi.org/10.1145/2556288.2557421

Skeba, P., & Baumer, E. P. S. (2020). Informational friction as a lens for studying algorithmic aspects of privacy. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), Article 101. https://doi.org/10.1145/3415172

Straus, R. A. (2019). *Mastering focus groups and depth interviews: A practitioner's guide*. Paramount Market Publishing.

Surowiecki, J. (2007, May 21). Feature presentation. *The New Yorker*. https://web.archive.org/web/20141003144234/https://www.newyorker.com/magazine/2007/05/28/feature-presentation

Thakkar, P. K., He, S., Xu, S., Huang, D. Y., & Yao, Y. (2022). "It would probably turn into a social faux-pas": Users' and bystanders' preferences of privacy awareness mechanisms in smart homes. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (Article 404). New York: ACM. https://doi.org/10.1145/3491102.3502137

Weidmüller, L. (2022). Human, hybrid, or machine? Exploring the trustworthiness of voice-based assistants. *Human-Machine Communication*, *4*, 85–110. https://doi.org/10.30658/hmc.4.5

Xu, K., Chan-Olmsted, S., & Liu, F. (2022). Smart speakers require smart management: Two routes from user gratifications to privacy settings. *International Journal of Communication*, *16*(0). https://ijoc.org/index.php/ijoc/article/view/17823

Zeng, E., Mare, S., & Roesner, F. (2017). End user security and privacy concerns with smart homes. *Proceedings of the Thirteenth USENIX Conference on Usable Privacy and Security* (pp. 65–80). USENIX Association. https://dl.acm.org/doi/10.5555/3235924.3235931

Zheng, S., Apthorpe, N., Chetty, M., & Feamster, N. (2018). User perceptions of smart home IoT privacy. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), Article 200. https://doi.org/10.1145/3274469

HMC

HUMAN-MACHINE

COMMUNICATION