"Nobody's Happy": Design Insights from Privacy-Conscious Smart Home Power Users on Enhancing Data Transparency, Visibility, and Control

Sunyup Park, *University of Maryland, College Park*
Anna Lenhart, *University of Maryland, College Park*
Michael Zimmer, *Marquette University*
Jessica Vitak, *University of Maryland, College Park*

Abstract

As smart home technologies continue to grow in popularity and diversity, they raise important questions regarding ways to increase awareness about data collection practices and empower users to better manage data flows. In this paper, we share insights from 32 privacy-conscious smart home power users—individuals who have invested significant time, money, and technological prowess in customizing their smart home setup to maximize utility and meet privacy and security needs. We explore the drawbacks and limitations power users experience when balancing privacy goals with interoperability, customizability, and usability considerations, and we detail their design ideas to enhance and extend data transparency, visibility, and control. We conclude by discussing the importance of designing smart home technologies that both address these considerations and empower a wide range of users to make more informed decisions about whether and how to implement smart technologies in their homes, as well as the wider need for greater regulation of technologies that collect significant user data.

1. Introduction

Smart home devices (SHDs) have gained popularity in recent years, offering a convenient and efficient way to control and monitor various aspects of one's home remotely. SHDs range from smart speakers and thermostats to security systems and lighting, and they can be integrated with other smart devices, hubs, and apps to create a fully automated smart home environment.

While SHDs offer significant benefits—ranging from convenience and cost efficiency to added security and accessibility—they have also led to increased data privacy risks. In particular, researchers and privacy advocates have raised questions regarding the vast amounts of data devices collect about users and their environments, often without their full knowledge or consent [2,54]. This data includes information about household members' activities, movements, habits, and preferences, which can potentially be misused by hackers, manufacturers, government agencies, or others [31].

Ensuring data privacy and security requires continuous vigilance from consumers to be aware of the data flows across and between devices and the privacy policies and practices of the companies they purchase SHDs from, and to take necessary steps to secure smart home environments. Beyond these factors, research highlights that creating holistic approaches to protecting the privacy of smart home environments that address the different platforms, end-users, and data flows is very challenging [9].

In seeking to address common privacy concerns, prior research has largely evaluated the privacy attitudes and behaviors of "average" or "everyday" smart home users. However, power users—those who "use the devices more innovatively, efficiently, and thoroughly than ordinary users" [57] (p. 1743)—engage in a wider range of practices to maximize the utility of SHDs and implement strategies to track and manage data flows beyond what devices natively provide. Because this population is heavily engaged in researching device options and spending time and energy optimizing setup to balance functionality and privacy, "privacy-conscious power users (PCPUs)" are uniquely positioned to provide feedback and insights that everyday users may not consider.

In this paper, we share insights from focus groups with 32 privacy-conscious smart home power users to better understand the limitations of current smart home options and identify key areas for improving data access and control. By evaluating the limitations of current technologies and eliciting their design ideas for improving or enhancing data management and control, our focus on PCPUs provides a unique perspective on how to better design smart home technologies to match the needs of a full range of users. Thus, we focus our analysis on two research questions:

RQ1: What drawbacks and limitations do smart home power users identify in their current smart home setup?

RQ2: What design features do smart home power users want to see developed or expanded in future tools and platforms to enhance data management and control?

USENIX Symposium on Usable Privacy and Security (SOUPS) 2023. August 6 -- 8, 2023, Anaheim, CA, USA.

.

Our findings highlight major tensions between the privacy and security goals our participants have and the customizability and interoperability of devices and hubs. Even though our participants spent significant time, energy, and money on their smart home setups, they found themselves sometimes limited in their options and having to make tradeoffs between privacy and functionality. Participants also described challenges managing devices with multiple users, whether other household members or visitors.

To address these challenges, our participants suggested three core areas for design improvements: data transparency, visibility, and control. Importantly, these power users stressed the need for interfaces and device management controls that accommodate a range of skill levels and privacy preferences, recognizing that most users lack the technical skills or desire to use advanced network management features or customizations to protect their privacy. At the same time, our participants also wanted advanced features, customizations, and data visualization options to accommodate their goals as power users. We conclude by reflecting on the barriers to accommodating these ideas and building tools that can provide control and privacy enhancement for a wide range of user types, and we consider how emerging standards and legislation might help in promoting data rights.

2. Related Work

2.1. Privacy Risks with Smart Home Technology

There are three major categories of privacy risks associated with SHDs. First, we consider the *location* of devices within the home. Consumers have significant concerns about data being collected in their homes [5,12,17,28,35] and find it more sensitive than data collected in public spaces [17]. Privacy concerns vary by room, with bedrooms [10,12], bathrooms [10,29,35], and children's spaces [51] being among the most concerning.

Second, people's concerns vary based on the *type* of data being collected. For example, audio data is frequently captured from smart speakers and TVs. Dunbar et al. [14] describe three categories of attitudes toward audio data collection: pragmatists, who have few concerns, are willing to trade privacy for benefits, and generally trust companies; guardians, who attempt to minimize data collection and safeguard data that is collected; and cynics, who rarely change default settings and lack a clear understanding of when data is collected and how it is processed. Video data, such as that collected from smart cameras in and around the home, also raises concerns [10,17,51], while research suggests smart home users have few concerns about raw data from more simple sensors (e.g., temperature, light) [25].

Third, researchers have found that *who* is collecting data is important to consumers [5,27,28]. Many consumers consider the reputation of technology companies when making decisions to purchase IoT devices [32,54], and users generally express trust that device manufacturers will collect data for legitimate purposes [26,56].

2.2. Privacy Design Work on IoT & Smart Homes

Usable privacy and security researchers have developed and evaluated various mechanisms to mitigate users' privacy concerns associated with IoT and SHDs. Perhaps the best-known design is privacy nutrition labels, which seek to increase transparency and support informed decision-making through standardized information about data collection and use. This work was initially carried out by Kelley et al. [22], before also being implemented by Apple in 2020 to provide standardized privacy labels for apps in the AppStore [39]. Emami-Naeini et al. [16] extended these labels to IoT devices, providing two layers of information to consumers: a more general label on the device packaging that contains information about security mechanisms, and a second label (accessible online) containing more detailed information about data collection and use.

The deployment of privacy nutrition labels has, however, hit roadblocks. There remains little incentive for developers of older apps to create or update privacy labels, and privacy labels themselves seem to be rarely updated once created [30]. Furthermore, even when developers think privacy labels are a positive thing, they are faced with challenges in creating them because of misunderstandings and the overall complexity of the task [30].

The IoT privacy and security community has also created tools aimed at increasing the visibility of data flows and provide users with additional privacy controls [47]. Tools such as IoTSense [7] and IoTSentinel [36] identify devices by analyzing network traffic, while IoT Inspector provides visualizations of device activity and traffic destinations [20]. Others have designed tools to increase the legibility of information flows and improve interpretability by providing users with more details about their connected devices and providing actionable choices [43,47]. Zeng and Roesner [55] designed a tool that includes location-based access controls as well as supervisory access controls to allow multiple users within a smart home to control their own privacy settings.

Information about data practices and data flows can be complemented with privacy notices as a part of privacy awareness mechanisms in smart homes. Privacy notices in the smart home context heavily focus on the use of visual and audio indicators on SHDs. For example, Song et al. [49] found that users were most interested in knowing the physical location of cameras and voice assistants, and preferred locator mechanisms that integrated visual and audio cues as the number of devices increased. In fact, many smart speakers now use lights to communicate the device's status to

users. However, Thakkar et al. [52] note that privacy notices in smart homes have a long way to go; current notices focus on individual stakeholders and devices, whereas the future of smart homes will inevitably consist of multiple stakeholders and devices.

2.3. Designing Smart Home Privacy Tools for Diverse Users

Researchers have also stressed the need to account for a diverse user base when designing SHDs and controls. For example, Chhetri and Motti [11] note that most SHDs lack user-friendly privacy controls, and they develop a framework to guide developers in building user-friendly privacy controls. In a similar vein, Kim et al. [23] created personas to help designers better understand potential users vulnerable to cybersecurity risks. Researchers have also identified several categories of design features to mitigate privacy harms, including transparency, privacy and security controls, and assistance for users [19,53]. In summary, features and diagnostic tools should be simple, proactive, preventative, and provide users with transparency and control [6,21].

Researchers have also considered the privacy needs of bystanders, including non-primary users as well as those whose data is captured from incidental interaction with devices. Yao and colleagues [53] found that bystanders' privacy concerns are more contextually dependent than users; bystanders wanted to cooperate with smart-home owners to negotiate privacy needs. Ahmad et al. [3] argue that devices should be designed to afford users and bystanders "tangible privacy" through features like camera covers, physical on/off buttons, and clear on/off indicators. On the other hand, Thakkar et al. [52] found that bystanders might not know or neglect users' privacy concerns when privacy awareness mechanisms are implicated; they suggest having a separate bystander mode within device controls. Similarly, Marky et al. [34] found that visitors' lack of awareness limits their ability to protect their privacy.

As a whole, research to date has identified several privacy risks associated with data collected by SHDs and frameworks for design tools to mitigate those risks. The present study extends this prior work by providing insights from smart home power users, who are particularly attuned to building systems that offer flexibility and customization without sacrificing data privacy.

3. Method

While prior research has evaluated users' privacy needs in smart homes, this paper considers a distinct customer segment: privacy-conscious power users (PCPUs). We argue that these users' perspectives can be especially useful when considering how to better design SHDs and features to achieve the privacy needs described above. Power users are enthusiastic about devices and are willing to put in the time and effort to find solutions to mitigate their privacy concerns [33,57]. PCPUs likely have more experience trying out a range of devices and solutions to minimize data collection and maximize their ability to monitor and control data flows. They may also have insights into how these devices can better serve non-technical users, as they may have experience customizing their homes to accommodate non-primary users and bystanders.

This paper presents data from 10 focus groups with 32 privacy-conscious smart home power users. Focus groups are especially useful for developing a deeper understanding of how people with a shared experience feel about an issue [24]. Additionally, they enable a variety of perspectives and immediate follow-up from other participants and facilitators, which can be helpful for design-based inquiries and idea generation [24]. In our case, we used focus groups to bring together smart home enthusiasts who engaged in various approaches to managing their devices to understand their perspectives on the drawbacks of existing devices and interfaces. The group discussions allowed us to also solicit input into how to improve current and future smart home technology to accommodate diverse needs and provide users with greater awareness and management of data flows.

3.1. Recruitment and Study Design

This paper is part of a larger research project evaluating the privacy concerns and practices of smart home users. In summer 2021, we recruited people to participate in virtual focus groups to discuss their use of smart home technologies. After receiving IRB approval from the University of Maryland, we began posting recruitment messages on social media, including Twitter and smart home-related subreddits and Facebook Groups, inviting people who want to "*talk about how they use devices in their homes, the types of data these devices collect and share, and how we can design tools to better visualize this data and provide consumers with more control over their data*." The message directed potential participants to a short survey that collected demographics, details about their home environment, general privacy attitudes, and SHDs used.

We received 441 responses over one week; after removing spam responses, we had 277 potential focus group participants. We used two types of purposeful sampling—criterion and maximum variance [42]—to create a prioritized participant pool based on three factors. First, we looked at the devices respondents said they used. Given our interest in more advanced users, we removed from consideration anyone who selected a single type of device and prioritized those who used several different types of devices. Second, we looked at various items in the survey that would suggest a person was privacy conscious. This included attitudes toward privacy as well as managing devices to address privacy concerns. We prioritized people who reported engaging in privacy-enhancing device management (e.g., moving

Table 1: Participant IDs, Descriptive Data, and Smart Home Details

ID	Gender	Race	Age	Devices Used[1]	Integration Platform[2]	Advanced Network Management?[3]
P1	M	White	37	1,5,6,8,11	HK	No
P2	M	Black	37	1,2,3,5,6,7,8,11	HA, HK	Yes
P3	F	Black	45	1,2,4,5,6,8,10,11,12	HK	No
P4	M	White	48	1,8	none	Yes
P5	M	White	52	1,2,3,4,5,6,7,8,10,11	HK	No
P6	M	White	52	1,6,11,12	ST	No
P7	F	White	37	1,2,3,6,11	HK	No
P8	M	White	38	1,2,4,5,6,10,11	none	No
P9	M	White	35	1,2,3,4,5,6,7,8,10,11	HK	Yes
P10	n/a	n/a	39	1,2,3,4,5,6,7,10,11	HK	No
P11	M	White	34	2,3,4,6,7,11,12	HA	Yes
P12	M	White	27	5,6,8	HK	Yes
P13	M	East Asian	55	1,2,5,6,10,11	HK+HB	No
P14	M	White	40	1,6,8,11	HA	Yes
P15	M	White	33	1,3,4,5,6,7,8,10,11,12	HA	Yes
P16	M	E Asian, White	38	1,2,3,4,5,6,8,10,11,12	HK+HB	Yes
P17	M	White	24	1,2,6,8,11	none	No
P18	M	White	47	1,5,6,10,11,12	ST	Yes
P19	M	White	20	1,5,6,8,11,12	HK	Yes
P20	M	White	39	1,2,3,4,5,6,8,9,10,11,12	ST, HA	Yes
P21	F	White	29	1,2,6,8,10,11,12	HK, HA	No
P22	M	White	32	1,4,6,8,10,11,12	HK, HA	No
P23	NB	East Asian	25	1,6	none	No
P24	M	White	32	1,3,5,6,7,8,10,11	HK+HB	Yes
P25	M	White	40	1,2,4,6,8,10,11	HK+HB	Yes
P26	M	White	28	1,2,4,5,6,8,10,11,12	HK	No
P27	M	American Indian	36	1,2,5,6,8,10,11,12	HK	Yes
P28	M	E&S Asian	20	1,3,4,6,8	HK	No
P29	F	White	30	1,2,3,6,7,8,10,11	none	Yes
P30	F	White	42	1,2,3,4,6,10,11	HA, HK	Yes
P31	M	White	23	1,2,6,8	HK+HB	No
P32	M	E Asian, White	47	1,4,5,6,8,11	HK+HB	No

[1] Smart devices participants used: 1) speaker; 2) thermostat; 3) vacuum; 4) doorbell; 5) security camera; 6) lighting; 7) blinds; 8) TV 9) refrigerator; 10) door locks; 11) sensors; 12) other.

[2] Smart home hubs participants used: HK (Apple HomeKit), HB (Homebridge), HA (Home Assistant), ST (Samsung SmartThings).

[3] "Yes" is assigned to participants who used 1+ advanced network management strategies (i.e., setting up a Pi-hole or private DNS, flashing devices with custom firmware to run locally, setting up multiple routers to isolate devices, setting up firewalls).

devices out of private spaces, not using certain brands). Based on this, we had an initial set of 129 people we wanted to contact. We then applied the third criterion, which was to maximize diversity across gender, race, and home environment (e.g., own vs. rent; live alone vs. with others). This led to us prioritizing non-male and non-white respondents, who were under-represented in the pool.

Using our prioritized list, we began inviting people in batches to participate in 60-minute Zoom-based virtual focus groups in August 2021. We kept sessions small (3-4 participants) to ensure everyone had ample time to speak and to address the challenge of deciphering nonverbal communication virtually [50]. At least two authors attended each session. The research team debriefed after sessions and discussed if we were hearing new ideas and determining when we had reached data saturation [45]. In total, 82 people were contacted, 38 people signed up for a focus group, and 32 people participated in one of 10 sessions (see Table 1). Participants were compensated with a US$30 gift card.

Each session started with participants sharing general thoughts about their devices, how they built out their home environment, challenges and drawbacks they experienced, and concerns they had about devices. Participants then completed two brainstorming activities using Google Jamboard. We first asked them to map the types of data their devices collected onto a grid that captured the perceived sensitivity of that data along one axis and their desire to control and see data flows on the other. Following a discussion, we then asked them to brainstorm ideas and add post-it notes regarding the features they thought would be useful in visualizing or sharing data from their SHDs (see Figure 1; in this session, a team member organized post-its into clusters as participants added them). All participants described their Jamboards and the transcripts were analyzed. Due to time constrains, some sessions skipped the post-it note portion of the brainstorm session, moving straight to discussion. See appendix for full protocol.

3.2. Data Analysis

Audio from sessions was transcribed via Rev, then uploaded to Atlas.ti for qualitative coding. Using Miles, Huberman, and Saldaña's [38] approach to guide our analysis, we conducted two cycles of coding. We first developed an initial codebook based on the focus group protocol, the detailed notes taken during each session and our research questions. Each team member coded one transcript using the initial codebook, adding memos with questions and suggestions for new or collapsed codes. The team then discussed this initial process and refined the codebook. Following this, each transcript was then coded by two authors to ensure all relevant codes were applied.

Coded excerpts were then exported to Excel for secondary coding, following Braun and Clarke's [8] approach for thematic analysis, as well as Saldaña's [44] technique for "themeing the data." For each code, one team member reviewed all coded excerpts, taking notes on emergent patterns in the data. Through multiple rounds of reading and taking notes, team members began categorizing excerpts from each code and extracting themes, then writing a detailed analytic memo to describe each theme and provide examples [38]. These memos were discussed by the full team before organizing them into findings.

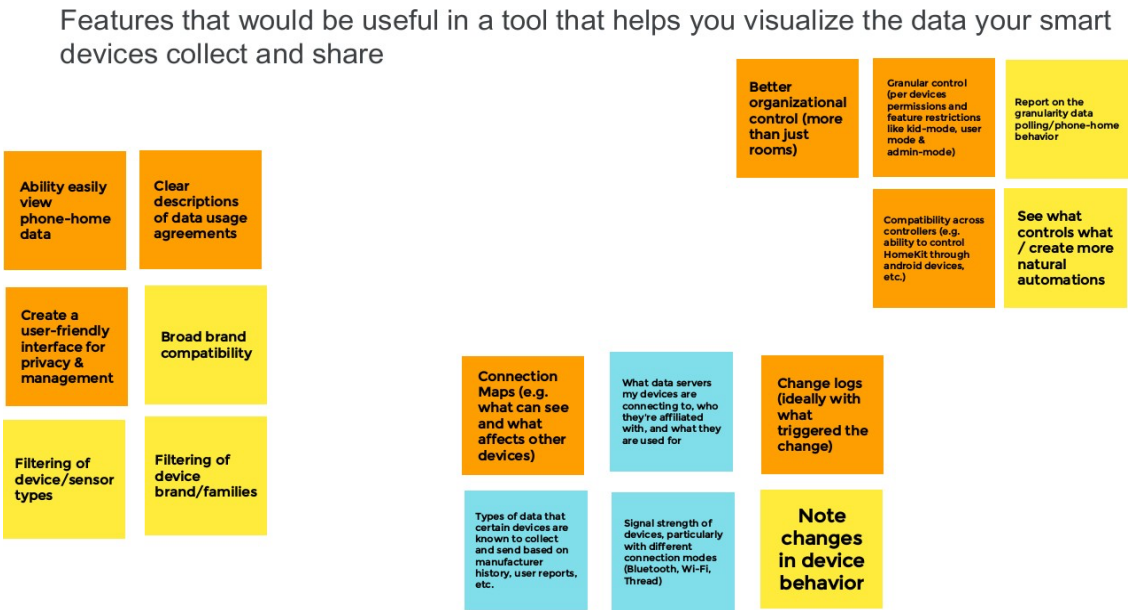

Figure 1. Jamboard screenshot from a focus group session, second design activity (brainstorming design features).

Table 2: Number of Participants Using SHDs/Hubs

Smart Home Device Usage			
Lighting	31	Speaker	30
Sensors	27	TV	23
Thermostat	19	Door locks	18
Security camera	18	Doorbell	15
Vacuum	13	Other	12
Blinds	8	Refrigerator	1
Smart Home Hub/Integration Platform Usage			
HomeKit	21	Home Assistant	8
Homebridge	6	Smart Things	3

In this paper, we focused primarily on two codes: Drawbacks (RQ1) and Design Features (RQ2). Additional codes were explored to supplement our analysis, including Bystanders and Data Concerns. We also conducted an additional round of analysis on the final brainstorming activity and Jamboards to further delineate the range of design features participants discussed into transparency, visibility, control and layers, taking note of why those features are important and the drawbacks they address.

4. Findings

4.1. Drawbacks and limitations to current smart home privacy and security management options (RQ1)

As our sample included many Apple HomeKit users, as well as participants who took complicated steps to manage data flows (see Tables 1 and 2), it is unsurprising they described conducting extensive research on privacy and security features before purchasing devices and using advanced configurations and customized settings during setup. That said, even with significant time and energy spent researching devices that aligned with their privacy and security needs, participants identified several drawbacks and limitations of smart home technologies currently available. Below, we describe three core tensions participants highlighted. Participants also raised general usability issues, but we focus solely on those connected to privacy and security.

4.1.1. Balancing privacy and security with functionality and interoperability

Most participants relied on HomeKit-compatible smart devices, with many noting they chose Apple because they trust the company and its commitment to data privacy and security. For example, P19 selected HomeKit because *"it's supposed to be really secure,"* and P13 indicated that when comparing smart device ecosystems, *"Apple was the one that had the most privacy built into it."* Participants also justified spending more money for HomeKit to achieve greater data security, with P5 saying: *"[Apple] costs you more, but you have that little bit more of a peace of mind that there's a little bit more control."*

Several participants specifically noted the privacy and security controls available for HomeKit-enabled routers, which offer a simple mobile interface that shows which devices are operating locally versus those connected to the internet. As P10 described, *"If you want it to have no internet access or just be able to connect to get firmware updates or whatever, there's different levels. It seems like the sort of thing that is ideal at the router level, because that's what's sitting between all your devices and the rest of the internet."* P19 also commented on this functionality, noting that the HomeKit control panel *"shows you all of your accessories, all of your hubs... Each accessory has the option to let it communicate freely, let it communicate to only a specific subset of domains that are strictly relevant to its operation, or only let it communicate locally."*

Others noted, however, that Apple's focus on simple interfaces can limit users' ability to manage data flows. P28 noted that default settings for devices may share more data than a user wants, while P19 acknowledged:

> *I feel like Apple largely has to appeal to the lowest common denominator… there are people who really want to get into the nuts and bolts of things, and that extra data is really valuable. And even though it might be a little overwhelming to the average user, it creates a value proposition there for people who are really, really interested in really protecting their privacy because that's their whole thing.*

Some participants acknowledged that enhanced privacy and security came with additional tradeoffs. Many HomeKit users who described using it for its enhanced privacy and security features also described frustration with its limited interoperability. P13 described challenges getting devices to work how he wanted them: *"To a certain extent, I'm sacrificing a lot of potential functionality and incurring greater cost, for the sake of being in a 'more private' environment."* This resonated with many participants, who noted general limitations when attempting to balance privacy and security with functionality and interoperability. P2 summarized this limitation, saying he spends significant time *"finding devices and figuring out the compatibility and the privacy"* and *"even then, I do buy some stuff that doesn't work the way I thought it would. That's really annoying."*

P2's comment reflects participants' experiences compromising their privacy and security preferences to achieve their automation needs. He said he bought Google Home Minis because he wanted to send text-to-speech commands through the HomePod, and the Minis had physical buttons for turning the microphone off. Likewise, P4 resorted to using an Amazon Echo for streaming music, saying, *"It wasn't my first choice for a smart device, but its capabilities*

were better than what I could find with my first choice [HomePod]. So yeah, that was a privacy tradeoff for me."

4.1.2. Balancing usability with customizability

Participants often sought to overcome limitations of interoperability by employing their technical skills, but HomeKit users faced additional drawbacks due to its general lack of customizability. Apple has long sought to simplify their products for greater usability, but for our PCPUs, this was seen as a drawback. P3 said she felt limited in what she could do in the Apple ecosystem: *"Being an old-school hacker geek, I don't like being told what I can't do. I like having the flexibility to play around."* P13 acknowledged that most people wanted (and needed) simple interfaces, but also wanted options for people with more advanced skill levels—who understood the risks and were willing to take them. He said, *"I wish Apple especially would be less about gatekeeping, and simplifying, and making things dumbed down for everybody at the expense of the few people who want to try more advanced things."* Likewise, P19 expressed frustration with the inability to easily tinker with devices, noting, *"If its biggest selling point wasn't that it was so secure, I would ditch it in a heartbeat."*

Participants described steps they took to overcome interoperability and customizability limitations with HomeKit by integrating third-party, open-source products like Homebridge or Pi-holes into their smart home ecosystem. These tools provide the ability to customize and control smart devices that may not be designed to work natively with the HomeKit ecosystem. They often require more effort and skill to install and configure than most non-power users have. P2 used a Pi-hole to block ads and track requests from external servers; one challenge he described was that *"you have to figure out which domain is associated with which device and how many there are."* P11 said one problem with these options is they're *"not necessarily user-friendly and you have to be willing to pick up coding or programming to make it work… I went into it without knowing how to program and I nearly threw a computer through a window at one point because of that."*

Participants often compared HomeKit to other products when describing customization limitations. P20, who used both Samsung's SmartThings as well as Home Assistant, said SmartThings *"had a great third-party dashboarding app I could build out an overview of my house and make it simple and easy for people to walk in and interact with."* Others suggested that attempts to balance usability with customizability have left everyone unsatisfied, such as P29: "*I think we're in this really weird state where, specific to smart home technology, it's a little too basic for the IT technology nerds, and a little too complex for the run of the mill user. So nobody's happy."*

4.1.3. More users, more challenges

A third set of drawbacks arose when multiple people were using devices. This complicated both smart device operation and participants' ability to manage privacy and security; in fact, many participants described struggling to balance their privacy and security needs with ensuring device usability and accessibility for other household members and guests.

The easiest solution for this challenge was to have a single household member (primary user) set up and manage devices. P19 said he and his partner came to an agreement where *"I'm just dealing with the whole thing. [My partner said,] 'I'm just trusting you with this. I'm not even going to try to understand. Just do it. I know you'll make it work.'"* Likewise, P30 said her husband trusts her with device setup and management because *"[he] knows that I'm a privacy person,"* while P26 said his wife lets him make purchasing decisions because she *"knows I care about the aesthetics of stuff, and I always run it by her."*

Multiple users can also create interoperability challenges. P25 noted their Apple HomeKit setup worked fine until a new (Android using) roommate joined the household: *"The thing about Apple devices in general is if you aren't in the Apple ecosystem, your friends with their 'dirty green bubbles,' they don't play well together. …I had to make a number of workarounds to make certain that everyone could still access devices."* P7 shared how she nearly switched ecosystems after purchasing a Sonos sound system, which wasn't compatible with HomeKit; in the end, she kept HomeKit rather than a more fragmented approach because she felt the latter would be less user-friendly for her daughter.

Having multiple users—including other household adults, children, and guests—led participants to consider ways to set up spaces with smart technology others could use while keeping them separate from the main ecosystem, especially when these non-primary users were not technically savvy. P22 used HomeKit as his primary hub, but he *"got an Alexa for just the guest room and got some lights that are connected over Bluetooth, so I feel confident they're local and offer that to the guest users,"* adding that this is because *"HomeKit does not have native support for limited guest users."* Multiple participants added physical (smart) switches as backups, especially for less-technical household members and guests. P16 struggled to find a smart switch for his partner that was not too complicated. He *"ended up picking a brand that acted like a light switch, and I did a lot of research to make sure that if it lost connection or if it failed to be smart, it could do everything dumb that it needed to do by itself."* P17 described *"exposing my partner to different things and easing them into that but having a physical fallback for them."* And P9 described negotiating with his less-technical wife: *"I think it's about creating off-ramps. When we moved into the new house, we just went with powered switches rather than smart bulbs. That way, at the end of the day, you can go over to the wall and hit a damn button and get the lights to turn on."*

4.1.4. Summarizing RQ1 findings

The drawbacks and limitations expressed by our power users reveal the challenges of balancing privacy and security with other features. Our participants were frustrated by difficulties in managing situations when multiple users might interact with smart devices. Some noted that if they are frustrated—with their knowledge, skills, and desire to tinker—then "average" users will be even more overwhelmed. P6 highlighted this sentiment: *"those who are not very tech-savvy, learning and understanding what does what, when you have lots of different devices, buttons, and lights and sensors, it can definitely overwhelm. Someone who's not technology-driven, you really need to create a cheat sheet for them."*

4.2. Design Feature Recommendations to Enhance Smart Home Ecosystems (RQ2)

RQ2 details the design features participants wanted to see developed or expanded that would address the drawbacks and limitations described in RQ1. In sharing their ideas, participants considered how such features would balance their needs and the needs of average users, which P29 captured when she said, *"I think there needs to be a better balance between dumbing things down so people can get [devices] up and running quickly, and being able to set up your smart home how you like it."*

Our participants wanted as much information as possible; however, they acknowledged that most users do not want or need so much information. Therein lays a challenge for designing tools that spanned all types of users; as P19 noted, *"There are people who really want to get into the nuts and bolts of things. And that extra data is really valuable. And even though it might be a little overwhelming to the average user, it creates a value proposition for people who are really interested in protecting their privacy because that's their whole thing."* At the same time, participants suggested that these features could help non-power users by *"bring[ing] the novice up to speed a little bit on what is really happening with their data"* (P6) and they would be *"something I would want to share with my family members or my partner, who I'm trying to convince"* (P17).

Below, we detail how participants balanced competing needs across different user types when describing features to enhance data transparency, visibility, and control.

4.2.1. Increase data transparency in a simple and standardized way to help users make informed decisions

Given that participants described investing significant time and energy researching SHDs, it is unsurprising they wanted device manufacturers and app developers to be more transparent regarding data collection and use practices to help them make more informed decisions both before purchasing and while using smart home technologies. Participants discussed two primary ways to make information more transparent: improved product labeling regarding data practices and improved notifications.

Our participants were largely dissatisfied with the limited information manufacturers and developers provided about data practices, and they wanted summary statements at multiple consumer touchpoints (device packaging, app, website). They believed that providing detailed information on data practices would help consumers make informed decisions *before* purchasing SHDs, including whether to purchase a device and where in their home they'd feel comfortable placing a device. P13 explained, *"You don't know until after you buy [a SHD] whether or not it's any good in terms of security or capability… so knowing ahead of time would be helpful."* Likewise, P31 said, *"the more information we can get as consumers, the better choices we can make.*

…give us all the information, be open, be transparent."

Participants described several core pieces of information that would aid them in decision making, including the types of data the device collects, how and when data is collected, and what purpose the data would be used for. Several participants specifically mentioned the need for privacy nutrition labels. P31 summarized the benefits of these labels, saying:

> *You look at your food, everything that's pre-packaged has nutritional labels on it, right? And I think if we could have some sort of electronic digital nutritional data like, "Hey, this uses this much electricity per hour, it sends out this kind of Z-Wave and Zigbee… this is the type of data we're collecting." I think that kind of nutritional label for electronics is what us, as consumers, are looking for.*

Participants liked that nutrition labels both provide key information for decision-making and do so in a clear, standardized way. P10 described Apple's privacy labels (which Apple unveiled eight months earlier) as *"distilling [information about data collection] down to something that's easily digestible and easy to compare one app to another";* they described wanting something similar for SHDs to make it *"easy to compare one device to another."* P17 reemphasized that the privacy labels should be concise and clear to help users parse out complex data practices. He said, *"I'd like a tool, like a nutrition label of sorts, that can say, 'Hey, this is the data that this device collects. This is why you should care or not care about it, and how often it does it.'"* The IoT privacy labels [16] reflect our participants' wants and needs for SHD data transparency. Additionally, participants suggested having an objective third party provide information about a device's data practices, rather than labels generated by companies, which resonates with the need for privacy ratings [21]. For example, P13 wanted *"a third-party reviewer or objective source to help folks who want to know more [information]."*

In addition to information about data practices, some partic- ipants wanted greater transparency regarding the identity of—and relationships between—device manufacturers and brands to show how data flows across other platforms or services shared by the same company. For example, P26 mentioned that *"some of these smaller companies like Aqara, a lot of stuff is just rebranded from a larger compa- ny from overseas that American consumers might not neces- sarily be aware of."* He said this was important information for making purchasing decisions because *"you might not want to use a specific company's products, but then you don't know that they've just white labeled that to somebody else."* P30 shared similar concerns, mentioning how the brand Tuya makes smart home technologies that *"all call home just by a random server in China"*—something she wanted to know before purchasing. Many participants de- scribed avoiding brands or devices from certain countries or companies because they didn't trust them to properly handle their data.

Participants also suggested ways to improve pop-ups and notifications when asking users to consent to data practices while setting up and managing SHDs. Notifications give users a chance to make informed decisions while using devices, but participants felt most lacked information to sufficiently inform and guide users. For example, P14 said, *"I find allow permissions like 'Do you allow permission for this app? That app?' woefully un-detailed. It's like, 'this app needs access to your camera.' Well, why does it need access to my camera? Why does it need access to my location? Some of the things just seem completely random."*

Like P14, many participants wanted better explanations for why devices were collecting data. P25 compared it to access requests on mobile apps, saying, *"it can give an explanation, ideally, like we need to access your water sensor to determine if the ground is wet outside."* Permissions could give users more information when deciding whether to agree to data collection. P32 summed up the idea by stating, *"if we could just have a little paragraph saying, 'this is what you're getting. This is where your data is going, and these are the people that are going to be using it,' then I can say, 'I'm cool with that' or 'I'm not cool with that.'"*

4.2.2. Greater visibility to manage device status and data flows

Along with increasing transparency about data collection and usage practices, participants underscored the importance of ongoing visibility into their smart home system and data flows, including device status; types of data being sent within and out of their network; where data is going; and how frequently data is being collected.

Building on their earlier comments regarding interoperability challenges (Section 4.1.1), participants expressed a strong desire for a standardized and centralized means to monitor their smart home data. P32 encapsulated this desire, saying: *"I want something that shows me everything in one place. Right now I have a smattering of ecosystems: Apple, Amazon, and Google. I would love for something in one app, something that I can just go and do everything."* To create a centralized tool for smart home data visibility, like P32 wanted, it would need to support a range of network protocols and technologies (e.g., HomeKit, Zigbee) to facilitate communication between devices. While Homebridge and Home Assistant already support this level of interoperability, our participants emphasized that these open-source tools require advanced networking skills beyond the out-of-the-box functionality associated with corporate ecosystems.

Within this centralized tool for managing data visibility, our participants wanted to quickly view current device status, see where they were located, and easily access their devices. Thinking about how to visually display device status, P15 noted a filtering mechanism: *"you might want to visualize it in a couple of ways. You might want to say, show me all the light switches in my house... or show me all the equipment that's operating in my kitchen right now, so it might not just be light switches in that case."* Participants also wanted to know how the house was automated and which devices or commands trigger other devices. Additionally, P28 wanted 'signal strength' to be in the overview as an indicator of device connection, to see *"if there are any particular problem spots where things might not be connecting properly."*

Several participants mentioned the importance of monitoring network traffic flow to know the data types and amount of data being shared by SHDs, as well as where device data went. P2 used a Pi-hole to *"track all the requests each site makes"* and wanted the feature to be *"as detailed as it could be and easily organized"* so he could see what data was being transmitted. P3 wanted *"to see what domains [device data] is going to and how much. I would love to see a time-of-day graph, where I could correlate, this is when I just got home so I'm seeing a large spike, and just the other day I wasn't even home and look what was going on."* Both participants' wants and needs build on to the idea of IoT Inspector [20], which visualizes network activities to identify security and privacy risks in the smart home environment.

As noted in Section 4.1, many participants wanted to keep as much of their data local as possible. In cases where a SHD required a cloud connection, participants like P10 wanted a *"visualization to show where in the world your data is going."* P10 further explained the broader concern motivating this feature request: *"I think some people might be surprised where some of their data is going, what countries it's going to. And you're like, why is my stuff going to that country? I just want to turn my lights on."*

To accompany visualizations of data flow such as where the data is being transmitted, our participants wanted additional information on *why* data was being collected. Thinking of features in a visualization tool for his parents, P9 focused on visualizing connections to services like Google AdWords. He said, *"the ability to visualize and understand [these con-*

nections], 'hey, you brought this thing online and it attempted to make 15 connections out. One was for a firmware update, one [was for] word upload logs, and the oth- ers were to shovel data out to known ad tracking agencies.'" P9 felt this information alone was simple and easy enough for his parents to interpret.

Like P9, many of our participants had other users in mind (e.g., bystanders) when thinking of ways to improve data visibility and control, which we will further elaborate in the next section. One reason our participants likely think of other users is they know firsthand the difficulty in parsing transaction logs and therefore the need for more user-friendly interfaces. As P24 noted: *"a lot of times I look at logs, or I look at lists of internet access, and I'm actually like, what is all this stuff, why are they going to all these different servers, and I just have no idea which server is which."* As such, even these advanced users don't always know why a device transmitted data to a certain domain.

4.2.3. Controls and notifications to manage smart home data

Building from the need for increased transparency and visibility, participants noted that these features were useful but still limited, especially when many devices and apps follow a "take it or leave it" approach—you can allow data collection, or you cannot use it. P22 brought this up when discussing privacy labels, saying, *"Apple's really good about saying, 'this app wants to access your photos' and allowing you control. But it's also like, well you have a nutrition label, but what's your options? I guess the question is: how low of a level can you disable stuff and still use the app? So I think that's part of the problem; it has a nutrition label and it's basically take it or leave it."* This dilemma between data minimization versus functionality is not new. For example, one of the central issues that Aretha [47] faced when providing a firewall as a control mechanism to manage data flows in smart homes was that it accidentally interfered with SHD operation. Additionally, our participants had many ideas for additional controls, including the ability to create allow/block lists for certain types of data to certain destinations, and more robust notification (e.g., alert, alarm) features that would allow users to manage data flows.

Many participants had advanced networking skills and used Pi-holes and private DNS setups to create lists of trusted and untrusted domains. That said, some expressed frustration with how these controls are often hindered by limited details on domain addresses, as well as the low accessibility of these tools. P2 captured this frustration when he said: *"If you try to wholesale block a device from connecting... it may stop working. But if you actually had the list of domains it was trying to hit, you could go through and block 90% of them and still keep your device working. So you're limiting your exposure while still getting the benefit of the device."* P15 described how greater network traffic visibility could lead to more impactful controls: *"Hopefully I'm able to say, look, this is only communicating with itself in my home hub. But maybe it's also communicating to my light switches for some reason—that's probably not a good idea. I should block that. Maybe it's communicating to some server that I don't know. I should block that."*

Others noted how better controls to create and manage block lists would provide peace of mind, such as P25's desire to *"easily put [IP addresses] in a box, if you don't feel so great about it"* or P24's wish that *"if in the certification process of the HomeKit app, or whatever they have to list to Apple, 'here's all the servers we're going to be needing to talk to and for this reason.' Then Apple says, 'okay, cool,' and they whitelist those and anything else gets blocked."*

Some participants wanted notifications (e.g., flags, alerts) to alarm users about unwanted activities based on the controls users created. For example, P25 suggested, *"if a Samsung device is contacting something other than a Samsung endpoint, that might be a red flag."* P4 suggested that after configuring a device, users could *"verify that with what the device is seeing on the network and say, 'Hey, you've told your TV, 'don't call home.' Hey look, it's calling home. Did you know that?'"* Once users are alerted to unusual activity on their network, participants wanted options to control their data, like P11, who mentioned *"a way to easily shut [data sharing] off, or shut off features, like, alright, here's the updates, but not sending out logs or…other stuff."*

Many participants wanted different settings for different types of users such as guests, domestic workers, and children. P22 mentioned that he used Alexas in guest rooms because Apple didn't yet have *"proper guest support."* P29 gave an example of babysitters (*"you want [babysitters] to be able to control lights or unlock the door, but you don't want them to be able to switch settings around")* and kids (*"you want [kids] to be able to control their room, but maybe not change the temperature level in the house").* On the other hand, P16, who self-described as a power user and said his wife had no interest in device management, said he'd rather take the initiative in a single location to manage data where *"I could go to and see a master status screen of everything going on in my house for the internet, the network traffic, data logs, power usage. I would love that."*

4.2.4. Summarizing RQ2 findings

While the PCPUs in our study benefited from their advanced technological skills to manage their smart device networks, they still looked to device manufacturers to provide better data transparency, visibility, and control. A common theme among our participants was that a user should not need special networking and security skills to understand how smart devices collect data or with whom data might be shared; manufacturers need to provide greater transparency and visibility as a default. Further, participants recognized that control over data flows they were able to create through customizations should similarly be available to all users regardless of technical proficiency. They

acknowledged that if they had to take extraordinary steps to see and manage data flows in their homes, then standard tools and interfaces will be insufficient to ensure usable data security and privacy controls for all.

5. Discussion

As the smart home landscape has expanded, usable privacy and security scholars have explored ways to make devices and automations more user friendly while encouraging expression of data rights like privacy, interoperability, and autonomy. In this study, we provide insights from privacy-conscious smart home power users who use both embedded features and third-party platforms to monitor traffic, build custom dashboards, and create custom domain block lists. By engaging with this unique set of users, we gained a deeper understanding of what features currently exist to help users understand and interact with their smart home data, the drawbacks of these tools and, most importantly, how the design of SHDs can be improved to increase the visibility and control of data flows to enhance user privacy—both for more technically savvy users as well as non-power users.

Our findings emphasize and extend many design features identified in prior work [6,10,11,17,19,21,47,53,54]. Our participants were uniquely well-positioned to identify critical limitations of existing privacy-preserving design features that prevented them from effectively mitigating privacy concerns and security risks raised by SHDs. Knowing that these power users—who were highly motivated to customize their homes to maximize benefits while minimizing external data flows—expressed significant frustration with currently available options signals an urgent need to develop more user-friendly features and controls that are accessible for everyday smart home users.

Below, we consider two aspects of smart home design that should be addressed to move beyond P29's sentiment that "nobody's happy" to one where all users can be both satisfied and confident regarding their ability to manage the privacy of their smart home data flows.

5.1. Designing for Context

A key takeaway from our findings is that PCPUs wanted smart home technologies that could better balance their functionality and interoperability needs with their desired level of privacy—and they wanted such features to be usable (and understandable) by non-power users too. In short, PCPUs recognized that transparency alone is insufficient, especially since it often takes certain technical skills to make sense of data flows within a smart home network. Rather, they pushed for a broader focus on providing all users with the ability to assess data flows in comprehensible ways and within certain contexts of use in their smart home environments.

This desire aligns well with the contextual approach to privacy championed by Nissenbaum [40,41], where the appropriateness of personal information flows is contextually bound by factors such as the actors and purposes for such flows. Our participants built custom features and setups to manage their privacy contextually—allowing some data to be collected and transmitted only within contexts deemed appropriate. A smart TV sending data to the manufacturer's IP address might be acceptable, but sharing the same data with an unknown actor was deemed inappropriate.

While transparency alone isn't sufficient, it remains important, and a major challenge when designing for context is the general lack of transparency from device providers regarding what data they collect and how they use it. As we note above, P14 bemoaned the lack of information from devices and apps regarding *why* they wanted certain permissions. This lack of transparency makes it more challenging to manage devices effectively. Many of the issues our participants raised, such as ensuring *all* users can easily understand what data is being collected, who and where it is being shared, and for what reasons, will require companies to provide contextual information in a structured data format.

However, things may be changing. New international standards like Matter [13] have the potential to increase device interoperability and ease the task of designing centralized data visualizations and controls that support all smart device manufactures—addressing some of the frustrations expressed by our participants. Beyond these standards, new regulatory measures may be required if companies still determine that such disclosures are not prudent based on existing market incentives. Recent proposals in the U.S. (e.g., American Data Privacy and Protection Act; Terms-of-service Labeling, Design and Readability Act) would require greater transparency and disclosures for how technology platforms collect and user data [15,37]. The future of such laws remains unclear, and we urge smart device companies to respond to the prompts of the PCPUs in this study in advance of any regulatory requirements.

5.2. Designing for Users at Different Skill Levels

A second design challenge speaks to a knowledge and skills barrier. Our participants' descriptions of how they managed their SHDs—often through advanced network management approaches or complicated automations—points to a need for simpler solutions that account for variations in contextual factors like who is interacting with devices (e.g., children, guests) and device location (e.g., a speaker in a bedroom is different than a speaker in the kitchen).

Our PCPUs repeatedly noted that any design enhancements that stem from their experiences and recommendations must be flexible for a diverse range of users and stakeholders. Their statements resonated with prior work suggesting that users want privacy tools that are simple, proactive, and provide more control options [21]. While our participants often wanted as much data as possible, they acknowledged that

most users would be overwhelmed with so much information—confirming experimental findings showing how new smart device users struggle with complex interfaces and dashboards [1]—and would benefit by simpler features to facilitate data control.

Specific to increasing data transparency and visibility, our participants recognized that for maximum usability, information about data practices should be provided in a summarized, digestible format, with the option for more information for those who seek it. Such a solution aligns with Emami-Naeini et al.'s [16] approach to *layered labels*, which provide two types of information: a primary layer containing the most important content, and a secondary layer containing more detailed information. Layered labels also have the potential to facilitate learning, encourage discussion of data flows with household members and bystanders, and prompt companies to be even more transparent.

Our participants further discussed expanding these labels to include more information about the company collecting data, including if they are part of a conglomerate, whether any ownership or branding changes have recently taken place, and what location data might be sent to. This aligns with prior work highlighting users' interest in the relationships between companies handling data [47]. All users would benefit from such expanded transparency within the smart home ecosystem; by taking a layered label approach, a range of information can be made available across multiple layers to avoid overwhelming users less interested in technical details.

In terms of smart home data visualization and control, our participants wanted a centralized location to monitor and control their smart home data, including device status and network traffic flow. Furthermore, participants wanted to create custom allow/deny network traffic lists and wanted notifications to be automated based on that list. Lastly, our participants wanted different modes of control for other users such as secondary users and bystanders (e.g., guests, visitors, children, and domestic workers).

Previous scholarship has explored smart home data visualization and control tools, and our participants' feedback offers insights for further development. IoT Inspector [20] labels smart home network traffic and produces tables and charts for users to monitor their smart home data. Our participants explained how this type of tool could be enhanced with filtering capabilities for device type, communication endpoints, etc. Similarly, Aretha [47] provides the daily ebbs and flows as well as aggregated smart home data to users; however, users found the control mechanism difficult to use because there were too many endpoints to comprehend. Our participants described struggling parsing out domain lists from smart home data management tools and suggested that manufacturers provide a default network traffic list that their products require to function. This will allow users of every technical skill level to start on and build upon creating their preferred network traffic list. Furthermore, to serve a variety of users and their various privacy preferences, assigning different roles and responsibilities might be an idea. One example of this is Kratos+ [48], a multi-user access control mechanism with a priority-based access-policy negotiation technique. Kratos+ applies a policy negotiation algorithm that automatically solves and optimizes conflicting user access requests based on users' set priorities on different devices. In addition to conflict resolution, users are notified when changes are made or when their requests are rejected. Although Kratos+ is de- signed for access controls, we can think of a similar mecha- nism to resolve conflicting privacy needs in smart homes.

5.3. Limitations

Participants were recruited largely through popular online discussion forums on Reddit and Facebook. This recruitment method increased the possibility of biases within our sample based on the socio-demographic characteristics of who are active in such online spaces. Future work could seek to obtain a more diverse set of smart device power users as well as seek out bystander viewpoints to directly assess their privacy concerns and strategies regarding exposure to smart devices.

6. Conclusion

With the growing adoption of smart home technologies, companies have emphasized making their products simple and user-friendly, often to the detriment of providing users with full transparency, visibility, and control over the data these devices capture and share. Complementing and extending previous studies that explore how everyday users of smart devices think about and address data privacy concerns, this paper engages specifically with privacy-conscious power users (PCPUs) to gain a clearer understanding of the steps taken by those with advanced technical skills to manage their smart homes. We identify design recommendations inspired by these power users and prompt device manufacturers to consider how such enhanced levels of data visibility and control should not be restricted only to those with the skills to customize their smart environments. The data privacy and security afforded by these suggestions should benefit all users.

The smart device ecosystem continues to evolve, and the growing use of artificial intelligence to better learn and adapt to users' behavior and preferences [4,46] only increases the need for the expanded collection of user data by device companies. At the same time, new standards promise to make smart devices more ubiquitous and easier to use, likely yielding in fewer opportunities for users to have full visibility or control into how data is collected and used. While the PCPUs in our study might make do, they also acknowledged that "nobody's happy" when it takes extensive technical skills to maintain privacy, or more typical users are left without usable means to manage their privacy.

References

1. Jacob Abbott, Jayati Dev, Donginn Kim, Shakthidhar Gopavaram, Meera Iyer, Shivani Sadam, Shrirang Mare, Tatiana Ringenberg, Vafa Andalibi, and L. Jean Camp. 2022. Privacy Lessons Learnt from Deploying an IoT Ecosystem in the Home. In *Proceedings of the 2022 European Symposium on Usable Security* (EuroUSEC '22), 98–110. https://doi.org/10.1145/3549015.3554205
2. Noura Abdi, Xiao Zhan, Kopo M. Ramokapane, and Jose Such. 2021. Privacy Norms for Smart Home Personal Assistants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3411764.3445122
3. Imtiaz Ahmad, Rosta Farzan, Apu Kapadia, and Adam J. Lee. 2020. Tangible Privacy: Towards User-Centric Sensor Designs for Bystander Privacy. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2: 1–28. https://doi.org/10.1145/3415187
4. Amos. 2022. Artificial Intelligence Is the Next Step for Smart Homes. *Unite.AI*. Retrieved February 14, 2023 from https://www.unite.ai/artificial-intelligence-is-the-next-step-for-smart-homes/
5. Noah Apthorpe, Yan Shvartzshnaider, Arunesh Mathur, Dillon Reisman, and Nick Feamster. 2018. Discovering Smart Home Internet of Things Privacy Norms Using Contextual Integrity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2: 1–23. https://doi.org/10.1145/3214262
6. Nata Barbosa, Zhouhao Zhang, and Yang Wang. 2020. Do Privacy and Security Matter to Everyone? Quantifying and Clustering User-Centric Considerations About Smart Home Device Adoption. *Usenix*. Retrieved from https://www.usenix.org/conference/soups2020/presentation/barbosa
7. Bruhadeshwar Bezawada, Maalvika Bachani, Jordan Peterson, Hossein Shirazi, Indrakshi Ray, and Indrajit Ray. 2018. IoTSense: Behavioral Fingerprinting of IoT Devices. Retrieved from http://arxiv.org/abs/1804.03852
8. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2: 77–101. https://doi.org/10.1191/1478088706qp063oa
9. Joseph Bugeja, Andreas Jacobsson, and Paul Davidsson. 2016. On Privacy and Security Challenges in Smart Connected Homes. In *2016 European Intelligence and Security Informatics Conference)*, 172–175. https://doi.org/10.1109/EISIC.2016.044
10. George Chalhoub, Martin J Kraemer, Norbert Nthala, and Ivan Flechais. 2021. "It did not give me an option to decline": A Longitudinal Analysis of the User Experi- ence of Security and Privacy in Smart Home Products. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3411764.3445691
11. Chola Chhetri and Vivian Genaro Motti. 2022. User-Centric Privacy Controls for Smart Homes. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2: 349:1-349:36. https://doi.org/10.1145/3555769
12. Eun Kyoung Choe, Sunny Consolvo, Jaeyeon Jung, Beverly Harrison, and Julie A. Kientz. 2011. Living in a glass house: a survey of private moments in the home. In *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11*, 41. https://doi.org/10.1145/2030112.2030118
13. Connectivity Standards Alliance. 2022. Matter Arrives Bringing A More Interoperable, Simple And Secure Internet Of Things to Life. *CSA-IOT*. Retrieved from https://csa-iot.org/newsroom/matter- arrives/
14. Julia C. Dunbar, Emily Bascom, Ashley Boone, and Alexis Hiniker. 2021. Is Someone Listening?: Audio-Related Privacy Perceptions and Design Recommendations from Guardians, Pragmatists, and Cynics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3: 1–23. https://doi.org/10.1145/3478091
15. Gilad Edelman. 2022. Congress Might Actually Pass ADPPA, the American Data Privacy and Protection Act | WIRED. *Wired*. Retrieved February 11, 2023 from https://www.wired.com/story/american-data-privacy-protection-act-adppa/
16. Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the Experts: What Should Be on an IoT Privacy and Security Label? In *2020 IEEE Symposium on Security and Privacy (SP)*, 447–464. https://doi.org/10.1109/SP40000.2020.00043
17. Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. 2017. Privacy Expectations and Preferences in an IoT World. *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*: 399--412.
18. Margaret Hagan. 2016. User-centered privacy communication design. In *Proceedings of the Symposium on Usable Privacy and Security*, 22–24. Retrieved from https://www.usenix.org/conference/soups2016/workshop-program/wfpn/presentation/hagan
19. Julie M. Haney, Susanne M. Furman, and Yasemin Acar. 2020. Smart Home Security and Privacy Mitigations: Consumer Perceptions, Practices, and Challenges. *NIST*. Retrieved from https://www.nist.gov/publications/smart-home-security-and-privacy-mitigations-consumer-perceptions-practices-and
20. Danny Yuxing Huang, Noah Apthorpe, Frank Li, Gunes Acar, and Nick Feamster. 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: 1–21. https://doi.org/10.1145/3397333

21. Haojian Jin, Boyuan Guo, Rituparna Roychoudhury, Yaxing Yao, Swarun Kumar, Yuvraj Agarwal, and Jason I. Hong. 2022. Exploring the Needs of Users for Supporting Privacy-Protective Behaviors in Smart Homes. In *CHI Conference on Human Factors in Computing Systems*, 1–19. https://doi.org/10.1145/3491102.3517602
22. Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. 2009. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security - SOUPS '09*, 1. https://doi.org/10.1145/1572532.1572538
23. Euiyoung Kim, JungKyoon Yoon, Jieun Kwon, Tiffany Liaw, and Alice M. Agogino. 2019. From Innocent Irene to Parental Patrick: Framing User Characteristics and Personas to Design for Cybersecurity. *Proceedings of the Design Society: International Conference on Engineering Design* 1, 1: 1773–1782. https://doi.org/10.1017/dsi.2019.183
24. Richard A. Krueger and Mary Anne Casey. 2014. *Focus Groups: A Practical Guide for Applied Research*. SAGE Publications, Inc, Los Angeles.
25. Albrecht Kurze, Andreas Bischof, Sören Totzauer, Michael Storz, Maximilian Eibl, Margot Brereton, and Arne Berger. 2020. Guess the Data: Data Work to Understand How People Make Sense of and Use Simple Sensor Data from Homes. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376273
26. Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 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
27. Scott Lederer, Jennifer Mankoff, and Anind K. Dey. 2003. Who wants to know what when? privacy preference determinants in ubiquitous computing. In *CHI '03 extended abstracts on Human factors in computing systems,* 724-725. https://doi.org/10.1145/765891.765952
28. Hosub Lee and Alfred Kobsa. 2016. Understanding user privacy in Internet of Things environments. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, 407–412. https://doi.org/10.1109/WF-IoT.2016.7845392
29. Christian Leichsenring, Jiajun Yang, Jan Hammerschmidt, and Thomas Hermann. 2016. Challenges for smart environments in bathroom contexts. In *Proceedings of the 1st Workshop on Embodied Interaction with Smart Environments* (EISE '16), 1–7. https://doi.org/10.1145/3008028.3008033
30. Tianshi Li, Kayla Reiman, Yuvraj Agarwal, Lorrie Faith Cranor, and Jason I. Hong. 2022. Understanding Challenges for Developers to Create Accurate Privacy Nutrition Labels. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), 1–24. https://doi.org/10.1145/3491102.3502012
31. Wenda Li, Tan Yigitcanlar, Isil Erol, and Aaron Liu. 2021. Motivations, barriers and risks of smart home adoption: From systematic literature review to conceptual framework. *Energy Research & Social Science* 80: 102211. https://doi.org/10.1016/j.erss.2021.102211
32. Yuting Liao, Jessica Vitak, Priya Kumar, Michael Zimmer, and Katherine Kritikos. 2019. Understanding the Role of Privacy and Trust in Intelligent Personal Assistant Adoption. In *Information in Contemporary Society* (Lecture Notes in Computer Science), 102–113. https://doi.org/10.1007/978-3-030-15742-5_9
33. S. Marathe, S. Sundar, M. Bijvank, H. C. V. Vugt, and J. Veldhuis. 2007. Who are these power users anyway? Building a psychological profile. Retrieved July 6, 2022 from https://www.semanticscholar.org/paper/Who-are-these-power-users-anyway-Building-a-profile-Marathe-Sundar/1455563bf9242612c36f08e5a295aa139b8a1f04
34. Karola Marky, Sarah Prange, Max Mühlhäuser, and Florian Alt. 2021. Roles Matter! Understanding Differences in the Privacy Mental Models of Smart Home Visitors and Residents. In *20th International Conference on Mobile and Ubiquitous Multimedia*, 108–122. https://doi.org/10.1145/3490632.3490664
35. Faith McCreary, Alexandra Zafiroglu, and Heather Patterson. 2016. The Contextual Complexity of Privacy in Smart Homes and Smart Buildings. In *HCI in Business, Government, and Organizations: Information Systems*, Fiona Fui-Hoon Nah and Chuan-Hoo Tan (eds.). Springer International Publishing, Cham, 67–78. https://doi.org/10.1007/978-3-319-39399-5_7
36. Markus Miettinen, Samuel Marchal, Ibbad Hafeez, N. Asokan, Ahmad-Reza Sadeghi, and Sasu Tarkoma. 2017. IoT SENTINEL: Automated Device-Type Identification for Security Enforcement in IoT. In *2017 IEEE 37th International Conference on Distributed Computing Systems,* 2177–2184. https://doi.org/10.1109/ICDCS.2017.283
37. Carrie Mihalcik. 2022. TLDR Act aims to make website terms of service easier to understand. *CNET*. Retrieved from https://www.cnet.com/news/politics/tldr-act-aims-to-make-website-terms-of-service-easier-to-understand/
38. Matthew B. Miles, A. Michael Huberman, and Johnny Saldaña. 2013. *Qualitative Data Analysis: A Methods Sourcebook*. SAGE Publications, Inc, Los Angeles, CA.
39. Lily Hay Newman. 2020. Apple's App "Privacy Labels" Are Here—and They're a Big Step Forward. *Wired*. Retrieved from https://www.wired.com/story/apple-app-privacy-labels/
40. Helen Nissenbaum. 2004. Privacy as contextual integrity. *Washington Law Review* 79: 119–157. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10/

41. Helen Nissenbaum. 2010. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford Law Books, Stanford, Calif.
42. Michael Quinn Patton. 2014. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. SAGE Publications.
43. Sarah Prange, Ahmed Shams, Robin Piening, Yomna Abdelrahman, and Florian Alt. 2021. PriView– Exploring Visualisations to Support Users' Privacy Awareness. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (CHI '21), 1–18. https://doi.org/10.1145/3411764.3445067
44. Johnny Saldana. 2021. *The Coding Manual for Qualitative Researchers*. SAGE Publications Ltd, Los Angeles.
45. Benjamin Saunders, Julius Sim, Tom Kingstone, Shula Baker, Jackie Waterfield, Bernadette Bartlam, Heather Burroughs, and Clare Jinks. 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity* 52, 4: 1893–1907. https://doi.org/10.1007/s11135-017-0574-8
46. Samad Sepasgozar, Reyhaneh Karimi, Leila Farahzadi, Farimah Moezzi, Sara Shirowzhan, Sanee M. Ebrahimzadeh, Felix Hui, and Lu Aye. 2020. A Systematic Content Review of Artificial Intelligence and the Internet of Things Applications in Smart Home. *Applied Sciences* 10, 9: 3074. https://doi.org/10.3390/app10093074
47. William Seymour, Martin J. Kraemer, Reuben Binns, and Max Van Kleek. 2020. Informing the Design of Privacy-Empowering Tools for the Connected Home. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376264
48. Amit Kumar Sikder, Leonardo Babun, Z. Berkay Celik, Hidayet Aksu, Patrick McDaniel, Engin Kirda, and A. Selcuk Uluagac. 2022. Who's Controlling My Device? Multi-User Multi-Device-Aware Access Control System for Shared Smart Home Environment. *ACM Transactions on Internet of Things*, 1–39. https://doi.org/10.1145/3543513
49. Yunpeng Song, Yun Huang, Zhongmin Cai, and Jason I. Hong. 2020. I'm All Eyes and Ears: Exploring Effective Locators for Privacy Awareness in IoT Scenarios. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20), 1–13. https://doi.org/10.1145/3313831.3376585
50. David W. Stewart & Prem Shamdasani. 2017. Online Focus Groups, *Journal of Advertising*, 46:1, 48-60, https://doi.org/10.1080/00913367.2016.1252288
51. Neilly Tan, Richmond Wong, Audrey Desjardins, Sean Munson, and James Pierce. 2022. Monitoring Pets, Deterring Intruders, and Casually Spying on Neighbors: Everyday Uses of Smart Home Cameras. In *CHI Conference on Human Factors in Computing Systems*, 1–25. https://doi.org/10.1145/3491102.3517617
52. Parth Kirankumar Thakkar, Shijing He, Shiyu Xu, Danny Yuxing Huang, and Yaxing Yao. 2022. "It would probably turn into a social faux-pas": Users' and Bystanders' Preferences of Privacy Awareness Mechanisms in Smart Homes. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), 1–13. https://doi.org/10.1145/3491102.3502137
53. Yaxing Yao, Justin Reed Basdeo, Oriana Rosata Mcdonough, and Yang Wang. 2019. Privacy Perceptions and Designs of Bystanders in Smart Homes. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW: 1–24. https://doi.org/10.1145/3359161
54. Eric Zeng, Shrirang Mare, and Franziska Roesner. 2017. End user security & privacy concerns with smart homes. In *Proceedings of the Thirteenth USENIX Conference on Usable Privacy and Security* (SOUPS '17), 65–80.
55. Eric Zeng and Franziska Roesner. 2019. Understanding and Improving Security and Privacy in {Multi-User} Smart Homes: A Design Exploration and {In-Home} User Study. 159–176. https://www.usenix.org/conference/usenixsecurity19/presentation/zeng
56. Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Smart Home IoT Privacy. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW: 1–20. https://doi.org/10.1145/3274469
57. Bu Zhong. 2013. From smartphones to iPad: Power users' disposition toward mobile media devices. *Computers in Human Behavior* 29, 4: 1742–1748. https://doi.org/10.1016/j.chb.2013.02.016

Appendix: Virtual Focus Group Protocol

Thank you for joining us today. The format of this session is a focus group. If you've never done one of these before, I have a set of questions I'd like to open up to discussion, but there's no formal method for answering. I encourage everyone to share their thoughts. My role is merely to facilitate the conversation; you all will be guiding it.

We're here today to talk about smart home technologies. This includes a wide range of devices, from smart thermostats and smart speakers, to connected TVs, fridges, vacuums, doorbells, security systems, toys, and more. We want to build a tool that helps consumers understand the types of data that are collected and used by these smart technologies, so the main goal of today's session is to learn from you about what information you think is important and what factors would make a tool like this useful to you.

We'll be recording the session today, but I want to assure you that whatever is being shared in this room today stays with us, and anything we use from this conversation will not be connected to your real name. That said, please treat this session as confidential and do not share things we discussed today with others. This session is scheduled to last 60 minutes. Does anyone have questions before we start?

Great, so let's get started. As a warm-up, let's talk about what types of smart technologies you use in your home. Can we go around the group and each of you share your name and then walk us through a normal day and **talk about the various devices you interact with and how you might use them. What do you like the most about them? What do you dislike about them?**

[Discussion]

For those of you who share your home with other people, smart devices may pose an interesting challenge because it's hard to "opt out" of using them. So we're curious, how do you make decisions about buying and using smart devices?

- Prompt (if needed): Is there one person who is "in charge" of managing these devices?
- Prompt (if needed): Do you have discussions with other household members before buying or setting up a device?

[Discussion]

Because you're using a range of devices, we can talk about building up an ecosystem of smart technologies that talk to each other and potentially share data. Thinking about that, one thing we want to hear more about is how you decide whether to connect smart devices to each other.

- Prompt (if needed): Do any of you struggle with the technical aspects of setting up and using these devices? Can you share an example of how that affected your decision on how to use a device?
- Prompt (if needed): Are there any devices you don't want to connect? Why?

[Discussion]

We're also interested in hearing about any times you've maybe been concerned about data being collected or transmitted by your devices. Are you ever worried about data being collected by one of your devices, or things a smart device might have "overheard" or collected without you knowing?

[Discussion]

Okay, we're going to spend the rest of the hour doing some design thinking activities. We're interested in ways to better share smart device data with consumers, and we want to think creatively about what that could look like.

[Design Thinking: Part 1]

Next, we want you to brainstorm all the types of data smart devices might collect about you and how you prioritize control over this data. Each of you have been assigned a Jamboard (virtual whiteboard) page. For the next few minutes list each type of data you can think of that is sent or received by each of your smart home devices (one per sticky note).

Place each sticky note on the grid provided. The grid has two axes capturing how sensitive a piece of data is to you and how much you want to be able to monitor and control that piece of data. So along the horizontal axis, place data you consider to be most sensitive on the right and along the vertical axis, place the data you would like to have more visualization or control over in the top half.

[Answer questions, give them three minutes to do this, then summarize the themes briefly and ask if anyone has things to add.]

[Design Thinking: Part 2]

To wrap up, I'd like you to get your thoughts on what types of features you would want in a tool that helps you visualize the data your smart devices collect and share. I realize this is kind of an abstract question, there are no wrong answers. You'll have three minutes to jot down as many feature ideas as you can come up with, then we can talk them through. *[time permitting, use Jamboard; otherwise, have a group discussion]*

[Facilitator note: As all the sticky notes are posted we can start to look for trends/themes that emerge and group them, this often leads to a more fruitful discussion. If short on time, skip the sticky noting part and just ask them to discuss features as a group.]

Wrap-up: *thank everyone for attending and let them know about getting gift cards and that we'll share results once this is written up.*