# "What *Can't* Data Be Used For?" Privacy Expectations about Smart TVs in the U.S.

Nathan Malkin*, Julia Bernd†, Maritza Johnson†, Serge Egelman*†
*University of California, Berkeley
Email: {nmalkin,egelman}@cs.berkeley.edu
†International Computer Science Institute
Email: {jbernd,maritzaj}@icsi.berkeley.edu

***Abstract*—Smart TVs have rapidly become the most common smart appliance in typical households. In the U.S., most television sets on the market have advanced sensors not traditionally found on conventional TVs, such as a microphone for voice commands or a camera for photo or video input. These new sensors enable features that are convenient, but they may also introduce new privacy implications. We surveyed 591 U.S. Internet users about their current understanding and expectations about how smart TVs collect and use data. We found a wide range of assumptions and opinions among our respondents, and a good deal of uncertainty about what's collected and how it is used. In addition, these assumptions and opinions varied between data types and sensors. One area where we found broad agreement was that it is unacceptable for the data to be repurposed or shared. But there was little understanding of the protections—or lack thereof—afforded by current laws and regulations to constrain such sharing. We hope that our findings will enhance end-user privacy by providing useful insights for smart TV manufacturers, regulators and lawmakers, and designers of privacy-enhancing technologies.**

## I. Introduction

Televisions have been an integral part of many households since the 1950s; by the late 1990s, 98% of households in the U.S. had at least one TV [48]. Today, most televisions on the market are "smart TVs": in 2017, as many as 70% of TVs being sold [8], adding up to nearly 70 million in the U.S. [35]. Smart TVs are Internet-enabled, and many have new features powered by sensors that would not be found on a conventional television. For example, you can find a movie by saying a line from it, or you can gesture to adjust the volume. Such features leverage sensors that collect data from the surrounding environment, such as microphones, video cameras, and motion sensors. The presence of new sensors and additional data collection pose security problems and privacy risks [18], [38], [5], [43], [23]. Do people today understand these risks? What are their current expectations about sensor data and how it is used? What do they find unacceptable, and what mitigations are needed?

Recent studies have found that people are most concerned with data capture in the home [33], [40]. Sensor-enabled smart TVs introduce exactly this risk: some of their features are only possible by capturing data from the environment surrounding the smart TV. Concerns related to smart TVs may be even higher, as many households have multiple TVs, sometimes in the bedroom. However, it's easy to overlook the new sensors and data-driven features. Unlike most other smart home devices, which are typically bought for their "smarts", a consumer might well purchase a smart TV without wanting or even knowing about those features, simply because most models now on the market have them. So users may be unaware of what data is being collected and what's being sent to various servers; after all, few users are even aware of how data flows when on Wi-Fi [29].

To explore these issues, we designed a survey to understand people's current understandings and expectations about smart TVs and the data they collect. What assumptions do people have about how data is collected, who has access, and what safeguards exist?

We found that there is a wide range in understanding and expectations, even among smart TV owners. For most of our questions, participants' understandings, preferences, and expectations of safeguards depended on what type of data we were asking about, with visual sensor data considered both most sensitive and most likely to be kept private. However, even within data types, our respondents did not tend to show clear agreement about where the data is processed (on the device or on a server) and who might have access to it (will a human view the data? will a third party have access?).

Confusion notwithstanding, our respondents were clear that it is unacceptable for the data to be repurposed for advertising or other uses and that they expected manufacturers to protect the data from hackers. Despite their wishes, over a third of respondents believed they had no legal recourse if data was repurposed or shared with third parties. Perhaps more concerning were those who did believe—incorrectly—that there are strong legal protections and who more often tended to view sharing and repurposing as unlikely.

In this paper, we present our survey results and discuss opportunities for enhancing user privacy in the home. Possible mitigations include changes that manufacturers or service providers could make to provide improved notice and better controls. We also discuss potential regulations to provide the safeguards that people believe are already in place and/or third-party tools to help users manage data collection and transfer by smart devices in the home.

European Workshop on Usable Security (EuroUSEC) 2018
23 April 2018, London, England
ISBN 1-1891562-54-1
https://dx.doi.org/10.14722/eurousec.2018.23007

## II. Related Work

Privacy researchers have long observed a discrepancy between people's stated privacy preferences and their actual behavior, a phenomenon known as the "privacy paradox" [31]. However, the reality is more nuanced than that. For example, a survey by Felt, Egelman, and Wagner [16] found that smartphone users' reactions to third-party apps accessing their information depends on the type of data and who would see it. Participants were least concerned about servers getting the data and most concerned about the general public. Experiments by the same authors [15] showed people are more likely to act in accord with their stated privacy preferences when they actually have options to choose from. Withholding information, or self-censorship, is a common privacy-preserving behavior [46], though it takes care to properly measure it [12]. And even though people have difficulty correctly managing all unwanted disclosures [34], [37], many people will take basic steps to eliminate the extremes [27].

Some people may forgo privacy-preserving behaviors because they mistakenly believe that there are adequate legal safeguards. An Annenberg survey in 2005 [49] found that a majority of U.S. consumers assumed—incorrectly—that there were laws preventing companies from sharing information about their purchases, and that the fact that a website has a privacy policy means it won't share their information. A similar survey in 2015 [50] found a smaller percentage of respondents believing these things, but still a majority (for at least some types of consumer data). Similarly, Hoofnagle and King [25] found in 2007 that Californians believed incorrectly that many types of consumer information were protected by law. Another survey by the latter authors [26] found similar results to Annenberg on privacy policies; a majority of online shoppers and a plurality of non–online shoppers believed incorrectly that privacy policies protect them from data sharing.

While users may be unduly optimistic about legal protections, they are actually pessimistic about unauthorized access. A recent Pew survey [36] found that many U.S. Internet users were not confident that the personal data kept by a variety of entities (from government agencies to online advertisers) would remain "private and secure." Interestingly for our purposes, cable TV companies and online video sites were among the entities for which respondents were most likely to say they were not sure. With regard to smart home technologies, a survey of IoT technologies found that about half of respondents believed there was a risk the devices would record private activities—though a smaller proportion thought there was a risk of "invasion of privacy" *per se* [53].

Of particular relevance to our survey are studies that explore differences in smart-device privacy expectations and concerns about data types, collection, or data recipient. In a survey on wearable devices, Lee *et al.* [33] found that participants were most concerned about photos and videos being shared, as well as financial and account information. Similarly, Aleisa and Renaud's survey of Saudi consumers [2] found that video recordings were second only to financial data and ID numbers in level of concern. A survey by Naeini *et al.* [40], [11] found that participants were most concerned about videos and personally identifying data (among data types), and about data collected via biometric scanners, followed by cameras and facial recognition (among device types).

Naeini *et al.* [40] also explored differences by place of collection, finding that data collected in the home was by far the biggest concern (even more concerning than public restrooms). Similarly, Lee *et al.* [33] found that photos and videos taken at home were among the most sensitive data types. Interestingly, a survey by Groopman and Etlinger [22] found that selling of data collected by IoT devices in the home was of slightly below average concern among the locations they queried—but that respondents felt the most strongly that they should be notified about in-home collection.

As in prior research, Lee *et al.* [33] found that respondents were less likely to be concerned about data from wearables being shared with apps than with humans. However, they note that differences based on recipient were generally weaker than differences based on data type. In contrast, interviews and in-home experiments by Choe *et al.* [7] found that participants' acceptance of sensor recording depended more on purpose and recipient—and on how data would be recorded and processed—than on data type. In qualitative interviews, Binns *et al.* [4] found that participants' concerns about data-sharing were dependent on their pre-existing relationships with and knowledge about the specific parties receiving the data.

However, as is frequently noted, concerns about privacy do not necessarily translate into action [31]. The privacy paradox may be even stronger for IoT devices than for conventional devices and online services. In interviews, Williams *et al.* [51] found that smart device users were slightly less concerned about privacy risks than non-users, but most strikingly, there was a much stronger disparity between concern and general privacy mitigation actions than for non-IoT-users.

In addition to this body of work on smart device privacy, a few researchers have looked at users' perceptions and concerns about smart TVs specifically. Consumer Reports [9] surveyed 38,000 smart TV owners and found that 51% were concerned about smart TV privacy and 62% about security.

Ghiglieri *et al.* [20] surveyed 171 smart TV owners, and potential owners, to examine awareness of privacy risks and concerns about data collection, sharing, and misuse. Very few respondents (16%) mentioned privacy risks in response to a free-answer question. Setting aside scenarios that involved hackers, respondents were most concerned about use of voice recognition data, followed by web browsing data, and least concerned about TV viewing history.

These preliminary findings necessitate a large-scale, systematic exploration of the range of users' understanding and concerns. What is specifically needed is a thorough investigation of smart TV owner and non-owners' attitudes and beliefs about particular sensors, data types, potential recipients, data uses, and the existing laws. A better understanding of these facets will help smart TV manufacturers, service providers, regulators, and designers to develop appropriate strategies to enable users to achieve their desired level of privacy in the dimensions most important to them.

## III. Methodology

In November 2016, we recruited participants to "help us better understand people's attitudes about Smart TVs" using Mechanical Turk. After asking whether respondents owned or

| # | Feature | Data Type | Description | N |
|---|---|---|---|---|
| 1 | voice recognition | audio | Voice recognition allows the user to speak various commands instead of using a remote control. | N = 90 |
| 2 | personalized recommendations | viewing history | Based on your viewing habits, Smart TVs might recommend new shows to watch that you might also like. | N = 109 |
| 3 | gesture recognition | video | Gesture recognition eliminates the need to use a remote control, by allowing you to use only your hands. | N = 86 |
| 4 | user recognition | photos | User recognition allows the Smart TV to differentiate between members of the same household. This allows the Smart TV to offer personalized recommendations for each viewer. | N = 110 |
| 5 | presence detection | photos | Presence detection means that the Smart TV will know when someone enters or leaves the room. | N = 91 |
| 6 | third-party apps | | Different content providers may offer their own apps that can be downloaded and installed on the Smart TV. | N = 105 |

TABLE I. OUR 591 RESPONDENTS WERE RANDOMLY ASSIGNED TO ONE OF SIX CONDITIONS.

were considering buying a smart TV (which we defined only as an Internet-enabled TV), we asked non-owners to pretend they owned one for the remaining questions. We advertised the IRB-approved survey to Turkers in the U.S. with an approval rate of $\geq$ 95% on 500 or more HITs. We compensated respondents $2.00; the survey took on average 15 minutes, 10 seconds.

In total, 626 respondents submitted surveys. We eliminated incomplete responses and those failing an attention check to get a final dataset of 591 people: 46% were female, 36% reported completing a Bachelor's degree, and the average age was 37 years old (range: 19–74). Of our respondents, 61% were smart TV owners, 27% were not owners but were considering buying one, and 12% were non-owners.

We designed the survey to measure current understanding and expectations about smart TVs and sensor data. To explore how these varied between different types of data, the survey was broken into six conditions. Each described a smart TV feature and asked the respondent to rate its utility. Five of the features required data that would be collected by the smart TV: audio for voice recognition, TV viewing history for personalized recommendations, video for gesture recognition, or photos for user recognition or presence detection. The sixth condition asked about third-party apps. The conditions and features are described in Table I; see the Appendix for the survey instrument.

The survey asked about how data was collected and where it was analyzed, asked respondents to rate the likelihood that various parties might gain access to the data, and then asked how acceptable it would be if that party accessed the data. Finally, we asked whether the data was likely to be combined for other uses, whether there are laws to limit data-sharing, and how data-sharing might impact the respondents' desire to buy smart TVs and use smart features.

Some of the questions were open-ended: how data is collected and how it might be repurposed (conditions 1–5), possible risks from malware associated with third-party apps (condition 6), what laws regulate data-sharing, and whether participants had thought about these issues previously. Common themes in the responses were identified and coded independently by two raters, who afterwards discussed any conflicting codes to resolve discrepancies. We report agreement using Kupper and Hafner's statistic for assessing interrater agreement for multiple-attribute responses [32].

## IV. SURVEY RESULTS

This section reports survey respondents' beliefs about how smart TV data is collected, analyzed, used, shared, and protected from hackers; preferences about sharing; and beliefs about legal protections.

### A. Beliefs about Data Flows and Uses

To determine users' beliefs about data flows, we asked a series of questions about when data leaves the device, who it might be shared with, and how it might be used.

*1) What Data Is Collected?:* Understanding how a feature works can be critical to understanding its privacy implications. We asked what the TV does while it is waiting to perform the feature—does it constantly record the room (audio, video, or photos), does it only listen/watch for certain commands, or does it do something else?

This open-ended phrasing elicited responses that could not reliably be thematically coded to categorize participants' models of the data collection process. However, the unclear and scattered nature of the responses was telling: users do not have a consensus about how smart TVs collect sensor data—even for voice recognition, a relatively familiar feature in comparison with some of the others. Understandings of voice and gesture recognition ranged from "I believe it only records your voice once it is activated for that purpose at that time" and "It is constantly monitoring video from the room and analyzing the stream without recording it" to "it records all audio constantly, waiting for commands" and "I would assume it records everything until it recognizes a gesture in its memory." Some explicitly expressed uncertainty: "It listens for specific commands, for sure, but it might do something else as well."

When asked about photos for user recognition and presence detection, many participants assumed they were motion-triggered: "It probably takes photos if it detects movement in the room." However, some respondents thought photo-taking might be constant: "I would assume it has to constantly take photos of the room, otherwise how would it know when you leave the room and then come back."

A number of respondents assumed data capture is under the user's control: "I think user recognition program [*sic*] only takes a photo when prompted." "You probably need to push a button telling it when to learn the commands." A few offered explanations related to performance, but came to very different conclusions: "It has to record all video in the room or how else could it know I am ever performing a gesture" versus "only watch for specific commands so normal movement wouldn't disturb it." For better or worse, some respondents based assumptions on their understandings of other devices: "I think it only watches specific commands, because I had an xbox Kinect, and it would do something of that nature."

Condition 2 asked whether, when the smart TV is analyzing viewing history, it examines all the shows the user has watched, it only checks to see if they've watched specific shows, or something else. Respondents in this condition—a feature that

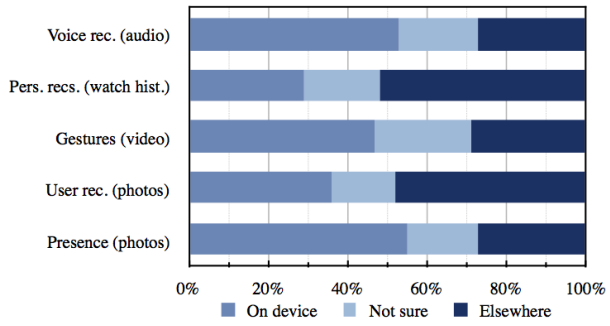


Fig. 1. Responses to "Do you believe that [the TV analyzes] this [data]...itself or does it transmit the [data]...to be analyzed elsewhere?" and "How confident are you...?" by condition.

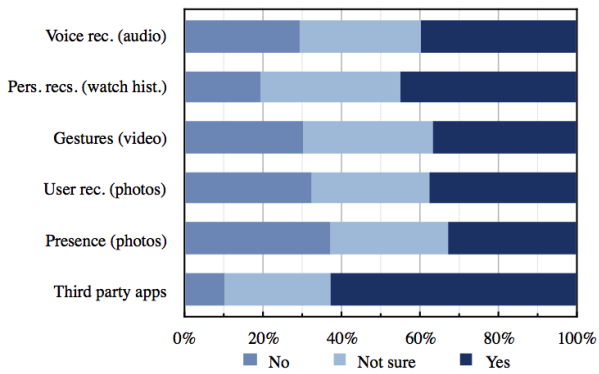


Fig. 2. Responses to "Do you believe that a human being will have access to your [data]?" by condition.

works similarly to well-known recommendation systems and doesn't involve sensors—exhibited the most consistency. Most said something like "the Smart TV examines all the shows that have been watched."

Some participants expressed concerns that provide concrete insights into how a feature's implementation might affect its acceptability: "I imagine it only watches for specific commands. I wouldn't want to use this technology if it recorded everything in the room, it would be an invasion of privacy." "I suspect it actually records all video in the room, which I do not like at all."

*2) Where Is the Data Analyzed?:* We asked respondents, "When performing [feature], the smart TV must record [data type]. Do you believe that it analyzes this [data type] for [feature] on the TV itself or does it transmit the data across the Internet to be analyzed elsewhere?" We found that beliefs about where analysis took place varied depending on the feature. In the case of audio to support voice recognition, video for gesture controls, and photos for presence detection, roughly half of respondents in those conditions believed the data would be analyzed on the device (see Figure 1). In the remaining conditions—personalized recommendations and user recognition—the leading belief was that analysis took place off-device (52% and 48%). The possible responses included "not sure" as well; "not sure" responses ranged between 15% (user recognition) and 24% (gesture recognition). The difference in proportions per condition is significant (Pearson's $\chi^2(8) = 30.687$, $p = 0.0001597$).

Most respondents were not particularly confident in their knowledge of where the data was analyzed, though again, confidence varied by condition. Slightly more than half the respondents in the personalized recommendations condition rated themselves as confident in their answer (55%). Respondents in the presence detection condition were the least confident (37%). Differences between smart TV owners, prospective owners, and non-owners were not significant for beliefs (per Pearson's $\chi^2$) nor confidence (per Kruskal-Wallis).

*3) Who Might Have Access to the Data?:* Previous research has shown that privacy concerns often depend on who has access to the data (see §II). We therefore asked, "Do you believe that a human being will have access to the [data type] recorded by your smart TV?" Across conditions 1–5, about 39% of respondents thought a human being would have access

| | Access Likelihood | | | | | Acceptability | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | C1 | C2 | C3 | C4 | C5 | C1 | C2 | C3 | C4 | C5 |
| advertisers | 55% | 82% | 50% | 68% | 53% | 8% | 29% | 5% | 3% | 5% |
| law enforcement | 39% | 20% | 33% | 49% | 40% | 25% | 16% | 13% | 15% | 12% |
| TV networks | 44% | 70% | 37% | 35% | 33% | 10% | 52% | 10% | 6% | 10% |
| ISPs | 37% | 44% | 37% | 33% | 35% | 10% | 17% | 12% | 3% | 8% |
| cable providers | 40% | 74% | 48% | 40% | 36% | 13% | 46% | 14% | 5% | 13% |
| data aggregators | 59% | 74% | 62% | 67% | 52% | 9% | 18% | 9% | 5% | 4% |
| hardware manufacturers | 66% | 48% | 67% | 50% | 49% | 26% | 33% | 24% | 15% | 16% |
| financial companies | 29% | 27% | 20% | 25% | 21% | 4% | 5% | 1% | 1% | 4% |
| hackers | 44% | 27% | 49% | 55% | 52% | 6% | 2% | 2% | 1% | 2% |

Fig. 3. Left: Responses of *somewhat* or *completely likely* to "How likely is it that [data] will be accessed by the following parties?" by condition. Right: Responses of *somewhat* or *completely acceptable* to "How acceptable would it be if [data] was accessed by the following parties?" by condition.

to the data; 29% answered no and 32% were unsure (see Figure 2). Differences between conditions were not statistically significant (per Pearson's $\chi^2$).

We asked respondents how likely it would be for various parties to gain access to the data: hackers, advertisers, financial companies, hardware manufacturers, ISPs, data aggregators, cable/satellite TV providers, law enforcement, and TV networks. There was some skepticism about some of the parties getting access, especially law enforcement and financial companies. Otherwise, respondents generally thought that advertisers, cable providers, data aggregators, and hardware manufacturers were likely to access the data (see Figure 3, left side). Comparing the proportion of affirmative responses across conditions, it appears that respondents in the personalized recommendations condition thought that third-party access to the data was more likely than the other conditions. (The margin of error for the proportions across conditions ranges from 7.2 - 10.6%, 95% CI.)

*4) How Is the Data Used?:* To probe participants' awareness that their data could be used for purposes other than making the relevant feature work, we asked three questions. We asked respondents in conditions 1–5, "If the Smart TV does upload [data type] to a server for analysis, do you expect [it/them] to be used for other purposes?" In total, only 37% of respondents answered Yes, while 40% answered No and 22%

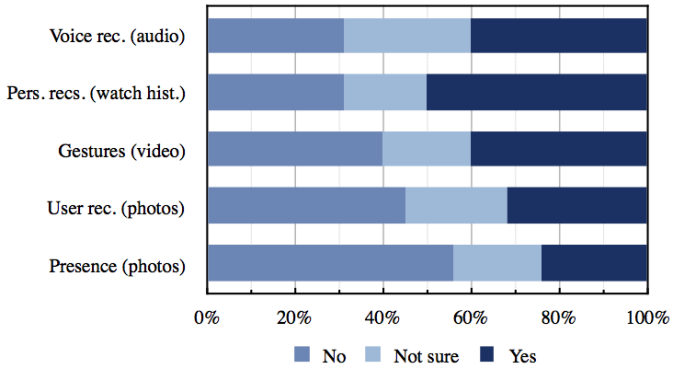


Fig. 4. Percentage of No and Yes responses to "Do you expect [the data] to be used for other purposes?" by condition.

were unsure. However, as Figure 4 shows, differences across conditions were stark (and significant; Pearson's $\chi^2(8) = 23.562$, $p = 0.0027$). A slight majority of respondents do expect their audio or video data to be repurposed, and a sizable majority expect viewing history to be repurposed. The low total of Yes responses is an effect of the two photo conditions. A combined 28% expect photo data to be repurposed, while 50% of respondents in those conditions do *not* expect photo data to be repurposed.

| Theme | % |
|---|---|
| Advertising; targeted ads; marketing | 30% |
| Sell it; share it with other parties | 10% |
| Improve TV performance, future products; training data for [feature] | 15% |
| Generate program ratings, recommendations | 5% |
| Analyze viewing habits (programs, commercials) | 12% |
| Demographic analysis; household demographics | 8% |
| Analysis; research; market research *(no details)* | 6% |
| Analyze people's possessions, spending habits | 6% |
| Analyze individual users to build a profile | 3% |
| Spying on people *(no details)* | 6% |
| Government or law enforcement spying or surveillance | 3% |
| Catch criminals, burglars; use as evidence of crimes | 2% |
| Anything; whatever the company wants | 3% |
| *Other purpose* | 10% |
| Don't know *(may be cross-coded)* | 10% |
| Nothing; no other uses | 9% |

TABLE II. SUMMARY OF CODED THEMES IN RESPONSES TO "WHAT OTHER PURPOSES MIGHT [THE DATA] BE USED FOR?" IN CONDITIONS 1–5. (PERCENTAGES DO NOT ADD UP TO 100%, AS MANY RESPONSES HAD MULTIPLE THEMES.) N=486.

We then asked participants what they thought the data might be used for (if it was repurposed). The common themes we identified are summarized in Table II (Kupper Hafner concordance: $KH = 0.697$). Most strikingly, 30% mentioned advertising or marketing of some kind—far more than any other theme we identified: "I am not sure if the photos are uploaded or not, but they could be used for targeted advertising. This is just a new form of finding things that people buy and using that information to provide advertisements to people that are more likely to buy said products." Some of the identified purposes can be viewed as immediate goals (e.g., some type of analysis) that feed into further goals: "To see if you watch certain channels or shows for advertising purposes."

Other frequent themes included using the data to improve TV performance or future products (15%): "I think it will be used to tweak the recommendation algorithm." "Understand human behavior in order to better develop products." In many such cases, especially for voice recognition, participants referred to developing training datasets: "I expect it will be added to a database and used by the program to teach it to recognize commands across more accents and speech patterns."

Many also mentioned analysis of viewing habits (12%): "Maybe it could be used for TV networks to figure out what shows are being watched." Selling or sharing with other parties was mentioned explicitly by 10%: "I think that perhaps it will be used to sell to ad companies or to show producers who are looking for deeper information into what people who watch their shows enjoy watching." Sharing was also implicit in some of the other suggested uses (such as court evidence).

Spying or surveillance by the government, law enforcement, the service provider, peeping toms, or users themselves was a common theme: "Collected by government and intelligence agencies for use as needed." "To see what other people are doing in your house." "We can be spied on for anything these days. Our TVs could be listening to our conversations with our friends. Are we god forbid terrorists? How are we going to vote? What groceries are we thinking about buying. This thing could check up on anything about us."

A mere 9% just reiterated that they did not think it would be reused; the rest of the participants came up with some possible uses even if they had said No to the previous question about the likeliness of reuse.

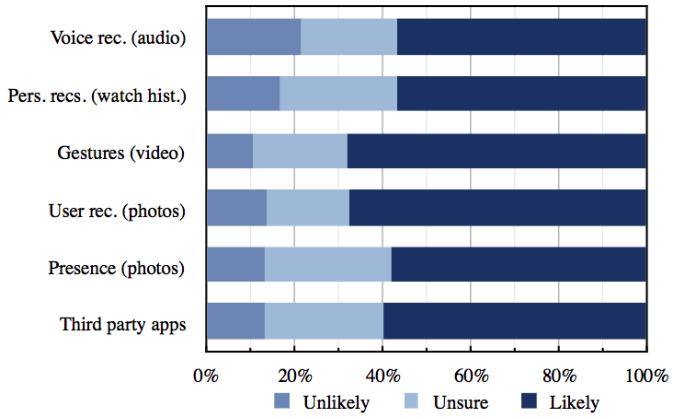


Fig. 5. Responses to "How likely is it that data...from your Smart TV will be combined...to create a detailed profile of your habits and interests?"

Finally, we asked participants in all six conditions, "How likely is it that data collected from your Smart TV will be combined with data collected from your other Internet-enabled devices (e.g., smartphone, tablet, laptop, etc.) to create a detailed profile of your habits and interests?" With this question—asking about a specific purpose, but without specifying data type—15% of participants thought it was unlikely, while 61% thought it was likely that the data would be combined in this way; see Figure 5.

Further exploration would be needed to determine whether this notable difference from the previous question about repurposing is really a matter of participants' perceptions about sharing of particular types of sensor data versus their perceptions about data-sharing in general. (Differences between conditions were not significant per Pearson's $\chi^2$, again, perhaps because the question did not ask about the specific data type.)

| Theme | % |
|---|---|
| TV malfunctioning; TV, apps stop working; viewing blocked | 42% |
| Security problem; virus; malware; getting hacked | 17% |
| Attacking other devices on home network or over the Internet | 10% |
| Showing content, ads user doesn't want | 10% |
| Taking over the TV *(no details)* | 6% |
| Credit card numbers, financial information stolen | 30% |
| Usernames, passwords stolen | 29% |
| Personal information stolen *(no details)* | 25% |
| Spying on user via the camera, other sensors | 11% |
| Tracking what user watches, what user does online | 10% |
| *Other risk* | 8% |
| Don't know *(may be cross-coded)* | 1% |
| It won't happen; no risks if it happens | 2% |

TABLE III. SUMMARY OF CODED THEMES IN RESPONSES TO "WHAT ARE SOME OF THE RISKS CAUSED BY MALICIOUS SOFTWARE...ON YOUR SMART TV?" IN CONDITION 6. N=105.

*5) Could the Data Be Stolen?:* Across conditions 1–5, roughly 45% of respondents believed a hacker might gain access to their data. (See Figure 3 for breakdown.) Condition 2 was least likely to believe a hacker would gain access to the [data] (Pearson's $\chi^2(8) = 31.276$, $p = 1.25e^{-4}$).

We asked respondents in condition 6, "By allowing new software to be installed on your Smart TV to enable third-party apps, the Smart TV may become vulnerable to malicious software being installed. How likely is it for hackers to install malicious software on your Smart TV?" Of 105 respondents, 50% believed that it was unlikely, 31% were unsure, and 19% believed that it was likely. Many of the participants were confident in their answer to this question (80% who answered Unlikely were confident and 75% who answered Likely).

We then asked, "What are some of the risks caused by malicious software running on your smart TV?". Table III lists the common themes we identified ($KH = 0.790$). Some respondents recognized the potential risk of having a TV with advanced sensors: "They might turn on my TV and get personal information. I think they might be able to turn on my game camera and microphone. It is just risky to have software that is malicious." But for most participants, this question did not seem to bring to mind the risk that a compromised TV could reveal sensor data to the attacker.[1] In particular, even where respondents mentioned information theft, most phrased it in terms of credit card numbers and account credentials: "I could have my card and bank information stolen." "The hacker may install a program that can collect information from the user like usernames and passwords."

## B. Preferences about Data Sharing and Repurposing

*1) Stated Preferences:* After asking about various parties' likelihood of gaining access to the data (see §IV-A3), we asked about the same parties, "How acceptable would it be if your [data], collected for [feature], was accessed by the following parties?"; see Figure 3 (right-hand side). Respondents showed the most acceptance for hardware manufacturers accessing the data—but still, 77% of respondents felt it would be unacceptable.

Overall, respondents felt strongly that it would be unacceptable for the data to be shared. Only in condition 2 (viewing history) were there any parties that a majority of respondents found acceptable: TV networks and cable providers. (The margin of error for acceptability proportions across conditions ranges from 1.9 - 9.4%, 95% CI.) Across all conditions, respondents were most opposed to hackers and financial companies accessing their data (an interesting pairing).

In responding to our open-ended question about what other purposes data collected by smart TVs could be used for, many participants took the opportunity to express the opinion that data *should not* be repurposed, for example: "They should not be used for other purposes, it seems like a violation of privacy." "They have no rights to use any pictures taken, they are the sole property of myself." Or only in a very limited way: "It should not be used for anything else other than the software in the TV to make it work better."

Some participants opined that repurposing should require permission, or at least be disclosed: "I hope it wouldn't use my voice without my permission." "I hope it doesn't do this unless it warns me that it does. I would not be happy if these photos were used for other purposes than what was stated." (Though others distrust disclosures: "I'm sure they'd like to say that they don't use it for anything else but who knows.")

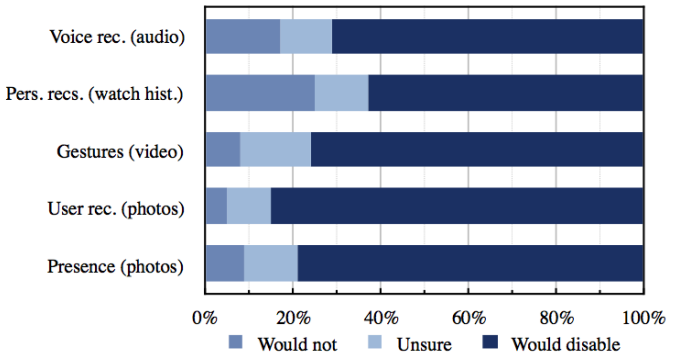


Fig. 6. Responses to "How likely would you be to disable the [feature]...to prevent your [data] from being shared with others?" by condition.

*2) Predictions about Protective Actions:* To gauge potential action based on preferences, we asked, "How likely would you be to disable the [feature] on your smart TV in order to prevent your [data] from being shared with others?" In every condition (1–5), the majority of respondents predicted they would (at least theoretically) disable the feature; see Figure 6. Sharing of watch history data was the least likely to incite action (63% would disable) and user recognition was the most likely (85%). Differences between conditions were statistically significant (Pearson's $\chi^2(8) = 27.361$, $p = 6.124e^{-4}$).

We asked if sharing would increase or decrease the likelihood that they would purchase a device from the same manufacturer in the future. Across conditions 1–5, 73% of respondents predicted they would be less likely to purchase another device from that manufacturer. Looking at individual conditions, as in Figure 7, we found that reactions differed depending on the feature (Pearson's $\chi^2(8) = 32.049$, $p = 9.126e^{-5}$). Respondents who saw the user recognition feature predicted the least willingness to purchase another device (85%). Respondents in the personalized program recommendations condition predicted the most willingness—but even there, 55% reported they would probably not purchase again.

[1] This survey condition did not specifically mention sensors anywhere.

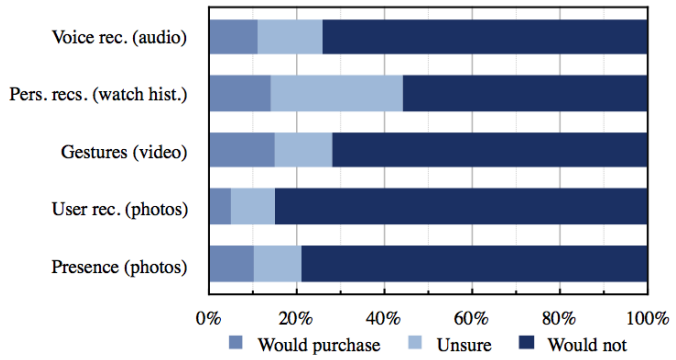


Fig. 7. Responses to "If a smart TV manufacturer shared your [data], collected for [feature], with others, would this [change] the likelihood that you would purchase...from this manufacturer again?" by condition.

### C. Beliefs about Legal Safeguards

To probe how beliefs about legal protection might affect users' attitudes and behaviors with respect to smart TVs, we asked all respondents, "Are there any laws or regulations that you believe prevent Smart TV manufacturers from sharing data collected in your home with other parties?"

| Theme | % |
|---|---|
| No; don't know of any | 35% |
| ∗Privacy laws or privacy rights prevent or constrain it | 18% |
| ∗Yes; there are laws *(no details)* | 6% |
| ∗Constitution; Fourth Amendment; right to privacy in one's home | 4% |
| ∗[Some other law] prevents or constrains it | 3% |
| ∗Wiretapping laws prevent or constrain it | 1% |
| ∗It's defined by their privacy policy/TOS; they have to abide by TOS | 9% |
| ∗They have to get users' consent | 7% |
| ∗They have to notify users | 3% |
| User waives rights by agreeing to privacy policy/TOS; TOS trumps law | 4% |
| User waives rights by buying, using TV | 2% |
| Existing laws have loopholes, are rarely enforced; companies ignore laws | 5% |
| Depends who they share it with | 3% |
| Depends on [some other factor] | 1% |
| Depends on the type of data | 1% |
| Don't know *(may be cross-coded)* | 19% |

TABLE IV. CODED THEMES IN RESPONSES TO "ARE THERE ANY LAWS OR REGULATIONS THAT...PREVENT SMART TV MANUFACTURERS FROM SHARING DATA COLLECTED IN YOUR HOME...?" FOR ALL CONDITIONS. N=591.

*1) Wide Range of Beliefs:* Table IV lists the major themes we identified in the responses ($KH = 0.652$). Many respondents (35%) said that they did not think there were any or that they did not know of any. In some of those cases, respondents referenced the lack of protection for data collected in other situations: "I doubt there would be any since there aren't any for using the Internet. They record info from our computers all the time." A sizable proportion of respondents (19%) said they simply did not know.[2]

On the other hand, many others did think their data was protected. In a few cases, participants (correctly) named laws that protect particular types of data, or mentioned laws that cover only specific situations: "I'd assume that it would be illegal in states with 2 party consent to recording, but in states without that I'm unsure." "I believe that law enforcement would need a warrant to collect data from a Smart TV manufacturer. But I suspect that commercial entities could buy it from them."

However, many respondents exhibited incorrect beliefs about laws they thought protected them.

*2) Incorrect Assumptions of Protection:* Most strikingly, 18% of respondents referred to some type of privacy laws that they assumed prevented data-sharing or at least limited or constrained it (for example, by requiring consent). In some cases, respondents simply mentioned broad "privacy laws": "Basic privacy laws that information can not be shared without our permission." "Yeah privacy laws." Other responses were more cautious: "I think there probably are some related to privacy of material given to public companies/people without our consent but I can't name them specifically."

Others respondents were more specific, but still incorrect—for example, citing the U.S. Constitution. "The Constitution protects against unlawful searches and I believe this would fall under that category." In many cases, these beliefs were an extension of incorrect assumptions that such laws cover other types of consumer data: "I would imagine there are laws just like the sharing of your email or phone number." A few responses demonstrated some misunderstanding about how the Internet is governed: "I would hope as an Internet based device, the laws of the Internet would apply just as equal [*sic*]."

Some participants had faith in the power of privacy policies and terms of service. For example, 9% of respondents believed these notices have legal standing to constrain companies' behavior (i.e., that terms of service are the main legal determinant of what companies can share). "I don't know other than their specific privacy policy that they send out regarding how they use your info." "I don't think there are any laws other then that if it states in the contract that they do not that [*sic*] they must adhere to that." Many respondents believed that permission or at least disclosure is required: "I assume they have to say something in the user agreement." (The actual situation is rather more complicated; see §VI-B.)

While some of those participants framed the legal force of privacy policies as protecting consumers, about 4% expressed a different perspective: "They'll just put it in their user agreement so you'll have to sign away your privacy rights." "I suppose all they'd have to do is have a terms of condition you needed to accept stating that they would be able to do this and then you'd have no choice but to agree so that you could use the TV or the third party app." Such comments about lack of choice echo previous findings about why consumers use privacy-compromising services; see discussion in §V-C.

The belief that terms of service have legal standing isn't entirely incorrect, but most end user agreements are written to be intentionally vague and unrestrictive. A number of participants commented on this, as well as pointing out that few people read those agreements: "The way companies word their agreements, they trick you into allowing them to break the privacy laws." "The ability to opt out is always embedded deep in some fine print that people don't take the time to read."

*3) Correlations and Consequences of Beliefs about Legal Protections:* We grouped the themes shown in Table IV into broad categories, separating responses that expressed the belief

[2]For all open-ended questions, we only coded responses as "don't know" if participants expressed their lack of knowledge explicitly, strongly, and separately from any speculation. We did not code cases where participants were simply expressing lack of certainty about their answer.

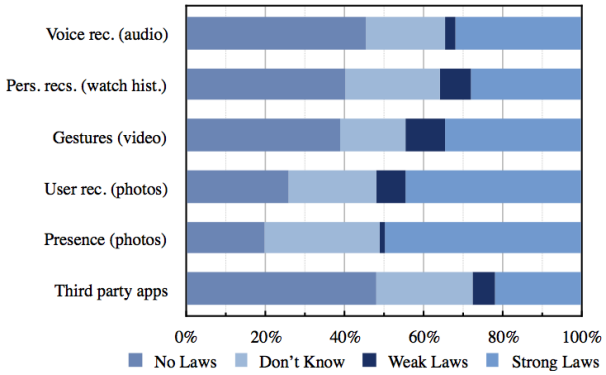


Fig. 8. Broad categories of responses to "Are there any laws or regulations that...prevent smart TV manufacturers from sharing data collected in your home...?" by condition. N=486.

that there are strong legal protections that prevent or constrain data-sharing (i.e., the themes marked with * in Table IV) from other types of responses (beliefs that there aren't any laws, beliefs that the laws/protections are weak or partial, or simply not knowing). There were 486 participants whose responses could be clearly categorized in this way—with 35% believing in strong legal protections of some kind. However, this result varied across conditions.[3] Participants in the conditions involving photos (C4 and C5) were most likely to think there are strong legal protections, and participants in the conditions involving viewing history (C2) and third-party app data (C6) were most likely to be skeptical ($p = 0.00167$, Pearson's $\chi^2$).[4] (See Figure 8.)

These groupings allowed us to investigate how incorrect beliefs about privacy laws can affect the concrete assumptions people make about what's happening with their data.

Among the participants whose responses we could clearly categorize, we found significant correlations between beliefs about privacy laws and understandings about data repurposing and sharing. Among participants who believed in strong legal protections against data-sharing, 28% considered it *unlikely* or *completely unlikely* that advertisers would have access to data collected by their smart TV, and 26% considered access by data aggregators unlikely (see §IV-A3). Among the participants who did *not* believe there are strong protections, those figures were 17% and 16%, respectively ($p = 0.00375$ and $p = 0.00671$).

Similarly, 44% of participants who believe themselves protected against data-sharing did not think that the data collected by the smart TV would be used for other purposes (see §IV-A4), as compared to 31% of participants who do not believe in such protections ($p = 0.00430$). (A similar effect was seen for the question about data being combined to make a profile, but it was not statistically significant.)

*4) General Feeling That Laws Should Be More Protective:* While the question asked about existing laws, many respondents took the opportunity to share the view that even if there aren't currently laws protecting them from data-sharing, they ought to exist: "I am not aware of any such laws but I feel there should be." "I do not know if there are any regulations but there should be. They shouldn't be allowed to send our personal information to others through what we watch." "I would hope there is something that prevents them from selling or giving away photos of the inside of people's houses. It is terrifying to think that anyone could see that." (Such responses were more frequent than responses saying there should not be any such laws by a factor of ten.)

Some respondents critiqued the current regulatory system as being too permissive. "I think there are some laws that prevent Smart TV manufacturers from sharing that kind of data but they are not that strict. They can freely do it behind our backs anyway." "I am sure there are, but I doubt these regulations are strictly monitored." Others simply took the opportunity to share concerns about data privacy in the current landscape: "I don't know how the laws work but I do know it is scary having a smart TV because they are so invasive." "I don't know the current laws, but now I am concerned after reading this! My kids watch the TV quite frequently and I don't like the idea of advertisers or others knowing anything about my kids. I want their ultimate privacy maintained."

On the whole, we found that a worryingly high proportion of participants believed that some type of privacy laws prevented companies from sharing their data with third parties. On the other hand, many others realize that data may be shared—but even though they may be opposed to it, they don't believe they can do anything about it.

## V. Discussion

In many ways, our results are consistent with the findings of prior surveys on privacy attitudes and behaviors. We find that people's expectations and opinions about data-sharing depend on context. We also find that, despite strong opinions about data sharing, people will continue to welcome new technology. At the same time, our data indicates some level of resignation, as people reported that certain sharing relationships were likely to happen even if they felt they were completely unacceptable.

### A. Expectations Based on Contextual Integrity

The theory of privacy as contextual integrity [3], [41] suggests that privacy violations happen when data flows counter to expectations and that people form their expectations based on the context in which the data was originally collected. Our participants were clear that they would find it unacceptable if smart TV data were shared with third parties or if the data was combined. However, many of them also held the notion that other parties accessing the data was inevitable. This was especially true for respondents in the personalized recommendations condition (C2) and to a lesser extent the voice recognition (C1) and gesture recognition (C3) conditions.

We know from prior work, and other fields, that self-reported predictions about future behavior are often inaccurate; however, it is worth noting that our respondents did indicate they would be willing to disable these features or avoid the manufacturer in the future if the data were shared (more than 60% reported they would disable the feature).

[3]The question asked generally about "data collected in your home", but responses may reflect the data type that was the focus of the condition because of question order (see the Appendix).

[4]Correlations in this section opposed responses that expressed a belief in strong legal protections against the other three response types grouped together.

### B. Developing Privacy Preferences with New Tech

When it comes to end-user privacy preferences, we often find there is no simple answer. To quote a 2016 Pew study, "These findings suggest that the phrase that best captures Americans views on the choice between privacy vs. disclosure of personal information is, 'It depends.' " [42]. People will choose to divulge personal information based on the benefit they stand to receive and attempt to evaluate the potential risks. But determining the potential risk is often much more difficult than determining the benefit.

Today a device that looks, and functions, more or less the same as an older television set—and that seems at first glance to have the same affordances—suddenly has new capabilities that may have privacy implications; cf. [39]. People are accustomed to watching and listening to a TV, but it's new territory to have a TV that watches and listens to you.

Formulating concerns about data collection and use requires that the user understand what is being collected, where it is going, and how it might be used. It is challenging to form an accurate mental model of new technology when data collection is not obvious and privacy notices are lacking [19]. Perhaps smart TVs would be less popular if people had a more accurate notion of what data was being collected and how it was being used—or if they did not incorrectly believe they were protected.

Looking at underlying technical knowledge about smart devices, Wilson, Hargreaves, and Hauxwell-Baldwin [52] also found a lack of consensus (or even individual consistency) about privacy-relevant features; a slight majority agreed with the statement that they were always on, but also with the statement that they only came on when activated. Zeng, Mare, and Roesner [54] found in one-on-one interviews that even people with highly equipped smart homes often have impoverished or inaccurate models of how the technology works (especially if it was personally not their idea to outfit the home)—resulting in impoverished or inaccurate understandings of security and privacy threats. Their participants often borrowed inapt models and mitigation strategies from computers and phones. Looking at self-evaluation, a majority of respondents in Groopman and Etlinger's survey [22] did not believe they had a good understanding of how companies use data collected from smart devices.

Similarly, our findings indicate two large areas of confusion that must be addressed if users are expected to self-moderate the data collection of a smart TV. First, smart TV owners need a better understanding of what data is being collected by the unit and how the data is used, including whether it is sent to third parties. Second, smart TV owners need a better understanding of how data uses or sharing practices are (or aren't) restricted by current laws, so that they understand what their personal responsibilities are to protect their privacy.

Interestingly, we did find that our respondents' expectations differed between types of data about what gets shared—and that they responded differently to the prospect of data being shared in different contexts. For example, people expected that their viewing history will be used in more ways and shared with more parties, whereas they expected less would be done with photos, audio, and video data. And at the same time, they found sharing of viewing history more acceptable. In that the photos, audio, and video were considered different, and perhaps more sensitive, our findings are similar to those reported by Lee *et al.* [33] and Naeini *et al.* [40].

It is particularly illuminating to compare variation across conditions for our question about legal protections (see §IV-C3) with the questions about the acceptability of data being accessed by various third parties (see §IV-B1): participants in the viewing-history condition were both most accepting of data-sharing and least likely to assume strong legal protections against it, while participants in the two conditions involving photos were both *least accepting of data-sharing* and *most likely to assume strong legal protections* against it. While further research is needed to confirm the connection, our results suggest that, where participants object most strongly to data-sharing, they are quickest to assume there must be a law against it.

### C. The Resignation Factor

We see an interesting effect when we compare responses about the likelihood of third parties accessing the data versus the acceptability of those same parties having that data (see Figure 3): respondents find it unacceptable for these parties to have access, but they also regard this access as likely. Furthermore, the exact purpose of data sharing is often opaque to users; as one participant expressed, "what *can't* data be used for?"

There are a few possible explanations for this effect, and additional research is needed to more fully understand it. As others have noted, people *are* learning that their data gets shared and repurposed [36], [50], but they don't necessarily condone it and they don't feel like they have control over how much information is collected or how it is used [36].

To further complicate matters, the vast majority of TVs currently on the market are smart TVs and people find (at least some of) the advanced features to be quite useful. People may therefore not feel like they have a real choice not to have their privacy invaded; cf. [15].

## VI. Opportunities to Enhance End-User Privacy

Our findings improve our understanding of what smart TV owners currently believe to be true about data collection and use; they also demonstrate users' expectations of how the data should be used and protected. Our results clearly indicate that people do not want their data to be repurposed or shared. Our findings are relevant to at least three audiences who are in a position to impact end-user privacy: manufacturers of smart TVs, regulators and lawmakers, and designers of privacy-enhancing technologies.

### A. Manufacturers of Smart TVs

We suspect that most smart TVs do not adequately signal operation of the advanced features. Our results indicate that users lack understanding of how data may be used to support each feature. This presents an opportunity for TV manufacturers to improve disclosures about the data collected and the ways that the data is used [21]. Internet of Things devices could benefit from typical recommendations on applying privacy by design principles [6] or improved notice and consent [45],

[44]. There have also been efforts to standardize notice and consent—in addition to improving technical controls and safeguards—by a cohort of privacy rights groups [10].

Smart TV manufacturers could also follow the guidance emerging from research on sensors found in smart-homes, for example designing signals that clearly communicate the collection and transmission of sensor data [54], [30], [28], [7], [21]. Our study reinforces prior findings that users' expectations vary according to different types of data; we therefore support recommendations to consider how to communicate that different types of data are being collected, when applicable.

Another way that smart devices in the home can mitigate potential privacy concerns is to minimize the amount of data sent off the device in the first place [21]. For example, Apple's Touch ID is stored on the device, mitigating concerns about Apple amassing a biometric database of users.

Our respondents were clear: they do not want the data from their smart TVs repurposed or shared. Smart TV manufacturers are in the best position to be the gatekeepers and to appropriately limit data access (*cf.* [23]). Potential mitigations could include using *privacy mediators* (such as those proposed by Davies et al. [14]) or deploying a third-party application platform whose design prioritizes appropriate limitations and strict safeguards.

### B. Regulators and Lawmakers

As we noted in §II, previous research has found that many people make incorrect assumptions or are unsure about legal protections on their personal data, which affects whether they actively take measures to protect it [25], [49], [50]. Similarly, in our survey, we found that most people either aren't sure how they are protected by current laws, or they assume they are—and that this plays out in specific assumptions about whether smart TV data is actually likely to be used or shared. It is also noteworthy that many respondents took our question about existing laws as an opportunity to air opinions about how there *should* be laws about the collection and use of personal data.

It is difficult, if not impossible, for end users to take full responsibility for managing their personal data and the devices that want to use personal data [47], [53], [23]. The benefits of new technology are more salient than the risks, and people find themselves adopting new technology hopeful that existing laws and regulations protect them, when in reality they are more limited than people realize [24], [50]. Even for a very informed user, it would be difficult to keep up with the current legal situation. What relevant laws there are (such as two-party consent to recording and biometric identification laws) vary from state to state even within the U.S., and their application to smart devices is still being contested (for example, whether facial recognition counts as a biometric) [50], [13], [11], [21].

Surveys like this one help to reveal how large the gap is between actual protections and beliefs. The fact that so many people trust that policymakers already *do* play an important role in protecting them suggests that policy must have an significant role in constraining data sharing and use.

### C. Designers of Privacy-Enhancing Technologies

Even if manufacturers and lawmakers act on this advice, there will be a need for new privacy-enhancing technologies. A smart TV may be one of many smart devices in a home, so it's not enough for a user to understand the TV's capabilities; they must also understand the interactions and implications of data from a number of devices. For years, experts have warned about a "mosaic effect" [17], in which new insights emerge from combining multiple datasets and sensors.

More third-party tools are needed to help people manage the digital exhaust generated by a smart home. (Aleisa and Renaud [1] provide a review of what is currently available.) Users could be notified when potentially sensitive data is being collected (e.g., [11], [39]). A third-party tool could also recommend configurations for individual devices based on the owner's preferences (e.g., [40]).

## VII. Concluding Remarks

While smart TVs are just one example of an Internet of Things device, they are the most common smart device in households today. As people welcome additional connected devices into their homes, we will see an increased need to help users understand the data collected. We will also see increased pressure to offer better data management tools and technically sound frameworks for public policy. Additional research is needed to investigate how to help consumers develop more accurate mental models of how these technologies work, in addition to research on protective tools and privacy-enhancing technologies.

Some potential research questions include: How do people use their understanding of older tech to understand smart devices with sensors? What analogies do they use (*cf.* [39])? Do people think about their smart devices as being part of the Internet (with all the same attendant privacy issues) or do they think of them as a separate phenomenon (with new privacy issues)? In what ways do privacy protection tools and frameworks need to be customizable to individual users and in what ways can they be the same across users?

At the time when we collected the survey data, not all of the features were common (yet), and respondents may not have had an experience with the features that the survey focused on, so a follow-up study may be useful once these features become more common to better understand how understanding evolves with exposure. A better understanding of the privacy threats and privacy-preserving behaviors employed would be gained by looking at the actual devices, models, and configurations, which were out of scope for this survey. A follow-up study could also investigate participants' expectations and beliefs vis-à-vis the smart devices and features that are actually in their home.

## Acknowledgments

This research was supported by the U.S. National Science Foundation grant, "Security and Privacy for Wearable and Continuous Sensing Platforms" (CNS-1514211). (Opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the NSF.) Our thanks to the panelists and audience at the U.S. Federal Trade Commission's workshop on smart TVs, held on December 7, 2016, where the last author presented a preliminary version of these findings.

## REFERENCES


[1] N. Aleisa and K. Renaud, "Privacy of the Internet of Things: A systematic literature review," in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017, extended version available at https://arxiv.org/abs/1611.03340. [Online]. Available: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1733&context=hicss-50

[2] ——, "Yes, I know this IoT device might invade my privacy, but I love it anyway! a study of Saudi Arabian perceptions," in *2nd International Conference on Internet of Things: Big Data and Security (IoTBDS 2017)*, 2017, pp. 198–205. [Online]. Available: http://eprints.gla.ac.uk/145777/

[3] A. Barth, A. Datta, J. C. Mitchell, and H. Nissenbaum, "Privacy and contextual integrity: Framework and applications," in *Proceedings of the 2006 IEEE Symposium on Security and Privacy (SP '06)*. Washington, DC: IEEE Computer Society, 2006, pp. 184–198. [Online]. Available: https://doi.org/10.1109/SP.2006.32

[4] R. Binns, J. Zhao, M. Van Kleek, N. Shadbolt, I. Liccardi, and D. Weitzner, "My bank already gets this data: Exposure minimisation and company relationships in privacy decision-making," in *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI '17), Extended Abstracts*. New York: ACM, 2017, pp. 2403–2409. [Online]. Available: http://doi.acm.org/10.1145/3027063.3053255

[5] J. Bugeja, A. Jacobsson, and P. Davidsson, "On privacy and security challenges in smart connected homes," in *Proceedings of the European Intelligence and Security Informatics Conference (EISIC)*, Aug 2016, pp. 172–175. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7870217

[6] A. Cavoukian and C. Popa, "Embedding privacy into what's next: Privacy by Design for the Internet of Things," April 2016, accessed: 17 February 2018. [Online]. Available: https://www.ryerson.ca/content/dam/pbdce/papers/Privacy-by-Design-for-the-Internet-of-Things.pdf

[7] E. K. Choe, S. Consolvo, J. Jung, B. Harrison, S. N. Patel, and J. A. Kientz, "Investigating receptiveness to sensing and inference in the home using sensor proxies," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. New York: ACM, 2012, pp. 61–70. [Online]. Available: http://doi.acm.org/10.1145/2370216.2370226

[8] Consumer Reports, "Consumer reports TV buying guide: Getting the right TV," October 2017, accessed: 15 February 2018. [Online]. Available: https://www.consumerreports.org/cro/tvs/buying-guide/index.htm

[9] ——, "Samsung and Roku smart TVs vulnerable to hacking, Consumer Reports finds," *Consumer Reports*, February 2018, accessed: 9 February 2018. [Online]. Available: https://www.consumerreports.org/televisions/samsung-roku-smart-tvs-vulnerable-to-hacking-consumer-reports-finds/

[10] Consumer Reports, Disconnect, Ranking Digital Rights, The Cyber Independent Testing Lab, Aspiration, and Philanthropic Partners, "The digital standard," accessed: 16 February 2018. [Online]. Available: https://www.thedigitalstandard.org/the-standard

[11] A. Das, M. Degeling, X. Wang, J. Wang, N. Sadeh, and M. Satyanarayanan, "Assisting users in a world full of cameras: A privacy-aware infrastructure for computer vision applications," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. IEEE, July 2017, pp. 1387–1396. [Online]. Available: http://openaccess.thecvf.com/content_cvpr_2017_workshops/w16/papers/Satyanarayanan_Assisting_Users_in_CVPR_2017_paper.pdf

[12] S. Das and A. Kramer, "Self-Censorship on Facebook," in *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 2013, pp. 120–127.

[13] N. Davenport, "Smart washers may clean your clothes, but hacks can clean out your privacy, and underdeveloped regulations could leave you hanging on a line," *John Marshall Journal of Information Technology & Privacy Law*, vol. 32, pp. 259–297, 2015. [Online]. Available: https://repository.jmls.edu/jitpl/vol32/iss4/2/

[14] N. Davies, N. Taft, M. Satyanarayanan, S. Clinch, and B. Amos, "Privacy mediators: Helping IoT cross the chasm," in *Proceedings of the 17th International Workshop on Mobile Computing Systems and Applications (HotMobile '16)*. New York: ACM, 2016, pp. 39–44. [Online]. Available: http://doi.acm.org/10.1145/2873587.2873600

[15] S. Egelman, A. P. Felt, and D. Wagner, *Choice Architecture and Smartphone Privacy: There's a Price for That*. Berlin & Heidelberg: Springer, 2013, pp. 211–236. [Online]. Available: https://doi.org/10.1007/978-3-642-39498-0_10

[16] A. P. Felt, S. Egelman, and D. Wagner, "I've got 99 problems, but vibration ain't one: A survey of smartphone users' concerns," in *Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM '12)*. New York: ACM, 2012, pp. 33–44. [Online]. Available: https://dl.acm.org/citation.cfm?id=2381943

[17] E. Felten, "Defining privacy: Hearings before the privacy and civil liberties oversight board," 2014, accessed: 6 March 2018. [Online]. Available: https://www.c-span.org/video/?322698-1/discussion-defining-privacy

[18] E. Fernandes, J. Jung, and A. Prakash, "Security analysis of emerging smart home applications," in *2016 IEEE Symposium on Security and Privacy (SP)*, May 2016, pp. 636–654.

[19] FTC Staff, "Internet of things, FTC staff report," January 2015. [Online]. Available: https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf

[20] M. Ghiglieri, M. Volkamer, and K. Renaud, "Exploring consumers' attitudes of smart TV related privacy risks," in *Proceedings of the 5th International Conference on Human Aspects of Information Security, Privacy, and Trust (HAS)*, ser. Lecture Notes in Computer Science, T. Tryfonas, Ed. Cham: Springer, 2017, pp. 656–674.

[21] S. Gray, "Always on: Privacy implications of microphone-enabled devices," Future of Privacy Forum, Tech. Rep., April 2016. [Online]. Available: https://www.ftc.gov/system/files/documents/public_comments/2016/08/00003-128652.pdf

[22] J. Groopman and S. Etlinger, "Consumer perceptions of privacy in the Internet of Things: What brands can learn from a concerned citizenry," Altimeter Group, June 2015, accessed: 17 February 2018. [Online]. Available: http://www.altimetergroup.com/pdf/reports/Consumer-Perceptions-Privacy-IoT-Altimeter-Group.pdf

[23] J. Hong, "The privacy landscape of pervasive computing," *IEEE Pervasive Computing*, vol. 16, no. 3, pp. 40–48, 2017. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7994573/

[24] C. J. Hoofnagle, J. King, S. Li, and J. Turow, "How different are young adults from older adults when it comes to information privacy attitudes and policies?" *SSRN eLibrary*, 2010.

[25] C. J. Hoofnagle and J. King, "What Californians understand about privacy offline," May 2008, available via Social Science Research Network. Accessed: 3 June 2015. [Online]. Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1133075

[26] ——, "What Californians understand about privacy online," September 2008, available via Social Science Research Network. Accessed: 25 February 2018. [Online]. Available: http://ssrn.com/abstract=126213

[27] M. Johnson, S. Egelman, and S. M. Bellovin, "Facebook and privacy: It's complicated," in *Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS '12)*, ser. SOUPS. New York, NY: ACM, 2012, pp. 9:1–9:15. [Online]. Available: http://doi.acm.org/10.1145/2335356.2335369

[28] S. Kazi, O. Kohanteb, T. Saensuksopa, O. Tong, , and H. Yang, "Decoding sensors: Creating guidelines for designing connected devices," Summer 2015, accessed: 7 March 2018. [Online]. Available: http://www.signifiers.io/summer.pdf

[29] P. Klasnja, S. Consolvo, J. Jung, B. M. Greenstein, L. LeGrand, P. Powledge, and D. Wetherall, ""When I am on wi-fi, I am fearless": Privacy concerns & practices in everyday wi-fi use," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '09. New York, NY, USA: ACM, 2009, pp. 1993–2002. [Online]. Available: http://doi.acm.org/10.1145/1518701.1519004

[30] O. Kohanteb, O. Tong, H. Yang, T. Saensuksopa, and S. Kazi, "signifiers.io guidelines for designing connected devices," 2015, web page. Accessed: 26 February 2018. [Online]. Available: http://signifiers.io/signifiers/guidelines.html

[31] S. Kokolakis, "Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon," *Computers & Security*, vol. 64, pp. 122 – 134, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167404815001017

[32] L. L. Kupper and K. B. Hafner, "On assessing interrater agreement for

multiple attribute responses," *Biometrics*, vol. 45, no. 3, p. 957, Sep 1989.

[33] L. Lee, J. H. Lee, S. Egelman, and D. Wagner, "Information disclosure concerns in the age of wearable computing," in *Proceedings of the NDSS Workshop on Usable Security (USEC '16)*. Internet Society, 2016. [Online]. Available: http://wp.internetsociety.org/ndss/wp-content/uploads/sites/25/2017/09/information-disclosure-concerns-in-the-age-of-wearable-computing.pdf

[34] Y. Liu, K. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing Facebook privacy settings: User expectations vs. reality," in *Proc. of Internet Measurement Conference (IMC)*. ACM, 2011.

[35] J. Lynch, "Nearly 70 million U.S. households now have a connected-TV streaming device," *Adweek*, November 16 2017, accessed: 15 February 2018. [Online]. Available: http://www.adweek.com/tv-video/nearly-70-million-u-s-households-now-have-a-connected-tv-streaming-device/

[36] M. Madden and L. Rainie, "Americans' attitudes about privacy, security, and surveillance," May 2015, accessed: 9 February 2018. [Online]. Available: http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

[37] M. Madejski, M. Johnson, and S. M. Bellovin, "A study of privacy settings errors in an online social network," in *Proceedings of the 4th IEEE International Workshop on Security and Social Networking*, ser. SESOC '12, 2012.

[38] B. Michéle, *Smart TV Security: Media Playback and Digital Video Broadcast*. Cham: Springer, 2015. [Online]. Available: https://doi.org/10.1007/978-3-319-20994-4

[39] A. Montanari, A. Mashhadi, A. Mathur, and F. Kawsar, "Understanding the privacy design space for personal connected objects," in *Proceedings of the 30th International BCS Human Computer Interaction Conference: Fusion! (HCI '16)*. Swindon, UK: BCS Learning & Development Ltd., 2016, pp. 18:1–18:13. [Online]. Available: https://doi.org/10.14236/ewic/HCI2016.18

[40] P. E. Naeini, S. Bhagavatula, H. Habib, M. Degeling, L. Bauer, L. F. Cranor, and N. Sadeh, "Privacy expectations and preferences in an IoT world," in *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. Santa Clara, CA: USENIX Association, 2017, pp. 399–412. [Online]. Available: https://www.usenix.org/conference/soups2017/technical-sessions/presentation/naeini

[41] H. Nissenbaum, "A contextual approach to privacy online," *Daedalus*, vol. 140, no. 4, pp. 32–48, Fall 2011.

[42] L. Rainie and M. Duggan, "Privacy and information sharing," January 2016, accessed: 7 March 2018. [Online]. Available: http://www.pewinternet.org/2016/01/14/privacy-and-information-sharing/

[43] R. L. Rutledge, A. K. Massey, and A. I. Antón, "Privacy impacts of IoT devices: A smartTV case study," in *Proceedings of the IEEE 24th International Requirements Engineering Conference Workshops (REW)*, Sept 2016, pp. 261–270. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7815633/

[44] F. Schaub, R. Balebako, and L. F. Cranor, "Designing effective privacy notices and controls," *IEEE Internet Computing*, vol. 21, no. 3, pp. 70–77, May 2017. [Online]. Available: https://doi.org/10.1109/MIC.2017.75

[45] F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor, "A design space for effective privacy notices," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. Ottawa: USENIX Association, 2015, pp. 1–17. [Online]. Available: https://www.usenix.org/conference/soups2015/proceedings/presentation/schaub

[46] M. Sleeper, R. Balebako, S. Das, A. L. McConahy, J. Wiese, and L. F. Cranor, "The post that wasn't: Exploring self-censorship on facebook," in *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, ser. CSCW '13. New York, NY, USA: ACM, 2013, pp. 793–802. [Online]. Available: http://doi.acm.org/10.1145/2441776.2441865

[47] D. J. Solove, "Privacy self-management and the consent dilemma," *Harvard Law Review*, vol. 1880, 2013.

[48] M. Stevens, "History of television," in *Grolier Multimedia Encyclopedia*, 2000, accessed: 3 March 2018. [Online]. Available: https://www.nyu.edu/classes/stephens/History%20of%20Television%20page.htm

[49] J. Turow, L. Feldman, and K. Meltzer, "Open to exploitation: America's shoppers online and offline," Annenberg Public Policy Center of the University of Pennsylvania, Tech. Rep., June 2005, accessed: 3 June 2015. [Online]. Available: https://repository.upenn.edu/asc_papers/35

[50] J. Turow, M. Hennessy, and N. Draper, "The tradeoff fallacy: How marketers are misrepresenting American consumers and opening them up to exploitation," Annenberg Public Policy Center of the University of Pennsylvania, Tech. Rep., June 2015, accessed: 24 February 2018. [Online]. Available: https://ssrn.com/abstract=2820060

[51] M. Williams, J. R. C. Nurse, and S. Creese, ""privacy is the boring bit": User perceptions and behaviour in the internet-of-things," in *15th International Conference on Privacy, Security, and Trust (PST 2017)*, 08 2017.

[52] C. Wilson, T. Hargreaves, and R. Hauxwell-Baldwin, "Smart homes and their users: A systematic analysis and key challenges," *Personal and Ubiquitous Computing*, vol. 19, no. 2, pp. 463–476, Feb. 2015. [Online]. Available: http://dx.doi.org/10.1007/s00779-014-0813-0

[53] ——, "Benefits and risks of smart home technologies," *Energy Policy*, vol. 103, pp. 72–83, April 2017. [Online]. Available: https://doi.org/10.1016/j.enpol.2016.12.047

[54] E. Zeng, S. Mare, and F. Roesner, "End user security and privacy concerns with smart homes," in *Proceedings of the Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. Santa Clara, CA: USENIX Association, 2017, pp. 65–80. [Online]. Available: https://www.usenix.org/conference/soups2017/technical-sessions/presentation/zeng

## APPENDIX

### A. Conditions

The survey had six conditions. Conditions 1-5 were designed to directly compare privacy perceptions based on the data type and use to support a specific feature: (1) audio to support voice recognition, (2) viewing history to support personalized recommendations, (3) video to support controlling the TV via gesture, (4) photos to support user recognition, and (5) photos to support presence detection. Condition 6 diverged from the other conditions to ask similar questions about third-party applications and code on a smart TV. The following feature descriptions were given to participants in the respective conditions.

**Condition 1: voice recognition** Some Smart TVs feature voice recognition. Voice recognition allows the user to speak various commands instead of using a remote control. For example, to change channels, search for content, or even turn the Smart TV on or off.

**Condition 2: personalized recommendations** Some Smart TVs feature personalized recommendations. Based on your viewing habits, Smart TVs might recommend new shows to watch that you might also like. This feature makes it easy to discover new content that you might otherwise not be aware of.

**Condition 3: gesture recognition** Some Smart TVs feature gesture recognition. This feature eliminates the need to use a remote control, by allowing you to use only your hands. By making hand gestures, you can navigate onscreen menus, change channels, adjust the volume, and more. More importantly, you do not need to worry about finding the remote control.

**Condition 4: user recognition** Some Smart TVs feature user recognition. This feature allows the Smart TV to differentiate between members of the same household. This allows the Smart TV to offer personalized recommendations for each viewer.

**Condition 5: presence detection** Some Smart TVs feature presence detection. Presence detection means that the Smart TV will know when someone enters or leaves the room. For instance, the TV may automatically turn on when someone sits down in front of it. This can also save on power, by automatically turning the TV off when no one is present.

**Condition 6: third-party apps** Some Smart TVs feature third-party apps. Different content providers (e.g., HBO, Netflix, Amazon, etc.) may offer their own apps that can be downloaded and installed on the Smart TV. This means that new features can be added without requiring the user to buy a new TV, as well as to offer access to new content providers as they become available.

### B. Survey Instrument for Condition 1

*Except where noted, Conditions 1–5 were very similar, varying only in the underlined portions of questions. Contact the authors for a copy of the full survey for all conditions.*

This survey is about your attitudes surrounding Smart TVs. Please read the following information below, that gives background information on Smart TVs. Please pay close attention, as there will be several questions about what you just read.

*What is a "Smart TV"?* A Smart TV is a television set that is connected to the Internet. This allows you to view content from Internet-based providers, such as YouTube or Netflix.

1) Is there a smart TV in your home?
   - Yes
   - No, but I'm considering buying one
   - No
   - I'm not sure
2) What brand of smart TV is it?
3) Approximately when was this Smart TV acquired?
4) Using the scale below, how useful a feature is having voice recognition on a Smart TV?
   - Very useful
   - Useful
   - Unsure
   - Not useful
   - Not at all useful
5) When performing voice recognition, the Smart TV must record audio of you giving it commands. Do you believe that it analyzes this audio to detect commands on the TV itself, or does it transmit the audio across the Internet to be analyzed elsewhere?
6) How confident are you of your previous answer?
   - Very confident
   - Confident
   - Neutral
   - Unconfident
   - Very unconfident
7) *Condition 1:* When waiting to accept commands for voice recognition, does the Smart TV constantly record all audio in the room, does it only listen for specific commands, or does it do something else?
   *Condition 2:* When analyzing your viewing history to make personalized recommendations, does the Smart TV examine all the shows that you have watched, does it only check to see if you have watched specific shows, or does it do something else?
   *Condition 4:* When waiting to perform user recognition, does the Smart TV constantly take photos of the room, does it only take photos under certain conditions, or does it do something else?
   Please explain to the best of your ability:
8) If the Smart TV does upload recorded audio to a server for analysis, do you expect it to be used for other purposes?
   - Yes
   - No
   - I'm not sure
9) What other purposes might it be used for?
10) *Condition 1:* Do you believe that a human being will have access to in-home audio recorded by your Smart TV?
    *Condition 2:* Do you believe that a human being will have access to your viewing history?
    - Yes
    - No
    - I'm not sure
11) How likely is it that in-home audio, recorded for voice recognition purposes, will be accessed by the following parties: Hardware manufacturers (e.g., Samsung, LG, Sony, etc.), ISPs, Cable/Satelite TV Providers, Financial companies, Law enforcement, TV networks, Hackers, Data aggregators (i.e., companies that resell consumer data), and Advertisers?
    - Completely likely
    - Somewhat likely
    - Neutral
    - Unlikely
    - Completely unlikely
12) How acceptable would it be if in-home audio, recorded for voice recognition purposes, was accessed by the following parties? *[parties same as previous question]*
    - Completely acceptable
    - Somewhat acceptable
    - Neutral
    - Unacceptable
    - Completely unacceptable
13) Are there any laws or regulations that you believe prevent Smart TV manufacturers from sharing data collected in your home with other parties? Please explain to the best of your ability.
14) How likely would you be to disable the voice recognition feature on your Smart TV in order to prevent audio recorded in your home from being shared with others?
    - Definitely would not disable
    - Might not disable

- Unsure
- Might disable
- Definitely would disable

15) If a Smart TV manufacturer shared in-home audio, collected for voice recognition purposes, with others, would this increase or decrease the likelihood that you would purchase a device from this manufacturer again?
    - Definitely would not purchase
    - Might not purchase
    - Unsure
    - Might purchase
    - Definitely would purchase

16) How likely is it that data collected from your Smart TV will be combined with data collected from your other Internet-enabled devices (e.g., smartphone, tablet, laptop, etc.) to create a detailed profile of your habits and interests?
    - Extremely likely
    - Likely
    - Neutral
    - Unlikely
    - Extremely unlikely

17) Thinking about the questions in this survey, have you thought about these issues before? Please explain:

18) In what year were you born?

19) What is your gender?

20) What is your highest level of education?

### C. Survey Instrument for Condition 6

This survey is about your attitudes surrounding Smart TVs. Please read the following information below, that gives background information on Smart TVs. Please pay close attention, as there will be several questions about what you just read.

*What is a "Smart TV"?* A Smart TV is a television set that is connected to the Internet. This allows you to view content from Internet-based providers, such as YouTube or Netflix.

1) Is there a smart TV in your home?
    - Yes
    - No, but I'm considering buying one
    - No
    - I'm not sure

2) What brand of smart TV is it?

3) Approximately when was this Smart TV acquired?

4) Using the scale below, how useful a feature is having third-party apps on a Smart TV?
    - Very useful
    - Useful
    - Unsure
    - Not useful
    - Not at all useful

5) By allowing new software to be installed on your Smart TV to enable third-party apps, the Smart TV may become vulnerable to malicious software being installed. How likely is it for a hacker to install malicious software on your Smart TV?
    - Very likely
    - Likely
    - Unsure
    - Unlikely
    - Very unlikely

6) How confident are you of your previous answer?
    - Very confident
    - Confident
    - Neutral
    - Unconfident
    - Very unconfident

7) What are some of the risks caused by malicious software running on your Smart TV? Please explain to the best of your ability:

8) Do you believe that a human being will have access to data collected by your Smart TV?
    - Yes
    - No
    - I'm not sure

9) How likely is it that data collected by various third-party apps will be accessed by the following parties: Hardware manufacturers (e.g., Samsung, LG, Sony, etc.), ISPs, Cable/Satelite TV Providers, Financial companies, Law enforcement, TV networks, Hackers, Data aggregators (i.e., companies that resell consumer data), and Advertisers?
    - Completely likely
    - Somewhat likely
    - Neutral
    - Unlikely
    - Completely unlikely

10) How acceptable would it be if data collected by third-party apps was accessed by the following parties? *[parties same as previous question]*
    - Completely acceptable
    - Somewhat acceptable
    - Neutral
    - Unacceptable
    - Completely unacceptable

11) Are there any laws or regulations that you believe prevent Smart TV manufacturers from sharing data collected in your home with other parties? Please explain to the best of your ability.

12) How likely would you be to disable third-party apps from running on your Smart TV in order to prevent data captured from within your home from being accessed by others?
    - Definitely would not disable
    - Might not disable
    - Unsure
    - Might disable
    - Definitely would disable

13) If a Smart TV manufacturer shared data collected from within your home by third-party apps with others, would this increase or decrease the likelihood that

you would purchase a device from this manufacturer again?

- Definitely would not purchase
- Might not purchase
- Unsure
- Might purchase
- Definitely would purchase

14) How likely is it that data collected from your Smart TV will be combined with data collected from your other Internet-enabled devices (e.g., smartphone, tablet, laptop, etc.) to create a detailed profile of your habits and interests?

- Extremely likely
- Likely
- Neutral
- Unlikely
- Extremely unlikely

15) Thinking about the questions in this survey, have you thought about these issues before? Please explain:

16) In what year were you born?

17) What is your gender?

18) What is your highest level of education?

### *D. Codebook for Qualitative Questions*

Contact the authors for a copy of the codebook that was used to classify themes in the qualitative questions.