Ask the Experts: What Should Be on an IoT Privacy and Security Label?

Pardis Emami-Naeini
Carnegie Mellon University
pardis@cmu.edu

Yuvraj Agarwal
Carnegie Mellon University
yuvraj@cs.cmu.edu

Lorrie Faith Cranor
Carnegie Mellon University
lorrie@cmu.edu

Hanan Hibshi
Carnegie Mellon University
hibshi@andrew.cmu.edu

Abstract—**Information about the privacy and security of Internet of Things (IoT) devices is not readily available to consumers who want to consider it before making purchase decisions. While legislators have proposed adding succinct, consumer accessible, labels, they do not provide guidance on the content of these labels. In this paper, we report on the results of a series of interviews and surveys with privacy and security experts, as well as consumers, where we explore and test the design space of the content to include on an IoT privacy and security label. We conduct an expert elicitation study by following a three-round Delphi process with 22 privacy and security experts to identify the factors that experts believed are important for consumers when comparing the privacy and security of IoT devices to inform their purchase decisions. Based on how critical experts believed each factor is in conveying risk to consumers, we distributed these factors across two layers—a primary layer to display on the product package itself or prominently on a website, and a secondary layer available online through a web link or a QR code. We report on the experts' rationale and arguments used to support their choice of factors. Moreover, to study how consumers would perceive the privacy and security information specified by experts, we conducted a series of semi-structured interviews with 15 participants, who had purchased at least one IoT device (smart home device or wearable). Based on the results of our expert elicitation and consumer studies, we propose a prototype privacy and security label to help consumers make more informed IoT-related purchase decisions.**

Index Terms—**Internet of Things (IoT), Privacy and Security, Label, Expert Elicitation, Delphi.**

I. INTRODUCTION

With the rapid growth in the development and deployment of IoT devices worldwide, large numbers of privacy and security issues have come to light [88]. The popular press has reported on incidents such as Amazon Alexa sending private conversations to a random phone contact [58] and baby monitors getting hacked [95]. Even well-known IoT manufacturers have faced criticism for lack of transparency about their data practices. For instance, recent news articles reported that Google forgot to mention that its Nest Secure hub has a microphone [76] and Amazon revealed that in addition to artificial intelligence algorithms, human employees listen to a subset of audio recordings from Echo devices [90].

Surveys have indicated that consumers are indeed worried about the collection of their personal data and who it is being sold to or shared with [16], [97]. Privacy is of particular concern to users of IoT devices. Mozilla surveyed 190,000 participants around the world and found that privacy is the biggest IoT-related concern for 45% of the respondents [15]. In a survey conducted by the Economist Intelligence Unit (EIU) of over 1,600 IoT consumers, 92% of participants were concerned about their privacy and wanted to have control over personal information collected by smart devices [56], [81].

While consumers are increasingly interested in purchasing smart devices [43], [79], a recent survey conducted in the United Kingdom showed that security and privacy are key pieces of information that participants would like to consider when purchasing them [46]. Similarly, an interview-based study of IoT consumers reported that almost all participants desired to have information about the privacy and security of IoT devices at the time of purchase [37]. Currently, there is little public information about the privacy and security of IoT devices that consumers can access to inform their purchase decisions [11]. Some resources, such as the Mozilla "Privacy Not Included" website [71] and a report published by the UK Information Commissioner's Office [50], provide information about specific IoT devices or address limited privacy and security factors.

Critical privacy and security information could be provided to consumers by including it prominently on a privacy and security label accompanying the device. This could also increase consumers' trust in the device manufacturer [106]. In a May 2019 proposal, the UK Digital Ministers declared their intention to mandate security labels for IoT devices, with the goal of notifying consumers about security aspects of these devices [32]. However, this plan only covers three security practices: using no default passwords, having a vulnerability disclosure program in place, and specifying the lifetime of security updates. Other proposals for IoT privacy and security labels fail to specify the specific information that consumers should be presented with on the label [30], [38], [40], [62], [66], [74], [80].

Given consumers' scarce attention, presenting them with the most relevant security and privacy information in the most digestible form is crucial. To determine the most important information to include on IoT privacy and security labels, we solicited the opinions of privacy and security experts. In various fields, expert elicitation has been used effectively for research and decision-making [8], [54], [86], [102], [103], particularly under uncertainty and when necessary information cannot be obtained from other sources [21], [47], [55].

We conducted interviews and surveys with 22 privacy and security experts. To get different perspectives, we recruited experts from industry, academia, government, and non-

governmental organizations (NGOs). We also ensured that these experts come from different backgrounds related to IoT (software, hardware, and policy). We used the iterative Delphi methodology (explained further in Section III) to develop a consensus among the experts around important factors and an understanding of their reasons for or against including each factor. Overall, we found that differences in opinions were driven less by fundamental differences in beliefs, but rather by differences in work experience and priorities. For example, some experts were more knowledgeable about specific security mechanisms, standards, or regulations, and prioritized factors related to their area of expertise or their organization's mission. Prior research has shown that security experts might analyze the same artifacts differently depending on their background in specific security domains [48].

Most factors identified as important by experts are factors that they believe will inform consumers. Experts also identified some factors for inclusion that could inform experts only, mostly to be able to hold companies accountable.

Prior studies suggest that layered labels can be effective [23], [37], [83], [96]. A layered label includes a *primary* layer that presents the most important and glanceable content, followed by a *secondary* layer for additional information. In our study, we asked experts to specify the layer on which the information should be included on the label. They mostly recommend putting only information that would be understandable and important to most consumers on the primary layer.

We designed a prototype layered label based on our expert elicitation study. We then conducted semi-structured interviews with 15 consumers of IoT devices (smart home devices or wearables) and presented our prototype to them. We show that all of our participants had a clear understanding of the information presented on the primary layer of the label. Although some of the factors on the secondary layer of the label were less understandable to participants who lacked privacy and security expertise, all of our participants reported that they still want such information to be included on the label mainly to be as informed as possible. In addition, consumer participants reported that having all the important privacy and security factors, even unfamiliar ones, on the label would help them easily search online to find more information.

We make the following contributions in this paper:

- We distill an extensive list of privacy and security factors to identify the most important pieces of information to include on IoT labels.
- We partition the most important factors into two layers: the primary factors we want consumers to notice and consume at a glance, and the secondary factors that require more space to effectively convey risk to consumers.
- Based on our expert and consumer interviews and surveys, we propose a prototype IoT label that includes the most important factors with proposed groupings.

II. BACKGROUND AND RELATED WORK

In this section, we first outline prior research on understanding consumers' privacy and security related concerns for IoT devices. Next, we provide background on how labels have been used in different contexts to inform consumers' choices. Finally, we discuss reports and recommendations on privacy and security best practices for consumer IoT devices.

A. Consumers' Privacy and Security Concerns and Preferences

Numerous studies have demonstrated that consumers are increasingly concerned about the privacy and security of IoT devices [5], [15], [45], [87]. Researchers have shown that the extent of these concerns depends on factors such as the type of data collected, the purpose of data collection, and the retention of collected data. Lee and Kobsa [59], [60] explored factors related to IoT data collection and found that people are more concerned about data being collected in private locations compared to public locations. In addition, they found that people's concerns depend on the type of data being collected. In a previous study Naeini et al. used vignettes to study how different factors can explain variations in comfort level related to data collected by sensors. They found that factors such as the purpose of data collection and the retention time significantly impact people's privacy-related concerns [73].

Despite such concerns, consumers are still purchasing IoT devices, mostly due to their perceived convenience and features [18]. This is sometimes referred to as a "privacy paradox," thanks to the discrepancy between privacy concerns and actions taken to mitigate those concerns [1]. One cause for this could be that consumers are provided with little, or often no, privacy and security information about IoT devices prior to purchase [10], [11], [30], [37]. This prevents consumers from making informed IoT-related purchase decisions and increases the risk of privacy and security vulnerabilities, which may lead to high-profile and large-scale attacks targeting IoT devices [4].

B. Product Labels

Product labels, such as food nutrition and energy labels, have been used to aid consumers' purchase decisions. Food nutrition labels, in particular, were developed to decrease obesity by helping consumers purchase healthier food products [33]. Other objectives of food nutrition labels include encouraging food companies to compete to produce healthier products and allowing governments to support consumers' health-related behaviors without mandating specific nutritional requirements [53].

The effectiveness of food nutrition labels has been shown to depend on factors such as consumers' attention at the point of sale [92], whether the consumer is purchasing the product for children, aiming to lose weight, and/or purchasing the product for the first time [84], [104], their nutrition-related knowledge, health condition [25], [34], and socio-demographic factors. Studies have also shown the limitations of nutrition labels in effectively communicating nutrition information to consumers to improve eating habits [35], [49]. Despite these shortcomings, these studies demonstrate that food nutrition labels significantly inform consumers' purchase decisions [70], [75].

In the realm of privacy, researchers have explored the impact of privacy "nutrition labels" on websites. They found that privacy labels help users find important information faster and

more accurately, as compared to finding such information in traditional privacy policies [52].

C. Privacy and Security Guidelines and Best Practices

Tanczer et al. [93] conducted an extensive literature review on publicly available reports from industry associations and international organizations on their security-related proposals and recommendations for consumer IoT devices. They reviewed 17 industry reports (including from Intel, HP, and Consumer Technology Association) and policy reports (including from European Commission, International Organization for Standardization (ISO), and Alliance for the Internet of Things Innovation (AIOTI)). This review observed 19 overarching principles related to security best practices that were referred to at least twice in these reports. The most common principles (mentioned in at least 10 reports) were strong authentication by default, reliable and cryptographically signed security updates, encryption by default, and compliance and risk assessment. Some of the other principles were related to physical security, vulnerability reporting and disclosures, and secure device boot. The security factors that we synthesized based on our expert elicitation study covered all of the most frequent security principles mentioned in this literature review [93].

Tanczer et al. concluded that in general, the industry acknowledges the importance of selling safe and secure IoT devices in the market and would like to work alongside the government as part of their efforts. However, they are more interested in self-regulation than in government interventions [93]. For example companies can self-certify their IoT devices using a framework developed by IoT Security Foundation (IoTSF) that specifies five levels of compliance [51].

A recent UK government report argued against self-regulation, noting the lack of incentives for IoT companies to adhere to security best practices when designing IoT products. The report recommended that the government mandate specific requirements for IoT devices as a mechanism to improve the security of consumer IoT products [31]. These requirements are no default password, availability of a vulnerability disclosure program, and security updates. These recommended requirements all are included in our privacy and security label (see Figures 3 and 4).

Notably, all the reviewed reports above focused on devices' security mechanisms with few references to data privacy considerations. As consumers are concerned about both the privacy and security of their devices, in our work we asked experts and consumers about both privacy policies and security mechanisms that should be included on a label.

III. METHODOLOGY

We first conducted an expert elicitation study to specify the content of a privacy and security label for IoT devices. We complemented the expert study with a series of 15 semi-structured interviews with non-expert consumers and iterated on the label design.

A. Expert Elicitation Study

In the expert elicitation study, our overarching goals were to identify factors that experts believed would be useful to include on a privacy and security label for IoT devices and to understand the experts' rationale for selecting each factor. We conducted an in-depth, semi-structured interview study, followed by two rounds of surveys with 22 privacy and security experts. The process is depicted in Figure 2.

1) Participant Recruitment and Compensation: To capture a wide range of expert opinions, we recruited experts from academia, industry, government, and non-governmental organizations (NGOs) in the United States, with a diverse range of expertise: software, hardware, policy, standards, and user experience (UX). We recruited experts with whom the authors had interacted professionally or were recommended by other experts. We carefully selected experts, who are all well-known in their respective fields. More specifically, we looked for experts, who satisfied at least one of the following qualification criteria. Seven experts met two criteria.

- Computer science or engineering professor in the field of privacy and security.
- More than 10 years of research or practice in the field of privacy, security, or policy.
- Author of notable books in the field of privacy and security.
- Active involvement in cybersecurity standardization.
- Leading a corporate IoT product team.

After identifying the experts who met the qualifications we were looking for, we contacted them and invited them to either come to our institution for an in-person interview or join our interview study online over Skype. All the interviews were audio recorded and transcribed by a third-party service. We compensated experts with a $25 Amazon gift card.

2) Delphi Method: As defined by Delbecq et al., the *Delphi method* is "a method for the systematic solicitation and collection of judgments on a particular topic through a set of carefully designed sequential questionnaires interspersed with summarized information and feedback of opinions derived from earlier responses" [6]. This method of qualitative research was originally developed by Dalkey and Helmer in the 1950s and has been widely used to reach consensus between a group of experts without face-to-face interactions [29]. The Delphi method has been used in a number of studies related to policy design and implementation [2], social science [91], and human-computer interaction [69].

The Delphi method has three important features. First, the responses as well as group interactions in each round are anonymized. Second, the process involves multiple rounds of data collection procedures (e.g., interview, survey), and finally, in each round, the summary of the previous round is shown to experts as a means to reach consensus [19], [27], [28]. The study continues until consensus is reached, which generally occurs after three iterations [63].

3) Expert Interviews: The first phase of the Delphi method is open ended [72]. Therefore, our first step was to conduct
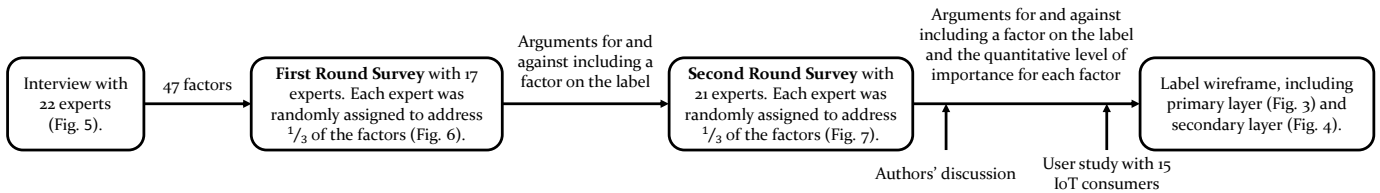
Fig. 1: We followed a three-round Delphi method, by conducting an interview study and two rounds of surveys. Finally, we designed a label prototype that captures the findings of the process, inputs from authors' multiple rounds of discussion, and a user study with 15 IoT consumers.

semi-structured interviews with privacy and security experts.

We began the interviews by introducing the idea of a privacy and security label and its similarity to a food nutrition label. Following the introduction to the study and its goals, we asked experts to provide their definition of privacy and security as it relates to IoT devices. Next, we asked experts to think about the content of the label and specify the information that they think should be on a privacy and security label for IoT devices. For each piece of information they specified, we asked them to consider whether it was relevant to consumers or experts. In an iterative process, we compiled a list of security and privacy factors suggested by the experts we interviewed, and added new factors suggested during each interview. Towards the end of each interview, we presented the full list of factors so that each expert reviewed their own factors, as well as the factors suggested by previously-interviewed experts (see Figure 5).

4) First Round Survey: The expert interviews resulted in an extensive list of privacy, security, and general factors that experts wanted to see on an IoT label. We then conducted a survey of the same experts to understand the rationale behind their preferences (see Figure 6). In order to decrease fatigue, we split the factors in the survey so that each expert was presented with one-third of all the factors. For each expert, the ordering of factors was randomized.

When introducing the survey to experts, we explained that in a layered IoT label, the first or primary layer would include the most important information, and the secondary layer would contain the information on the primary layer as well as additional helpful information. We also advised experts not to worry about the design of the label when answering the questions. Then, for each factor, we asked the experts to specify whether they believe that factor is important to include on the label, and to provide reason(s) that support their answer.

5) Second Round Survey: For agreement over the inclusion and exclusion of factors, we conducted a second survey with the same set of experts (see Figure 7). When introducing the second survey to the experts, we described our two key objectives for the content of the IoT label: to inform consumers, and to provide a means for holding companies accountable for their privacy and security practices. To reduce respondent fatigue, participants answered questions for one-third of all the factors, randomly chosen from the three categories—security, privacy and general—such that they saw approximately the same number of questions within each category.

In this second survey, we used data collected from our interviews and first survey. We presented each factor from our

dataset alongside the experts' reasons for inclusion or exclusion. Then, on a five-point Likert scale, we asked each expert to decide whether they believe the factor should be included on the label and to provide their rationale if different from what we presented. Next, we asked them to specify on which layer of the label they would like to place this factor and the rationale behind their choice. We asked them to classify the factor as being most relevant to privacy, security, or general information. Finally, we asked experts to provide any additional comments about the factor that came to mind.

In addition to the questions we asked for each factor, we asked experts about their opinion on separating or merging privacy and security sections on the label. At the end of the survey, we asked experts to state their privacy and security expertise and domain of knowledge, followed by some general demographic questions.

6) Data Analysis: We collected approximately 22 hours of interview audio recordings. We used thematic analysis to qualitatively summarize interview transcripts, following the approach suggested by Braun and Clarke [14]:

- Phase 1: A primary coder read the interview transcripts and took notes, listening to parts of the audio files as needed when the transcripts were incomplete.
- Phase 2: The primary coder created the initial codebook by examining the notes from the interview phase and the notes from Phase 1 above, listening for reasons for and against including factors on an IoT label. This step did not focus on finding patterns in the responses.
- Phase 3: The primary coder merged the smaller codes into broader themes. This step focused on finding patterns and themes from the long list of codes from Phase 2.
- Phase 4: The themes that emerged from Phase 3 were reviewed and discussed by the authors of the paper to resolve any disagreements. This step helped increase the validity of the themes. In an iterative process, some of the themes were removed from the codebook and some themes were merged into more general themes until we achieved consensus on the final themes.
- Phase 5: The finalized themes (reasons to include or exclude each factor) were moved into the final codebook.

The finalized privacy, security, and general factors were used as input to the first round of our survey, where we asked experts to provide us with their arguments. We then followed the same coding process described above to code the open-ended survey responses. After the first survey, we revised the themes (reasons

for and against including a factor on the label) in the codebook and presented them to the experts in the second survey.

We reached a point of saturation in terms of finding new factors after interviewing 20 experts. In other words, no new privacy, security, or general factors were mentioned by our participants in the rest of the interviews as well as the two follow-up surveys.

Thematic analysis is purely qualitative and inductive [14]. The literature showed that having more than one coder does not make the codes objective, since two coders could apply the same subjective perspective to the data [67]. Indeed according to a survey of CSCW and HCI publications from 2016 to 2018, only 6% of papers using thematic analysis used multiple coders and measured Inter-rater reliability [68].

An iterative, yet inductive, analysis approach in thematic analysis increases the reliability of the theme-finding process [42]. All the themes were iteratively and extensively discussed among the researchers in the group. For any disagreement, researchers traced the theme back to its corresponding subcodes and checked whether the source of disagreement was the subcodes that were used. If not, we traced the subcodes further back to experts' quotes from the transcriptions and we then decided on the appropriate subcodes and the appropriate themes arising from the subcodes. This iterative approach is recommended with qualitative methods that are high in subjectivity [42], [65]. We also improved the reliability by consulting with the expert participants using the second round survey mentioned above; this triangulation method is known as testimonial validity or member checking [82].

B. Semi-Structured Interviews with Non-Expert Consumers

We used the results of our expert study to inform the development of prototype designs for primary and secondary labels. We created prototype boxes for two fictitious brands of security cameras and included a primary-layer label on each box. We put the corresponding secondary-layer labels on a mock-up of an online shopping website. Next we conducted a semi-structured interview study with 15 non-expert consumers to gain insights into how they would use these labels and how well the labels convey risk.

1) Participant Recruitment and Compensation: We recruited participants by posting on Craigslist, Reddit, and our institution's recruitment website. Participants were required to be at least 18 years old and have purchased at least one smart home device or smart personal device. Prospective participants completed a short screening survey, in which we collected demographics and asked about what IoT devices they had purchased and how they purchased them. We invited a diverse sample of qualified participants to our lab for a 1-hour interview. Each interviewee received a $25 Amazon gift card.

2) Initial Questions: We showed participants a box for a hypothetical security camera that did not include a label and asked them what they could tell about the privacy and security of the device by looking at its box. We then asked participants whether they had ever seen an informative label on any product. Next, we presented them with one of the two labeled security

Fig. 2: A user study participant comparing the privacy and security practices of two hypothetical smart security cameras.

camera boxes and asked them what they could tell about the privacy and security of the device. We asked a number of questions to study participants' understanding of the content of the label and whether the information conveyed risk.

3) Risk Communication in Comparative Purchase Process: We showed participants the other labelled security camera box, and told them this camera had the same price and features, but with different privacy and security information. We asked them to compare these two products and discuss which has better privacy and security, which device would they purchase, and the information that helped them make this decision.

We then told participants about the secondary layer of the label, which can be accessed by scanning the QR code or typing in the URL on the first layer. After introducing the idea of the layered label, we asked participants whether they had ever seen one on any other product. We asked them to discuss the pros and cons of a one-layer and two-layer label.

4) Information Comprehension in Non-comparative Purchase Process: We asked participants to look at the information on the label of the product they decided not to purchase and to discuss their concerns and their understanding of the information.

5) Risk Communication in Non-comparative Purchase Process: We asked participants to specify the factors that seemed risky to them from a privacy and security perspective and discuss what kinds of risk they would be exposed to. We also asked how the product could be improved to reduce this risk.

6) Secondary-Layer Information Comprehension: We asked participants whether they would prefer to scan the QR code or type in the URL to look for additional information. Based on their preference, we scanned the QR code or typed in the

URL on the primary layer to show the secondary layer to participants. We then asked them to start from the beginning of the label and tell us what each factor means to them, how useful they believe each factor would be, and if they have any suggestions to make the information more understandable.

7) Label Format: We asked questions about the label format, including the separation of factors into privacy, security, and general information sections. We also asked participants to specify the factors that they believed are currently misplaced, and should be either removed from the label or moved to another section or layer of the label.

8) Purchase Behavior: Finally, we asked questions to understated participants' purchase behavior related to online and in store shopping.

9) Data Analysis: We collected about 15 hours of audio recordings, which we had transcribed. The first author was the primary coder who created the codebook and kept it updated throughout the coding process. To analyze the data, we used structural coding, which is appropriate for coding semi-structured interviews [64], [82]. We defined four structural codes (e.g., attitudes toward layered labels, reasons to include or exclude a factor from the label), which we divided into 13 subcodes (e.g., being as informed as possible, lack of relevance to privacy and security). Unlike thematic analysis, structural coding is more objective, and results in a codebook used for categorization [82]. Therefore, having more than one coder and using inter-rater reliability is helpful in testing the reliability of the codebook [42]. Each interview was independently coded by two researchers, who then discussed and iteratively revised the codebook. After resolving the coding disagreements, we reached the Cohen's Kappa inter-coder agreement of 84%. Cohen's Kappa inter-coder agreement of over 75% is considered as "excellent" rate of agreement [41]. In case of disagreement, we report on the results of the primary coder.

C. Ethics

Both expert and consumer studies were reviewed and approved by our Institutional Review Board (IRB). All participants provided their informed consent to participate in the surveys and interviews, to have audio recorded, and to have the recordings transcribed by a third-party transcription service.

D. Limitations

Expert elicitation is prone to overconfidence and cognitive biases [89]. To reduce overconfidence, in the second expert survey, we presented strong arguments for and against having each factor on the label so that experts could read the rationale provided by other experts before indicating their own preferences.

The experts interviewed in this study are not representative of the entire population of privacy and security experts. Our aim was to surface a wide variety of expert viewpoints. Therefore, we recruited experts with diverse expertise related to IoT security and privacy and from different sectors. We selected experts based on our inclusion criteria, as discussed in Section III-A1.

TABLE I: We conducted an expert elicitation study with 22 privacy and security experts. NGO stands for non-governmental organization and UX stands for user experience.

Expert ID	Privacy & Security Expertise	IoT Focus	Workplace
P1	Privacy	Policy, standards	Enterprise
P2	Privacy	Policy, UX	Enterprise, NGO
P3	Privacy	Software	University
P4	Privacy	Policy	University
P5	Privacy	Hardware	Enterprise
P6	Privacy	Policy, software, UX	Enterprise
P7	Privacy	Policy	NGO
P8	Privacy	Policy, privacy	NGO
P9	Privacy	Policy, privacy	NGO
S1	Security	Software	University
S2	Security	Software	University
S3	Security	Policy, software	Government, University
S4	Security	Policy, security	Enterprise
S5	Security	Hardware, security	Enterprise
S6	Security	Software	Enterprise
S7	Security	Policy	Enterprise, NGO
S8	Security	Policy	NGO
S9	Security	Hardware, software	Enterprise
B1	Both	Software	University
B2	Both	Policy, software	Enterprise
B3	Both	Policy, standards	NGO
B4	Both	Policy, software	Enterprise

In the follow-up surveys with experts, we presented each participant with only one-third of factors randomly sampled from the list of all privacy and security factors. This reduced respondent fatigue [7] and increased the quality of responses, at the cost of not being able to achieve a true consensus across all experts. In the expert study, our main objective was to collect the opinions of diverse experts, and not to reach perfect consensus. Therefore, we report themes that were only mentioned by a few experts.

We designed a label prototype based on findings from our three-round Delphi process. However, we did not conduct a fourth round of study to show experts the label and ask them for feedback. Although this would have helped us confirm experts' opinions about the factors in the context of a complete label design, it would have introduced confounding factors related to the design of the label, including, but not limited to, the order of sections on the label and the specific language used to convey the information. Since the expertise of the participants in our study was in the area of IoT security and privacy, and not in communications design, we limited the expert elicitation study to focus on the individual factors.

Our consumer study was a small-scale qualitative study designed to gain initial consumer feedback and assess the overall usefulness of the layered label approach in this context. Additional large-scale iterative design and testing is needed to refine and validate the label design.

IV. RESULTS

We conducted 22 one-hour, semi-structured interviews with IoT privacy and security experts with diverse backgrounds as described in Table I. We compiled a list of 47 privacy, security, and general factors that experts said they would like to see on the IoT label.

TABLE II: User study participants, demographics, and devices they have purchased.

Participant ID	Gender	Age	Technical Background	IoT Devices our Participants Have Purchased
C1	F	35-44	Y	Thermostat, TV, switches, lock, outlet
C2	F	35-44	Y	Vacuum cleaner, gaming consoles
C3	M	55+	Y	Thermostat
C4	F	45-55	Y	Smart speaker
C5	M	35-44	Y	Camera
C6	M	55+	N	Smart speaker, lights
C7	M	18-24	N	Smart speaker, lights, TV
C8	M	25-34	N	Activity tracker
C9	F	25-34	N	Smart speaker
C10	F	25-34	Y	TV, camera
C11	F	18-24	N	Smartwatch, activity tracker, camera
C12	F	35-44	N	Smart speaker
C13	M	25-34	N	Smart speaker, TV, plugs, vacuum cleaner
C14	F	25-34	N	Smartwatch, activity tracker, vacuum cleaner
C15	F	25-34	Y	Smart speaker, smartwatch

We followed the expert interviews with a qualitative survey to understand the reasons experts wanted to include or exclude each factor. Out of 22 invited experts, 17 answered the first survey, with each of them being asked to comment on one-third of the 47 factors. This survey took an average of 16 minutes to complete. We collected on average seven reasons for or against including each factor on the label. We then conducted thematic analysis on the arguments provided to arrive at two or three primary reasons for and against each factor.

In the second survey, we presented experts with the reasons for and against each factor from the previous two phases and asked them to rate their enthusiasm for including the factor on either the primary or the secondary layer of the label. 21 experts participated in the second survey, spending an average of 15 minutes. We identified 12 factors that most experts recommended including on the primary layer and 13 factors most experts recommended including on the secondary layer.

Based on the expert study and authors' discussions, we designed prototype privacy and security labels for hypothetical smart security cameras and presented them to a diverse sample of 15 non-expert consumers (see Table II). We asked them questions related to their understanding of the factors on the label and whether they conveyed risk. We iteratively improved the content of the label to make it more understandable, resulting in a final prototype label.

In this section, we first discuss experts' attitudes toward privacy and security. Next we present the factors experts wanted to include on the primary and secondary layers of the label or exclude from the label, followed by a discussion of how consumers perceived those factors. We continue by discussing consumers' attitudes toward the labels and their layered design. Finally, we present our prototype label design that we designed based on experts' and consumers' input.

A. Definition, Assessment, and Accountability

We started the interviews by asking experts to define privacy and security related to IoT devices. When defining security, almost all experts (21/22) mentioned the CIA triad of confidentiality, integrity, and availability. However, experts had different definitions for privacy. Some experts (9/22) defined privacy as having transparency and control over data practices and some experts defined privacy as the confidentiality aspect of security (8/22). Overall, experts' definitions for security were mostly passive and focused on hardware and software enforcement mechanisms. On the other hand, their definitions for privacy were active and centered around policy, control, and individuals' preferences and comfort. For instance, P7 compared privacy and security practices by saying: "If the privacy is done right, it would be more active than security because the consumer would be able to be in control." B3 explained how privacy and security are related by saying: "Security mechanisms are the things that enforce the technical controls that allow us the privacy we have."

Most experts (15/22) believed that security information is less tangible and understandable for consumers compared to privacy information, in part because it relies on technical mechanisms. S5 explained: "Consumers don't necessarily understand some of the abstract stuff about security that they don't see. Whereas when their privacy is breached, they are more aware of that."

In addition, almost all experts (19/22) reported that security practices are easier to measure and assess than privacy practices, as security is more objective and less controversial, while privacy is more subjective and context dependent. P5 explained that security is easier to quantify: "Security strikes me as less subjective, and, therefore, easier to measure. Which is to say that there could be certain standards. What sort of encryption exists on the device? What encryption is in the cloud? These are all fairly quantifiable. Whereas privacy is trickier, I think. And almost ethically, and morally from my point of view, there's a lot more gray area in this." This finding is aligned with the current efforts in IoT assessments and scoring, which are more focused toward security mechanisms than privacy practices [22], [39], [44], [51], [107].

Experts reported that IoT privacy and security labels could increase accountability (most frequently mentioned for factors on the secondary layer) and transparency. Seven experts suggested that increased transparency could be an incentive for companies to compete on privacy and security, leading to safer products. S4 explained: "There is value in forcing the company to write a list down even if the consumer doesn't understand it. If you said, 'list your open ports,' there would be an incentive to make them few."

Some experts (8/22) mentioned that IoT companies' accountability should be different for privacy and security breaches. They said that security breaches can happen accidentally, even if companies follow best practices. However, privacy violations could be intentional and IoT companies could even profit from them. P4 explained: "You can have the best intentions in the world, but if somebody comes up with some crazy hack overnight, you shouldn't be held responsible for it. You should be held responsible for fixing it. As opposed to if you intentionally share someone's data with a third party, it's not like, oh you could have prevented it but you chose to do it."

In the second survey, we asked experts to specify whether they prefer to see privacy and security factors in two separate sections or if we should combine them into one section. About half of the experts (10/21) believed privacy and security should be presented in separate sections. Most of these (9/10) said such separation would improve the readability and utility of the label and help educate consumers. P1 explained: "I lean toward the option of separation, because I'd like to see a streamlined label for most consumers to 'consume' as easily and quickly as possible." S8 concurred: "Consumers may have preferences for one aspect more than the other and stating them separately better enables consumer choice and education." However, the other half of the experts believed privacy and security factors should be combined into one section (11/21). For example, P4 believed that for some consumers, security seems more important than privacy. Thus, separating them on the label may cause consumers to focus only on security factors and ignore privacy information. Among those experts, who were more interested in combining privacy and security information into one section, almost all of them (9/11) mentioned that privacy and security are so related that it is not possible to completely separate these two concepts.

In the label that we presented to consumer interview participants, we grouped information into three main sections: security mechanisms, data practices, and general/more information. Participants preferred the proposed separation of sections and reported that these groupings made sense to them.

B. Factors to Include in the IoT Label

From the second survey, we found 30 factors that at least 4 out of 7 of the experts recommended including on the label (either on the primary or secondary layer) and 17 factors that at least 4 out of 7 experts recommended excluding from the label. Note that since only a third of the factors were shown to each expert (total 21 on the second survey), at least 4 responses constituted a majority.

The authors discussed the experts' arguments and preferences for each factor and made a decision as to whether or not each factor should be included and if so, on which layer of the label. In some cases, we made a decision that contradicted the majority of experts if we felt that their arguments could be accommodated in a different way.

1) Primary Layer: We found 12 factors that at least 4 out of 7 experts wanted to include on the primary layer:

- Privacy rating for the device from an independent privacy assessment organization
- Security rating for the device from an independent security assessment organization
- The date until which security updates will be provided
- Type of data that is being collected
- Type of sensor(s) on the device
- Whether or not the device is getting cryptographically signed and critical automatic security updates
- Types of physical actuations (e.g., talking, blinking) the device has and in what circumstances they are activated
- Whether or not the device is using any default password

- Frequency of data sharing (e.g., continuous, on demand)
- The warranty period of the device
- Level of detail (granularity) of the data being collected, used, and shared (e.g., identifiable, aggregate)
- Access control for device and apps (e.g., none, single-user account, multi-user account)

Experts were interested in including these factors on the primary layer because they considered them necessary for consumers to know, they convey critical information about the privacy and security of the device, and they inform consumers' purchase decisions. For example, P1 explained why the type of collected data should be included on the primary layer: "I think this is the most useful information to be provided to consumers for them to compare privacy risks of IoT devices."

All of our consumer participants understood the information presented on the primary layer and were able to talk about privacy and security implications of each factor. For example, consumers associated the expiration date of the device to its security updates lifetime. C3 mentioned "planned obsolescence" when talking about the security update lifetime: "I do like the fact that you say when the security updates will no longer be available, because that alerts people to the fact that this device is going to expire. People have thermostats that last for decades and it's useful to know that this is planned obsolescence." One of our participants brought up a point of skepticism, related to how long a company claims to provide security updates: "I'm skeptical because I know that tech startups can very rarely guarantee that their servers will be online for three or more years."

Among the factors experts believed should be included on the primary layer, there were three factors that we decided to either move to the secondary layer or exclude from the label. Note that we also removed privacy and security star ratings later in the process as mentioned in Section V. We expect to add them back when such assessments are available in the future.

First, we decided to move the physical actuation factor to the secondary layer because this information is not usually directly related to the privacy and security of the device. In the consumer study, all participants found this information useful from a safety point of view, but none wanted us to move this factor to the primary layer as they reported that the information conveyed by this factor does not have privacy and security implications for them.

Second, we decided to move frequency of data sharing to the secondary layer because prior work has shown that most people do not understand the privacy and security implications of the frequency of data sharing [9]. While almost all the consumers we interviewed were concerned about data sharing, only four mentioned privacy concerns related to the frequency of sharing.

Third, we decided to exclude device warranty period from the label because it has few, if any, privacy, security, or safety implications.

2) Secondary Layer: We found 13 factors that at least 4 out of 7 experts wanted to include on the secondary layer of the label:

- Retention time
- Purpose of data collection
- What information can be inferred from the collected data
- Supported standards (e.g., Wi-Fi, Zigbee)
- Where the collected data is stored
- Whether or not the collected data will be linked with data obtained from other sources
- Special data handling practices for children's data
- The control that users are offered (e.g., opt-in/out from data sharing)
- Data-collection frequency (e.g., once a month, on install)
- Whether or not the device can still function when Internet connectivity is turned off
- Relevant security and privacy laws and standards to which the device complies (e.g., ISO 27001, GDPR)
- Link to the device's key management protocol
- Resource usage in terms of power and data (e.g., kw, kbps)

Experts mentioned two common reasons to include a factor in the secondary layer rather than the primary layer: the factor requires detailed information to convey risk to consumers (mentioned by 6/7) or the factor does not convey critical information related to the privacy and security of the device (mentioned by 4/7).

Among the factors our experts wanted to include on the secondary layer, there were three factors that we decided to include on the primary layer instead: date of the latest firmware update, purpose of data collection, and where the collected data is stored.

Experts wanted to have the date of firmware update on the secondary layer mainly because these updates happen frequently and the information on the label can become outdated. On the other hand, consumers need to know the firmware version to which the label is applicable. Therefore, we believe the firmware version number and date information should be provided on both layers of the label.

Most experts (6/7) believed that it would be hard to fit all the purposes of data collection on the primary layer of the label. Therefore, they recommended including this information on the secondary layer. However, past research has shown that purpose of data collection is one of the most important factors consumers want to consider when making privacy decisions [12], [59], [61]. Purposes may be grouped into high-level categories that could be included on the primary layer. For example, the W3C's Platform for Privacy Preferences (P3P) standard identified 12 purpose categories [105]. The consumers we interviewed indicated that it was important to them to know the type of data collected and the purpose of collection when making device purchase decisions.

Experts stated that where the data is stored should be included on the secondary layer because it is not relevant to privacy or security. However, we believe local storage versus data being stored on the cloud can indeed have different privacy and security implications [85]. Therefore, we decided to include

this in the primary layer. Moreover, most consumer participants were able to reason about privacy and security implications of cloud versus device and discuss the trade-off between security and convenience. C10 talked about this trade-off by saying:

> The advantage of the cloud is that if the device is damaged, you can still access it. So it's going to be always available as long as you can access internet. The other issue with the cloud though is that, like, it can be hacked and also, who has access to that is less clear, or you have less control over that. But I can always access it from whatever device I have and it's convenient.

Although experts recommended including information about device resource usage on the secondary layer, we decided to remove this factor from the label due to its lack of privacy, security, or safety implications.

3) Factors with no Specific Layer: There were four remaining factors that at least 4 out of 7 experts were enthusiastic about having on the label, but their opinions were split between including them on either the primary or the secondary layer. These factors were:

- Who the data is shared with
- Who the data is sold to
- Whether or not the device can still function when data-driven smart features (e.g., the learning function of smart thermostat) are turned off
- Whether or not the device has parental control mode

For these factors, about half of the experts reported that they would like to include them on the primary layer since they are important privacy and security factors that consumers should know about before making purchase decisions. The other half of the experts were not enthusiastic about including these factors on the primary layer.

We decided to put the factors related to parental controls and device functionality when smart features are turned off on the secondary layer because they are not directly related to security or privacy.

Some experts (3/7) noted that who data is shared with or sold to is likely to change over time, and recommended putting these factors on the secondary layer where they could be updated more easily. We showed consumers a label with these factors on the secondary layer. However, all consumer participants expressed concern when they saw that their information could be shared and sold with third parties and 8 out of 15 said who data is sold to or shared with were among the most important factors that could inform their purchase decisions. Therefore we moved these factors back to the primary layer.

4) Factors to Exclude from the Label: There were six factors that at least 6 out of 7 experts believed should not be included on the label:

- List of device-compatible products
- Link to the device's software and hardware bill of material
- Link to the device's accompanying app(s)
- Whether or not the device manufacturer has a bug bounty program

- Where and when the device brand was incorporated
- Consumer Reports rating

The most common reasons experts said these factors were not suitable for the label were the lack of relevance to privacy and security and inability to convey risk to consumers. For example, S2 did not want the label to include whether or not the manufacturer has a bug bounty program as this factor does not offer adequate insight into security practices of the company: "This information is not too important on how the company does security analysis."

Almost all experts were opposed to including the Consumer Reports rating, mainly as they believed this organization's reputation does not stem from their privacy and security assessments. However, Consumer Reports is in the process of developing a digital privacy and security standard [22], so this may change in the future.

We decided to include whether or not the device manufacturer has a bug bounty program on the secondary layer. Since the word "bug bounty" was not immediately clear to consumer participants, we changed the wording to vulnerability disclosure and management, which was more understandable to them. When we presented this factor to our participants, 13 out of 15 associated this information with having good privacy and security practices, hence they were more inclined to trust the company who were transparent about their devices' discovered vulnerabilities and had taken steps to manage them. C3 explained: "This factor shows that this is a company that has a security process, and participates in public activities to educate the community on things that can go wrong." C5 wanted the IoT companies to disclose their devices' history of known vulnerabilities: "A lot of times, if the company had some kind of vulnerability, they maybe want to sweep it under the rug and not let anyone know about it. That's good that it shows you that they're being honest about what issues they've had in the past, and what they've done to address them."

We decided to include a link to the software and hardware bill of materials (mentioned on the label as software and hardware composition list) on the secondary layer since it can provide useful information related to security when it is available. When we asked consumers about this factor, most wanted it to be included and noted that even if they did not understand it, it could be useful to those with technical expertise.

Based on our experts' opinions, we initially excluded the list of device-compatible products from the label. In the consumer interview study, we asked participants to tell us about anything they thought was missing from the label that they would like us to add. The only factor that participants suggested was a link to the privacy statement of device-compatible products such as Alexa. As this was suggested by a couple of our early participants, we added a factor on the secondary layer to list compatible platforms with a link to their privacy policies.

C. Attitudes toward Labels and Layered Design

All consumer participants discussed how difficult it is for them currently to find information related to privacy and security of smart devices before purchasing them. They all reported that they would like to have an IoT security and privacy label available at the point of sale, mainly to be as informed as possible.

Most experts mentioned that IoT privacy and security labels are useful to inform consumers when making purchase decisions, which is in line with prior work [37]. P7 explained that a label can provide consumers with information they would not have otherwise:

> What's good about a label is that it empowers the consumer to make a more active decision about cybersecurity rather than just being completely helpless as to what the security of her device might be. Especially as more and more of this technology is designed for consumers, the average consumer doesn't have a privacy, security, or a legal department to review this stuff before they buy it. Enterprises do, but consumers do not, so someone's gotta be looking out for consumers and giving the consumers this information.

All consumer participants were familiar with layered labels, as they had seen QR codes on products such as food, drugs, or video games. Most of our participants (11/15) expressed positive attitudes toward the layered design, mainly because the amount of information that could fit in two layers would not fit on a single-layer label. Participants also appreciated the ability to easily gain further insights about the privacy and security practices of the device manufacturer. These participants reported that they engage in a combination of online and in-store shopping. Hence, they believed the layered label design would be useful to them throughout their purchase process.

Some consumer participants (4/15) thought a layered label would not be ideal, citing the inconvenience of using a phone to scan the QR code when shopping in a store, especially for the elderly. C15 explained: "These technologies for older generation, they are kind of tough. The idea of installing something to scan the QR code, it's going to be too much for them. I know they prefer to just read everything, put on their glasses and read everything line by line." Two participants expressed concern that companies might withhold important information from the primary layer and put it only in the secondary layer.

For each of the factors on the label, we asked consumer participants to tell us how they believe that factor would impact their purchase decisions and whether they would like us to remove the factor or add additional details to the factor. All participants understood the factors presented on the primary layer and were able to discuss the potential risks associated with all of these factors except the level of detail for data storage. Although they all understood the terms "identifiable" and "anonymous," participants did not associate identifiable data with risk in this context, perhaps because the utility of a security camera is increased if it can record videos in which people are identifiable. Further testing is needed to understand the impact of the interaction between the purpose of data collection and the granularity of data on consumers' privacy concerns.

As we expected, some of the information on the secondary layer did not convey privacy and security risks to consumer participants. However, participants still asked to see most of the factors that we included on the secondary layer because they wanted to be as informed as possible when purchasing a smart device.

Participants mentioned that they might search online for information about unfamiliar factors and the availability of our label would help them. C5 explained: "I don't know what TCP and UDP are. But it's interesting to have this here, because then I could go to Reddit and ask on there what that means and what the capabilities are."

There were only six factors that 2 or 3 consumer participants thought should be removed from the label. These secondary-layer factors were perceived by those participants as lacking relevance to privacy and security (physical actuations, hardware and software bill of material), not understandable (MUD compliant, key management protocols, open network ports), and not relevant to them (special data handling practices for children).

At least one consumer participant recognized that some of the factors in the secondary label might be useful to experts. C3 explained:

> Labels are both for customers and for experts such as tech journalists, consumer advocacy groups, who are capable of understanding it and who will click on the things, and if they see something that is questionable will raise it in the public press, will raise it with regulatory authorities, and otherwise. The label is not just for the consumer, but also there's another feedback process that works through experts to the extent that the information is available at all.

D. Prototype Privacy and Security Label

We used the results from our expert elicitation and consumer studies as well as recent IoT security standardization and certification efforts to inform the design of a prototype privacy and security label for a hypothetical smart security camera. Our design has primary and secondary layers, as shown in Figures 3 and 4, respectively. The secondary layer includes plus signs next to each item that can be clicked to reveal further details. We envision that the secondary layer would be accompanied by a computer-readable version of the label to enable automated processing and comparison between products, for example by personal privacy assistants [20] or search engines [24], [36]. Our website at www.iotsecurityprivacy.org has the latest label design.

V. DISCUSSION

Expert participants recommended including privacy and security star ratings on an IoT label, mostly because ratings would help consumers to more easily compare IoT devices based on their privacy and security practices. All of our consumer participants liked the idea of having privacy and security assessments from trustworthy organizations. Although
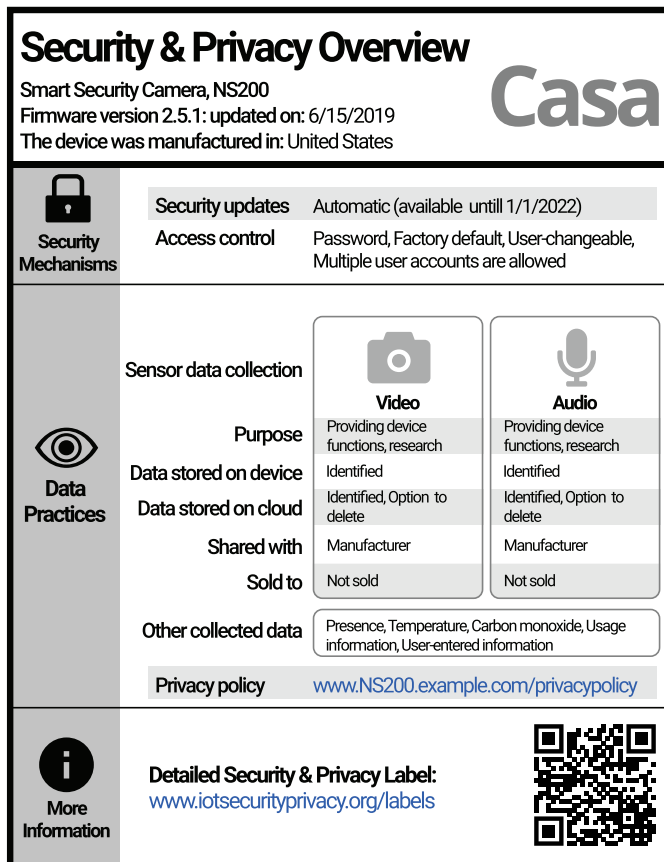

Fig. 3: Primary layer of the label. This layer is designed to be printed on product packaging or to appear on a product website. View our latest label design at www.iotsecurityprivacy.org.

we believe these third-party assessments would inform consumers' purchase behavior, we decided not to include them on our proposed label, as there is no organization currently doing these evaluations at scale for a wide range of IoT devices. We expect to add a place for assessment information once it is available.

We begin this section by providing a comparison between certifications and star ratings. Next, we discuss possible approaches to IoT privacy and security certifications.

A. Star Ratings vs. Certification Levels

Similar to the Energy Star rating system managed by the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Energy (DOE) [78], the idea of star ratings has been proposed for IoT devices to help consumers make informed purchase decisions [17], [77]. In a hearing of the U.S. Senate Committee on Commerce, Science, and Transportation's Subcommittee on Security [101], Senator Ed Markey suggested a 5-star security rating system for IoT products. In our study, while experts were supportive of privacy and security ratings on the label, they also mentioned two potential challenges of including them.

The first challenge experts brought up relates to the rating scale. Experts suggested that consumers might have trouble distinguishing a large number of ratings, yet a more granular

Security & Privacy Details

Smart Security Camera, NS200
Firmware version 2.5.1, updated on: 6/15/2019
The device was manufactured in: United States

Casa

Security Mechanisms

Security updates	Automatic (available until 1/1/2022)
Access control	Password, Factory default, User-changeable, Multiple user accounts are allowed
Security oversight	Audits performed by internal security auditors
Ports and protocols	www.NS200.example.com/port
Hardware safety	www.NS200.example.com/hwsafety
Software safety	www.NS200.example.com/swsafety
Personal safety	www.NS200.example.com/usersafety
Vulnerability disclosure and management	www.NS200.example.com/vulreport
Software and hardware composition list	www.NS200.example.com/BOM
Encryption and key management	www.NS200.example.com/key

Data Practices

	Video	Audio	Presence	Temperature	Carbon Monoxide
Sensor data collection					
Collection frequency	When user requests it	Continuous, Adjustable	Periodic, Option to opt-out	Continuous, Option to opt-in	Continuous, Option to opt-out
Purpose	Providing device functions, research	Providing device functions, research	Providing device functions	Providing device functions	Providing device functions
Data stored on device	Identified	Identified	De-identified	De-identified	De-identified
Local data retention time	Up to a month	Up to a month	Up to a year	Up to a year	Up to a year
Data stored on cloud	Identified, Option to delete	Identified, Option to delete	No cloud storage	De-identified	De-identified
Cloud data retention time	Up to a month	Up to a month	No cloud storage	Up to a month	Up to a month
Shared with	Manufacturer	Manufacturer	Not shared	Manufacturer, Third-party	Third-party, option to opt-out
Sharing frequency	Periodic, Adjustable	Periodic, Adjustable	Not shared	Continuous	Continuous
Sold to	Not sold	Not sold	Not sold	Third-party	Third-party, Option to opt-out

Other collected data	Usage information, User-entered information
Data linkage	Data may be linked with internal and external data sources
What could be inferred from user's data	No data inference
Special data handling practices for children	Yes
In compliance with	GDPR, ISO27001
Privacy policy	www.NS200.example.com/privacypolicy

More Information

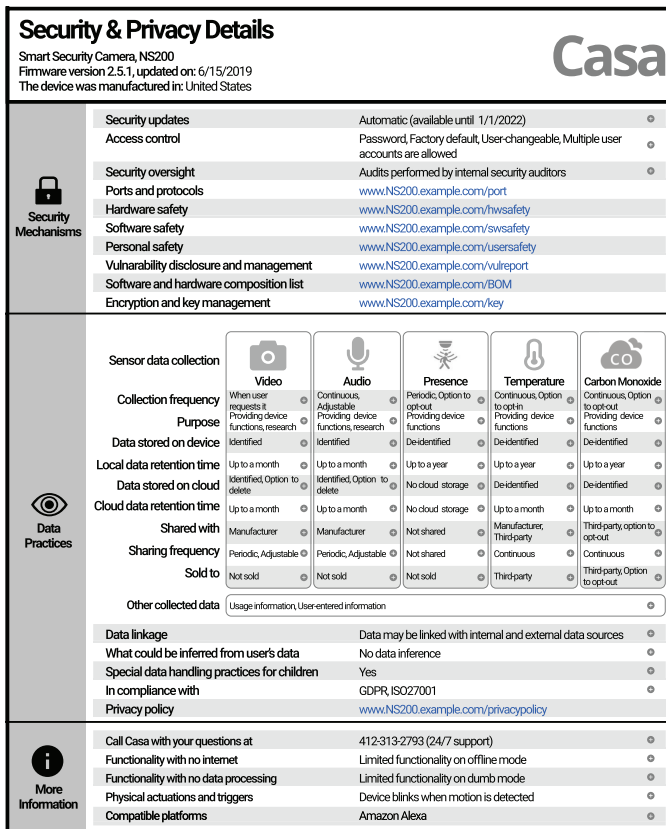|	
Call Casa with your questions at	412-313-2793 (24/7 support)
Functionality with no internet	Limited functionality on offline mode
Functionality with no data processing	Limited functionality on dumb mode
Physical actuations and triggers	Device blinks when motion is detected
Compatible platforms	Amazon Alexa

Fig. 4: Secondary layer of the label, which can be accessed by scanning the QR code or typing the URL on the primary layer. View our latest label design at www.iotsecurityprivacy.org.

scale could help manufacturers better differentiate their privacy and security practices. P1, who works in industry, discussed this issue: "I'm sure industry people, manufacturers, will want more in there. What would happen if you had something like this is it might start to grow based on features they want reflected in that rating. Then I can see it becoming a bigger and bigger scale."

Experts mentioned that ratings might pose an unhealthy incentive for IoT companies to achieve full-star ratings only to be able to compete in the market. Companies may be able to game the ratings in order to get all the stars and eventually all products will have all stars, whether they deserve them or not. S2, an academic, explained: "The problem I have with ratings like this is that everybody's gonna get a five star, because everybody's gonna figure out how to get the five star."

To address these challenges, some experts discussed the idea of having multiple certification levels (e.g., silver, gold, platinum) with a secure baseline or minimum standard instead of star ratings. This is similar to what the LEED standards use for rating energy efficiency and sustainability of buildings [57]. P8 explained: "I think consumers should know it passes the minimum security level. If I'm buying a space heater, I know they're not allowed to sell me one that will set on fire. I don't have to say, oh, it has a 70% score that it will set the house on fire."

Underwriters Laboratories (UL) published a 5-level IoT security standard (bronze, silver, gold, platinum, and diamond) in 2019 [99]. As of January 2020, no devices have been certified [100]. As manufacturers start having their devices certified, this certification could be added to the IoT label.

Since the lowest certification level indicates a safe device, there is a risk that manufacturers will aim to achieve the lowest level and not bother pursuing higher levels. Market competition may encourage manufacturers to pursue higher certification levels, especially for devices where the consequences of security breaches are most severe.

B. Privacy and Security Evaluation and Scoring

Over the past few years, a number of organizations and research teams have started to develop standards for IoT privacy and security evaluation and scoring. They include Consumer Reports [22], YourThings [3], [107], and UL [99].

1) Digital Standard: In 2017, Consumer Reports launched the Digital Standard to work toward providing a comprehensive standard that enables organizations to evaluate consumer IoT products. This standard focuses on four categories: security, privacy, ownership, and governance & compliance [22].

The security category of the Digital Standard includes build quality, data security, and personal safety.

Build quality refers to product stability and whether "software was built and developed according to the industry's best practices for security." The Cyber Independent Testing Lab (CITL) [26], a Digital Standard partner, is actively evaluating and scoring software of IoT devices according to a number of factors. Our label design includes a software safety features element where manufacturers can provide a URL with information related to software security.

Data security includes authentication, encryption, updatability, security audits, and vulnerability disclosure program. All of these factors are included on our label.

The personal safety category has not yet been defined in the Digital Standard, although developer notes indicate it will be related to avoiding abuse and harassment. Media reports suggest there are many incidents involving smart home devices being used for domestic abuse [13]. However, device manufacturers appear to be doing little to address the risks associated with abuse involving their devices [94]. We have included a factor called personal safety, which provides a place where device manufacturers can indicate available safeguards against abusive behavior once such safeguards have been implemented. Further discussions with experts are needed to determine how to address significant safety issues effectively on the label. As it was explained by S4: "Safety means if your car gets hacked, you die. The room that has a laser attached and if it gets hacked, it kills you. A drone can be reprogrammed to dive-bomb your child. I'm not sure how to capture that on the label."

The privacy section in the Digital Standard includes user controls, data use and sharing, data retention, and overreach. The assessment procedure for almost all the privacy factors in the Standard involves verifying the company's claimed data practices with actual data practices.

All the privacy factors mentioned in the Digital Standard are covered in our proposed label, except overreach. Overreach, or "collecting too much data" focuses on determining whether data collection is beneficial to the user, fully disclosed, the minimum necessary for functionality, and private by default. This seems like an area where a third-party assessment rather than a self report is likely warranted.

As some of the experts we interviewed mentioned, consumers may weigh privacy and functionality trade-offs differently. Thus it may be difficult to capture a single privacy rating that makes sense for all consumers. In addition to providing detailed information about data practices, a future privacy rating system could be customized based on a consumer's stated privacy preferences, which could change over time.

2) YourThings: Alrawi et al. [3] developed a security evaluation and scoring method for smart home devices. In their YourThings [107] initiative, they produce device scorecards with grades in four areas: device, mobile application, cloud endpoints, and network communication. While our label provides information related to all of the major areas of the YourThings rubric as well as some security and privacy factors not addressed by YourThings, the YourThings scorecard considers some additional security details, including some that do not lend themselves to self report. The YourThings scorecard offers a concise expert summary of device security issues, which could be useful to include on the label.

The YourThings rubric considers five device-related factors: upgradability, exposed services, vulnerabilities, configuration, and Internet pairing. We include all on our label.

The mobile section of the YourThings rubric includes sensitive data, programming issues, and "over-privileged," i.e., requesting excess permissions that are not used or required. Sensitive data is defined in the rubric to include "artifacts like API keys, passwords, and cryptographic keys that are hard-coded into the application." Our label includes factors related to sensitive data and programming issues such as software safety features and key management protocol. However, over-privileged is a factor better assessed by a third-party evaluator rather than being self-reported.

The cloud endpoints section of the rubric includes domain categories, TLS configuration, and vulnerable services. Some of the information needed to compute this score is included on our secondary layer of the label when fully expanded. However, details needed to evaluate TLS configuration are not included.

Finally, the network communication section of the rubric includes protocols, susceptibility to Man in the Middle (MITM) attack, and use of encryption. While the secondary layer of our label provides some of the information needed to compute this score, a third-party evaluation would be needed to provide a complete assessment.

The concise YourThings scores are useful for comparing devices, but users may need to drill down to obtain information relevant to their specific needs. For example, devices are penalized for not having automatic updates. While automatic updates are generally considered the most secure approach,

poorly timed updates can be problematic, potentially interfering with critical device functions.

3) UL: The 5-level UL certification process includes 44 requirements over seven categories: software updates, data & cryptography, logical security, system management, customer identifiable data, protocol security, and process & documentation [98].

While our proposed label includes factors from all seven categories, a third-party evaluation is needed to assess compliance with requirements. Our label can inform consumers about security and privacy, and goes into more detail about privacy issues than UL's customer identifiable data category. By including the UL certification in our label, we would offer users a single concise assessment of device security that complements the more detailed information provided on the label.

VI. CONCLUSION

We conducted a study with 22 privacy and security experts to elicit their opinions on the contents of IoT privacy and security labels. By following a three-round Delphi method, we found the factors that experts believed should be included on the label, and distributed them between primary and secondary layers of the label in three categories (security mechanisms, data practices, and general/more information). By conducting a series of in-depth semi-structured interviews with 15 IoT consumers, we iteratively improved the design of our proposed privacy and security label for IoT devices. Additional user tests should explore how consumers use the label in context as well as identify areas where wording or design can be further improved. Finally, we plan to develop a glossary of terms for consumers and an implementation guide for device manufacturers with detailed definitions of each factor and its possible values. The latest version of the label and implementation information, as well as a tool for generating the label are available at www.iotsecurityprivacy.org.

REFERENCES

[1] A. Acquisti, L. Brandimarte, and G. Loewenstein. Privacy and human behavior in the age of information. *Science*, 347(6221):509–514, 2015.

[2] M. Adler and E. Ziglio. *Gazing into the oracle: The Delphi method and its application to social policy and public health*. Jessica Kingsley Publishers, 1996.

[3] O. Alrawi, C. Lever, M. Antonakakis, and F. Monrose. SoK: Security evaluation of home-based IoT deployments. In *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2019.

[4] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, et al. Understanding the Mirai botnet. In 26th *USENIX Security Symposium (USENIX Security 17)*, pages 1093–1110, 2017.

[5] O. Arias, J. Wurm, K. Hoang, and Y. Jin. Privacy and security in internet of things and wearable devices. *IEEE Transactions on Multi-Scale Computing Systems*, 1(2):99–109, 2015.

[6] C. R. Atherton. Group techniques for program planning: A guide to nominal group and Delphi processes. by André L. Delbecq, Andrew H. Van de Ven, and David H. Gustafson. Glenview, Ill.: Scott, Foresman & Co., 1976.

[7] K. Backor, S. Golde, and N. Nie. Estimating survey fatigue in time use study. In *international association for time use research conference. Washington, DC*. Citeseer, 2007.

[8] R. Balebako, C. Bravo-Lillo, and L. F. Cranor. Is notice enough: Mitigating the risks of smartphone data sharing. *ISJLP*, 11:279, 2015.

[9] R. Balebako, J. Jung, W. Lu, L. F. Cranor, and C. Nguyen. Little brothers watching you: Raising awareness of data leaks on smartphones. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, page 12. ACM, 2013.

[10] J. Blythe and S. Johnson. The consumer security index for IoT: A protocol for developing an index to improve consumer decision making and to incentivize greater security provision in IoT devices. 2018.

[11] J. M. Blythe, N. Sombatruang, and S. D. Johnson. What security features and crime prevention advice is communicated in consumer IoT device manuals and support.

[12] B. Bonné, S. T. Peddinti, I. Bilogrevic, and N. Taft. Exploring decision making with Android's runtime permission dialogs using in-context surveys. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 195–210, 2017.

[13] N. Bowles. Thermostats, locks and lights: Digital tools of domestic abuse. https://www.nytimes.com/2018/06/23/technology/smart-home-devices-domestic-abuse.html, June 2018.

[14] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

[15] J. Caltrider. 10 fascinating things we learned when we asked the world "how connected are you?". https://goo.gl/92JDfq, November 2017.

[16] Centre for International Governance Innovation & IPSOS. 2016 CIGI-Ipsos global survey on internet security and trust. https://www.cigionline.org/internet-survey-2016, 2016.

[17] R. Chirgwin. Australia's IoT security rating might work, if done right. https://www.theregister.co.uk/2017/10/17/iot_security_rating_it_can_work_if_done_right/, October 2017.

[18] Cisco. Building trust and value in the data exchange between people, things and providers. https://www.jasper.com/sites/default/files/cisco_iot_survey_-_the_value_trust_paradox_final-1_2.pdf, December 2017.

[19] S. W. Cochran. The Delphi method: Formulating and refining group judgements. *Journal of Human Sciences*, 2(2):111–117, 1983.

[20] J. Colnago, Y. Feng, T. Palanivel, S. Pearman, M. Ung, A. Acquisti, L. F. Cranor, and N. Sadeh. Informing the design of a personalized privacy assistant for the internet of things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM.

[21] A. R. Colson and R. M. Cooke. Expert elicitation: using the classical model to validate experts' judgments. *Review of Environmental Economics and Policy*, 12(1):113–132, 2018.

[22] Consumer Reports. The standard. https://www.thedigitalstandard.org/the-standard.

[23] L. F. Cranor. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *J. on Telecomm. & High Tech. L.*, 10:273, 2012.

[24] L. F. Cranor, P. G. Leon, and B. Ur. A large-scale evaluation of us financial institutions' standardized privacy notices. *ACM Transactions on the Web (TWEB)*, 10(3):17, 2016.

[25] S. Crites Jr and S. N. Aikman. Impact of nutrition knowledge on food evaluations. *European journal of clinical nutrition*, 59(10):1191, 2005.

[26] Cyber Independent Testing Lab. How we evaluate. https://cyber-itl.org/.

[27] F. R. Cyphert and W. L. Gant. The Delphi technique: A case study. *Phi Delta Kappan*, 52(5):272–273, 1971.

[28] A. L. Dailey and J. C. Holmberg. Delphi-a catalytic strategy for motivating curriculum revision by faculty. *Community/Junior College Quarterly of Research and Practice*, 14(2):129–136, 1990.

[29] N. Dalkey and O. Helmer. An experimental application of the Delphi method to the use of experts. *Management science*, 9(3):458–467, 1963.

[30] Department for Digital, Culture, Media and Sport. Secure by design: Improving the cyber security of consumer internet of things report. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/686089/Secure_by_Design_Report_.pdf, March 2018.

[31] Department for Digital, Culture, Media and Sport. Mandating security requirements for consumer IoT products. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/798722/Secure_by_Design_Consultation_Stage_Regulatory_Impact_Assessment.pdf, May 2019.

[32] Department for Digital, Culture, Media and Sport. Plans announced to introduce new laws for internet connected devices. https://www.gov.uk/government/news/plans-announced-to-introduce-new-laws-for-internet-connected-devices, May 2019.

[33] A. K. Draper, A. J. Adamson, S. Clegg, S. Malam, M. Rigg, and S. Duncan. Front-of-pack nutrition labelling: Are multiple formats a problem for consumers? *The European Journal of Public Health*, 23(3):517–521, 2011.

[34] A. C. Drichoutis, P. Lazaridis, and R. M. Nayga. Nutrition knowledge and consumer use of nutritional food labels. *European Review of Agricultural Economics*, 32(1):93–118, 2005.

[35] A. C. Drichoutis, R. M. Nayga, Jr, and P. Lazaridis. Can nutritional label use influence body weight outcomes? *Kyklos*, 62(4):500–525, 2009.

[36] S. Egelman, J. Tsai, L. F. Cranor, and A. Acquisti. Timing is everything?: The effects of timing and placement of online privacy indicators. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 319–328. ACM, 2009.

[37] P. Emami-Naeini, H. Dixon, Y. Agarwal, and L. F. Cranor. Exploring how privacy and security factor into IoT device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 534. ACM, 2019.

[38] European Commission. Proposal for a regulation of the European Parliament and of the Council on ENISA, the "EU cybersecurity agency", and repealing regulation (EU) 526/2013, and on information and communication technology cybersecurity certification ("Cybersecurity Act"). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017PC0477&rid=1, 2017.

[39] European Union Agency for Cybersecurity. ENISA Good practices for IoT and Smart Infrastructures Tool. https://www.enisa.europa.eu/topics/iot-and-smart-infrastructures/iot/good-practices-for-iot-and-smart-infrastructures-tool.

[40] Federal Trade Commission. Comment to national telecommunications and information administration. https://www.ftc.gov/policy/advocacy/advocacy-filings/2017/06/ftc-comment-national-telecommunications-information, June 2017.

[41] J. L. Fleiss, B. Levin, and M. C. Paik. *Statistical methods for rates and proportions*. John Wiley & Sons, 2013.

[42] C. Franklin and M. Ballan. Reliability and validity in qualitative research. *The handbook of social work research methods*, 4:273–292, 2001.

[43] Gartner. Internet of things endpoint spending worldwide by category from 2014 to 2020 (in billion U.S. dollars). https://www.statista.com/statistics/485252/iot-endpoint-spending-by-category-worldwide/, August 2018.

[44] Global System for Mobile Communications. IoT Security Assessment. https://www.gsma.com/iot/iot-security-assessment/.

[45] J. Granjal, E. Monteiro, and J. S. Silva. Security for the internet of things: a survey of existing protocols and open research issues. *IEEE Communications Surveys & Tutorials*, 17(3):1294–1312, 2015.

[46] Harris Interactive. Consumer internet of things security labelling survey research findings. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/798543/Harris_Interactive_Consumer_IoT_Security_Labelling_Survey_Report.pdf, January 2019.

[47] H. Hibshi. *Composite Security Requirements in the Presence of Uncertainty*. PhD thesis, 2018.

[48] H. Hibshi, T. D. Breaux, M. Riaz, and L. Williams. A grounded analysis of experts' decision-making during security assessments. *Journal of Cybersecurity*, 2(2):147–163, 2016.

[49] S. Hieke and C. R. Taylor. A critical review of the literature on nutritional labeling. *Journal of Consumer Affairs*, 46(1):120–156, 2012.

[50] Information Commissioner's Office. Energy efficiency. hhttps://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/11/blog-the-12-ways-that-christmas-shoppers-can-keep-children-and-data-safe-when-buying-smart-toys-and-devices/, November 2017.

[51] IoT Security Foundation. IoT security compliance framework. https://www.iotsecurityfoundation.org/wp-content/uploads/2018/12/IoTSF-IoT-Security-Compliance-Framework-Release-2.0-December-2018.pdf, December 2018.

[52] P. G. Kelley, L. Cesca, J. Bresee, and L. F. Cranor. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pages 1573–1582. ACM, 2010.

[53] E. V. Kleef and H. Dagevos. The growing role of front-of-pack nutrition profile labeling: a consumer perspective on key issues and controversies. *Critical reviews in food science and nutrition*, 55(3):291–303, 2015.

[54] A. B. Knol, P. Slottje, J. P. van der Sluijs, and E. Lebret. The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health*, 9(1):19, 2010.

[55] P. M. Kuhnert, T. G. Martin, and S. P. Griffiths. A guide to eliciting and using expert knowledge in bayesian ecological models. *Ecology letters*, 13(7):900–914, 2010.

[56] V. Lara. What the internet of things means for consumer privacy. https://perspectives.eiu.com/technology-innovation/what-internet-things-means-consumer-privacy-0/white-paper/what-internet-things-means-consumer-privacy, March 2018.

[57] Leadership in Energy and Environmental Design. Green building leadership is leed.

[58] D. Lee. Amazon Alexa heard and sent private chat. https://www.bbc.com/news/technology-44248122, May 2018.

[59] H. Lee and A. Kobsa. Understanding user privacy in internet of things environments. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, pages 407–412. IEEE, 2016.

[60] H. Lee and A. Kobsa. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 276–285. IEEE, 2017.

[61] P. G. Leon, B. Ur, Y. Wang, M. Sleeper, R. Balebako, R. Shay, L. Bauer, M. Christodorescu, and L. F. Cranor. What matters to users?: Factors that affect users' willingness to share information with online advertisers. In *Proceedings of the ninth symposium on usable privacy and security*, page 7. ACM, 2013.

[62] T. Lieu. H.R.4163: Cyber Shield Act of 2017. https://www.congress.gov/115/bills/hr4163/BILLS-115hr4163ih.pdf, October 2017.

[63] B. Ludwig. Predicting the future: Have you considered using the Delphi methodology. *Journal of extension*, 35(5):1–4, 1997.

[64] K. M. Macqueen, E. McLellan-Lemal, K. Bartholow, and B. Milstein. Team-based codebook development: Structure, process, and agreement. *Handbook for team-based qualitative research*, pages 119–135, 2008.

[65] A. Madill, A. Jordan, and C. Shirley. Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British journal of psychology*, 91(1):1–20, 2000.

[66] E. Markey. S.2020: Cyber shield act of 2017. https://www.congress.gov/115/bills/s2020/BILLS-115s2020is.pdf, October 2017.

[67] D. F. Marks and L. Yardley. *Research methods for clinical and health psychology*. Sage, 2004.

[68] N. McDonald, S. Schoenebeck, and A. Forte. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. 2019.

[69] T. Miaskiewicz and K. Kozar. The use of the Delphi method to determine the benefits of the personas method-an approach to systems design. *SIGHCI 2006 Proceedings*, page 7, 2006.

[70] L. M. S. Miller and D. L. Cassady. Making healthy food choices using nutrition facts panels. the roles of knowledge, motivation, dietary modifications goals, and age. *Appetite*, 59(1):129–139, 2012.

[71] Mozilla. Shop safe this holiday season. https://foundation.mozilla.org/en/privacynotincluded/, 2018.

[72] J. W. Murry Jr and J. O. Hammons. Delphi: A versatile methodology for conducting qualitative research. *The Review of Higher Education*, 18(4):423–436, 1995.

[73] 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)*, pages 399–412, 2017.

[74] National Telecommunications and Information Administration. Communicating IoT device security update capability to improve transparency for consumers. https://www.ntia.doc.gov/files/ntia/publications/draft-communicating_iot_security_update_0426.pdf, April 2017.

[75] R. M. Nayga Jr. Nutrition knowledge, gender, and food label use. *Journal of Consumer Affairs*, 34(1):97–112, 2000.

[76] A. Ng and M. Wollerton. Google calls Nest's hidden microphone an "error". https://www.cnet.com/news/google-calls-nests-hidden-microphone-an-error/, February 2019.

[77] S. Nielsen. Should connected devices carry an IoT security-star rating? https://internetofthingsagenda.techtarget.com/blog/IoT-Agenda/Should-connected-devices-carry-an-IoT-security-star-rating, May 2019.

[78] Office of Energy Efficiency & Renewable Energy. Energy star. https://www.energy.gov/eere/buildings/energy-star.

[79] Park Associates. Purchase intentions for smart home devices increased by 66% year over year. https://www.prnewswire.com/news-releases/parks-associates-purchase-intentions-for-smart-home-devices-increased-by-66-year-over-year-300632082.html, April 2018.

[80] PETRAS Internet of Things Research Hub. Developing a consumer security index for consumer IoT devices (CSI). https://www.petrashub.org/portfolio-item/developing-a-consumer-security-index-for-domestic-iot-devices-csi/, September 2018.

[81] J. Phelps, G. Nowak, and E. Ferrell. Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing*, 19(1):27–41, 2000.

[82] J. Saldaña. *The coding manual for qualitative researchers*. Sage, 2015.

[83] 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)*, pages 1–17, 2015.

[84] D. Schor, S. Maniscalco, M. M. Tuttle, S. Alligood, and W. R. Kapsak. Nutrition facts you can't miss: The evolution of front-of-pack labeling: Providing consumers with tools to help select foods and beverages to encourage more healthful diets. *Nutrition Today*, 45(1):22–32, 2010.

[85] J. Sen. Security and privacy issues in cloud computing. In *Cloud Technology: Concepts, Methodologies, Tools, and Applications*, pages 1585–1630. IGI Global, 2015.

[86] S. Sheng, P. Kumaraguru, A. Acquisti, L. Cranor, and J. Hong. Improving phishing countermeasures: An analysis of expert interviews. In *2009 eCrime Researchers Summit*, pages 1–15. IEEE, 2009.

[87] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini. Security, privacy and trust in internet of things: The road ahead. *Computer networks*, 76:146–164, 2015.

[88] V. Sivaraman, H. H. Gharakheili, A. Vishwanath, R. Boreli, and O. Mehani. Network-level security and privacy control for smart-home IoT devices. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2015 IEEE 11th International Conference on*, pages 163–167. IEEE, 2015.

[89] A. Speirs-Bridge, F. Fidler, M. McBride, L. Flander, G. Cumming, and M. Burgman. Reducing overconfidence in the interval judgments of experts. *Risk Analysis: An International Journal*, 30(3):512–523, 2010.

[90] N. Statt. Amazon's Alexa isn't just AI — thousands of humans are listening. https://www.theverge.com/2019/4/10/18305378/amazon-alexa-ai-voice-assistant-annotation-listen-private-recordings, April 2019.

[91] H. J. Strauss and L. H. Zeigler. The delphi technique and its uses in social science research. *The Journal of Creative Behavior*, 9(4):253–259, 1975.

[92] J. M. Szanyi. Brain food: bringing psychological insights to bear on modern nutrition labeling efforts. *Food & Drug LJ*, 65:159, 2010.

[93] L. Tanczer, J. Blythe, F. Yahya, I. Brass, M. Elsden, J. Blackstock, and M. Carr. Summary literature review of industry recommendations and international developments on IoT security. *PETRAS IoT Hub, Department for Digital, Culture, Media & Sport (DCMS)*, 2018.

[94] L. Tanczer, I. Lopez Neira, S. Parkin, T. Patel, and G. Danezis. Gender and IoT research report. https://www.ucl.ac.uk/steapp/sites/steapp/files/giot-report.pdf, November 2018.

[95] A. Taylor. It may have been terrifyingly easy for a stranger to hack into a baby monitor and curse at a 2-year-old. https://www.businessinsider.com/foscam-baby-monitor-hack-2013-8, August 2013.

[96] The Center for Information Policy Leadership. Ten steps to develop a multilayered privacy notice. https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/ten_steps_to_develop_a_multilayered_privacy_notice__white_paper_march_2007_.pdf, March 2007.

[97] C. Thorun, M. Vetter, L. Reisch, and A. K. Zimmer. Indicators of consumer protection and empowerment in the digital world. https://www.vzbv.de/sites/default/files/downloads/2017/03/13/conpolicy_executive_summary.pdf, March 2017.

[98] UL. Methodology for marketing claim verification: Security capabilities verified to level bronze/silver/gold/platinum/diamond, UL MCV 1376. https://www.shopulstandards.com/ProductDetail.aspx?UniqueKey=35953.

[99] Underwriters Laboratories. IoT security rating. empowering consumers through product security labeling. https://ims.ul.com/IoT-security-rating.

[100] Underwriters Laboratories. Verify. https://verify.ul.com/search.

[101] U.S. Senate Committee on Commerce, Science, & Transportation. Strengthening the cybersecurity of the internet of things. https://www.commerce.senate.gov/public/index.cfm/hearings?ID=A6113AB7-E89B-48C7-B555-E3CBB1466040, April 2019.

[102] W. Usher and N. Strachan. An expert elicitation of climate, energy and economic uncertainties. *Energy policy*, 61:811–821, 2013.

[103] D. Veen, D. Stoel, M. Zondervan-Zwijnenburg, and R. Van de Schoot. Proposal for a five-step method to elicit expert judgment. *Frontiers in psychology*, 8:2110, 2017.

[104] E. L. Vyth, I. H. Steenhuis, A. J. Roodenburg, J. Brug, and J. C. Seidell. Front-of-pack nutrition label stimulates healthier product development: A quantitative analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 7(1):65, 2010.

[105] W3C. The platform for privacy preferences 1.0 (p3p1.0) specification. https://www.w3.org/TR/P3P/#PURPOSE, April 2002.

[106] K. L. Walker. Surrendering information through the looking glass: Transparency, trust, and protection. *Journal of Public Policy & Marketing*, 35(1):144–158, 2016.

[107] YourThings. Evaluating and scoring smart-home devices to improve security! https://yourthings.info/about.

You are probably aware that food products have nutrition labels on them that tell you about all the different nutrition facts about them. However, when you buy an IoT device, there is no information like that. We are trying to develop a label for IoT devices that would focus on privacy and security aspects of those devices, similar to nutrition label for foods. Today I am going to talk to you about the things that should be on this privacy and security label for IoT devices.

Before asking about the label, I want you to please
Q: define security in the context of IoT devices.
Q: Can you also define privacy related to the IoT devices.

We are going to talk about what are the different factors that you would like to put on the label. Obviously there is a limited amount of space on the label, but for now don't worry about that and let's just talk about what factors you would like to see.

Q: Please specify the **security factors** that you would want to see on the label.
Q: For each factor please specify the **levels or values** that you think this factor should have.
Q: For each factor please specify the **choices** consumers may be provided with by the manufacturer.
(We asked the same questions for privacy and general factors)

Q: Among the factors you mentioned, please specify the ones that are good for experts but probably aren't needed for consumers.

(Presenting interviewee with the factors (privacy, security, and general) from the previous experts we have talked to)
Q: Please read the factors that other experts provided and compare them with the factors that you mentioned. Let me know if there is any factor, level, or choice that is currently not on your list, but you find it important for consumers to know about and want to add that to your list. Why do you think this should be added?
Q: In what order do you think we should present the sections of privacy, security, and general information to consumers? Why?

Fig. 5: Questions we asked in our semi-structured interviews with 22 experts. Based on the interviews, we found 47 factors that experts expressed interest in including on IoT privacy and security label.

Please answer the following questions and provide us with the reasons that are most convincing to you. The reasons can be in your own words or alternatively, you can quote any resource that you think provides a well-explained reason. Please note that to design the privacy and security label, we will follow a layered approach. In a layered design, we will put the most important/understandable information on the first/primary layer and the additional important information will go on the secondary layer in case consumers are curious to know more or have privacy and security expertise and are looking for more information (the secondary layer can get updated if the value of any factor is changed). Therefore, you may believe that it is important to include a specific factor somewhere on the label, but not necessarily on the first/primary layer. Please also do not worry about the physical space on the label and how many factors should be included on it.

Q: For each factor, please select the statement that you believe is correct.
- Is an important factor to include on the label.
- Is not an important factor to include on the label.
- May or may not be an important factor to include on the label

(If experts respond that the factor is important or may or may not be important to include)
Q: Please specify your most convincing reason(s) as to why this is an important factor to put on the label. (Open-ended question)

(If experts respond that the factor is not important to include)
Q: Please specify your most convincing reason(s) as to why this is not an important factor to put on the label. (Open-ended question)

(For each participant, we asked the same questions for one third of the factors)

Fig. 6: Questions we asked on the first survey. In this stage, we asked experts to review the factors from the interview study and specify their preferences about including each factor, as well as their reasons.

For each factor, please read the arguments for and against having that factor on the label (primary or secondary layer) and answer the questions. We selected the factors as well as the arguments from the experts we interviewed.
The label is not only to educate consumers to make informed purchase decisions, but also to hold IoT companies accountable for their privacy and security practices. In addition, all of the factors need to have a precise definition and metric, but for now do not worry about that and assume that the definition and metric is specified.

Including a factor on the label that describes **whether or not the device is getting cryptographically signed and critical automatic security updates:**
- **is important** because:
  - Consumers understand the importance of automatic security updates.
  - New vulnerabilities in software are discovered very often and hence continuous checking for updates is needed.
- **is not important** because:
  - Signed security updates do not guarantee a well implemented and trustworthy update process.
  - Sometimes there are reasons not to have automatic security updates. These reasons include legal considerations or if the device is incapable of receiving updates.
  - Automatic updates should be a requirement and there is no need to explicitly mention them.

Q: I believe this factor
- Should definitely be on the label
- Should probably be on the label
- I have no preference about this factor
- Should probably not be on the label
- Should definitely not be on the label

Q: On which layer of the label do you think this factor should be presented?
- primary/first layer (Please specify your reason(s)) [text field]
- secondary/extended layer (Please specify your reason(s)) [text field]
- I have no preference over this (Please specify your reason(s)) [text field]

Q: Information about this factor is for (select all that apply):
- Consumers' understanding
- Experts' understanding
- For holding manufacturers accountable, e.g., using regulations
- Other (Please specify) [Text field]

Q: If the label had separate security, privacy, and general sections, this factor would be most appropriate for:
- The security section
- The privacy section
- Either the privacy or the security section
- The general information section
- None (Please specify) [text field]

Q: Please provide any additional arguments you may have for or against including this factor on the label. [Text field]

Q: Please provide any comments you have about this factor. [Text field]

(For each participant, we asked the same questions for one third of the factors)
Q: I believe:
- privacy factors and security factors should be presented separately on the label (Please specify your reason(s)) [Text field]
- privacy and security factors should be combined and shown on the label without separation (Please specify your reason(s)) [Text field]
- I have no preference toward presenting privacy and security factors (Please specify your reason(s)) [Text field]

Fig. 7: Questions we asked on the second survey. Here we presented experts with common arguments from the first survey and asked them again to make decision about each and provide us with their additional arguments.

[first without the label, then with the label] I want you to please take a look at this smart security camera package. Please let me know what can you learn about the privacy and security of this device?

[Risk Communication in Comparative Purchase Process]

Let's say you were trying to decide whether to buy the security camera that I just showed you. The store has this other security camera, which is the same price and has the same features. However, it has slightly different privacy and security related information, which is now being presented on the label on the package. Please take a look at these two devices and tell me

- What are the things that this company has done better over the other one and why do you think they are better?
- What are the things that this company has done worse and why do you think they are worse?
- Overall, which device would you purchase? Why?
- What is the most useful piece of information on this label that helped you make this decision?

[explain the idea of a layered label] Please scan the QR code or type in the URL on the primary layer to take a look at the secondary layer of the label.

- Except the devices that I showed you in our study, have you ever seen a product that has a layered label like we describe?
- What is your opinion on the idea of a layered label? What could be the advantages and disadvantages of a layered label as opposed to a simple label that does not have layers?

[Information Comprehension in Noncomparative Purchase Process]

- For each presented factor in the security mechanisms/privacy practices/general information section, can you please explain what the factor means?
- How useful do you believe this factor would be in your purchase decision making?
- Do you have any suggestions to make the information about this factor more understandable?

[Risk Communication in Noncomparative Purchase Process]

- What are the factors that you see here that seem risky to you from a security or privacy perspective?
- [If participants specify a factor as being risky] What kinds of risks do you think you could be exposed to by the information conveyed by this factor? What bad things could happen as a result of this information?
- If you were looking for a camera that was like this but less risky, what would you like to see instead of the current information for this factor?
- If features, price, and the brand were all what you wanted, based on the privacy and security information of this primary layer of the label, would you purchase this device? Why?
- What are the most useful factors on this label that helped you make this decision?

[Information Comprehension and Risk Communication of the Secondary Layer]

- For each presented factor in the security mechanisms/privacy practices/general information section, can you please explain what the factor means?
- How useful do you believe this factor would be in your purchase decision making?
- Do you have any suggestions to make this information more understandable?

[Format and layout considerations]

- Overall, what do you think about the current design consideration of separating out security mechanisms, privacy practices, and general information?
- What are the changes you would like to apply to this label to make it more usable and understandable to you and other consumers? The suggestions could be both related to the content of the label or the design of the label. Please specify your reasons for the change you proposed.
- Is there anything else that you think would be useful to have on the primary layer of the label?
- Is there anything else that you think would be useful to have on the secondary layer of the label?

[Purchase Scenario Selection]

- If you were going to purchase a smart device from a physical store, where would you go?
- If you were going to purchase a smart device from an online store, where would you go?

Which one is the kind of process that you would be most likely to follow if you were to purchase a smart home device at brick & mortar? Why? When reading the scenarios, please replace the HomeDepot, with [physical store participants mentioned] and the online store with [online store participants mentioned].

1. You walk into HomeDepot and go to the smart home aisle. You find the products you are interested in and see the primary layers of the labels on the packages. You stand there looking at the packages, but you don't go online to take a look at the secondary layer.
2. You walk into HomeDepot and go to the smart home aisle. You find the products you are interested in and see a QR code/URL on the package. You stand there and scan the QR code/type in the URL and look up the information of the secondary layer on your phone.
3. Before going to HomeDepot, you go online and do the comparison shopping. You come to HomeDepot already knowing which device you will buy.
4. You do the comparison shopping online, find the product you are interested in and purchase the device from Amazon.
5. Something else? Please explain.

Fig. 8: After the expert elicitation study and specifying the content of the IoT privacy and security label based on experts' input, we conducted in-depth semi-structured interviews with 15 participants, who had purchased at least one smart home device. We asked interviewees questions to study how much they understand the content of our designed labels and how the presented information conveyed risk to participants.