SoK: (Un)usable Privacy: the Lack of Overlap between Privacy-Aware Sensing and Usable Privacy Research

Yasha Iravantchi
University of Michigan
Ann Arbor, Michigan, USA
yiravan@umich.edu

Pardis Emami-Naeini
Duke University
Durham, North Carolina, USA
pardis@cs.duke.edu

Alanson Sample
University of Michigan
Ann Arbor, Michigan, USA
apsample@umich.edu

Abstract

As the number of smart devices increases in our lives, the data they collect to perform valuable tasks, such as voice assistant requests, comes at the cost of user privacy. To mitigate their privacy impact, emerging usable privacy-aware sensing (UPAS) research has relied on cross-disciplinary approaches that extend past the core focus of broader academic research communities, such as Security & Privacy or Human-Computer Interaction. These works incorporate privacy design principles, whereby systems include safeguards by combining usable privacy (UP) with privacy-aware sensing (PAS) design to protect users' privacy. To better understand this emerging area of research, we conducted a mixed qualitative and quantitative Systematization of Knowledge (SoK). With a thorough review of pertinent literature, resulting in 114 selected works (reduced from 10,122 across 12 venues), we found that, despite the similarity of these works, many are dispersed across multiple communities, utilize community-specific jargon and keywords, and minimally overlap in design and evaluation approaches, potentially hindering cross-pollination across communities and thereby slowing the growth of this emerging research area. Thus, these factors helped reveal a research gap in this space. We use these findings to present four research themes and provide community and design recommendations to encourage cross-disciplinary UPAS research.

1 Introduction

Despite the many useful tasks Internet of Things (IoT) devices perform, our ever-increasing daily interactions with them come at a non-trivial cost to user privacy (i.e., a privacy footprint). For example, many voice agents (e.g., Alexa) process voice commands in the cloud, which has resulted in leaked or mishandled audio recordings of their users [80]. This loss in privacy can extend beyond the owner of the device to bystanders or anyone within the sensing range of the device, which has resulted in numerous privacy incidents [29, 45, 50, 80] that have fostered mistrust with these devices and, ultimately, can hamper their adoption despite their utility.

Efforts to reduce the privacy footprint of these devices while maintaining their usability have fostered a distinct area of focus for both Usable Privacy (UP) and Privacy-Aware Sensing (PAS) researchers. Research exploring Usable Privacy considers the human element within privacy and focuses on usability contributions that improve end-user privacy [138]. These researchers have explored

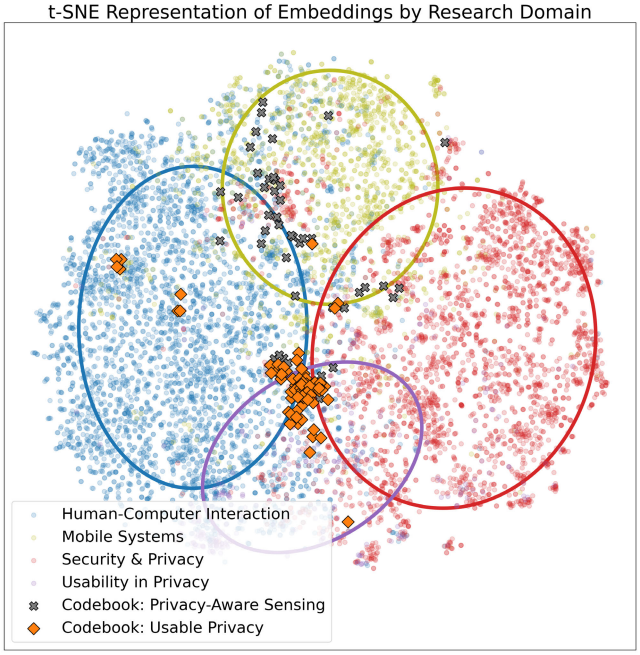

Figure 1: A t-SNE visualization of 10,122 works from the past 5 years spanning 12 relevant venues grouped into four broader research topics. The orange diamonds and gray crosses represent the Usable Privacy (UP) and Privacy-Aware Sensing (PAS) that form the 114 total works in the codebook. These works have three distinct clusters: one of UP works within Usability in Privacy, another of PAS works within the Mobile Systems community, and a small mixed cluster of UP and PAS works at the intersection of all four communities that denote emerging usable privacy-aware sensing research.

many avenues related to IoT sensing devices, including identifying contextual effects on users' perceptions of privacy [7], investigating approaches to mitigate their imposition in people's homes and living spaces by improving user control over devices [54], and providing user-accessible ways to examine and audit devices for whether they uphold user privacy preferences [97]. The Privacy-Aware Sensing community also has significant research investment in IoT devices in exploring mechanisms to restrict the amount and type of data sensors collect, such as rejecting speech content [60, 131]. These devices have been used in the home for ubiquitous and mobile computing applications, such as health and wellness monitoring [16, 17, 72, 131] and virtual assistant agents [4, 73, 102, 110]. Beyond collecting audio or visual data, the sensing community

Proceedings on Privacy Enhancing Technologies 2025(1), 472–490

https://doi.org/10.56553/popets-2025-0026

has explored utilizing other physical sensors (e.g., LIDAR [101], vibration [35], thermal [100], SA-waves [61]) to perform these valuable tasks in privacy-conscious ways that do not collect sensitive content such as speech or images of persons. These two distinct groups have independently contributed to improved user privacy.

The overlap between Usable Privacy and Privacy-Aware Sensing creates a natural pairing for cross-disciplinary Usable Privacy-Aware Sensing (UPAS) research to integrate usable privacy elements into hardware and system design, such that the user is aware of how the device operates and can control how the IoT device collects and stores information. Particularly established everyday examples of usable sensor privacy are devices that give users understandable notice and real control when their sensors operate (e.g., a webcam's illuminated LED or a mute button for a microphone). More recently, emerging UPAS devices include acoustic devices that allow users to select a reduction in information collection based on understandable human-centric amounts (e.g., filter out speech content or remove all human-audible content entirely) [60] or cameras that collect pixel data from only user-defined specific regions (e.g., remove bystanders from images) which can be visually audited [8], allowing users to find a compromise between a system's privacy and its utility and performance. Furthermore, this past year marked the first annual Sensors S&P workshop at SenSys 2023, highlighting a particular concern with "widespread deployment of networked sensors and the rapid advancements in black-box deep learning models capable of encoding large amounts of information" [1], demonstrating the emerging need for UPAS research.

This paper presents a Systematization of Knowledge (SoK) of usable privacy-aware sensing, aiming to bridge the gaps between UP and PAS research groups whose work often does not overlap within a single academic venue (i.e., conferences and journals) but rather span four broader academic venue groupings including Security and Privacy (S&P), Human-Computer Interaction (HCI), Mobile Systems (MS), and Usability in Privacy (UiP). This dispersion across many venues affects UP and PAS contributions, where they embed their academic venue's priorities and approaches towards improving privacy and how their privacy mitigations are evaluated—even affecting the language and jargon used to describe their contributions. These differences make it challenging for UP and PAS researchers from different academic communities to find each other and their work despite working in similar spaces, ultimately reducing intellectual cross-pollination and hindering collaborations that would lead to cross-disciplinary UPAS contributions. Based on an approach common to other SoK works [104, 109, 116, 126], we utilized our domain expertise as authors in both the UP and PAS spaces to identify and perform a detailed analysis of 114 selected works from a pool of 10,122 works across 12 academic venues. Based on these works, we identify key technical and methodological gaps and propose a framework for integrating user-centric privacy principles into sensor design and evaluation.

Our analysis revealed significant gaps in the existing literature on usable privacy-aware sensing. We identified four major themes: (1) the overly heavy reliance on user control and awareness to develop privacy mitigations in Usable Privacy (UP) contributions, (2) the tension between system functionality, user acceptance, and privacy in Privacy-Aware Sensing (PAS) contributions, (3) the prevalence of unvalidated assumptions about user needs in PAS research, and (4) the lack of feasibility evaluations in UP contributions.

Using statistical analysis, we found that beyond the works analyzed as part of our codebook, there is a shallow overlap between the four stakeholder academic venue groupings that would facilitate cross-disciplinary research. For example, while HCI and S&P overlap with UiP and MS, neither significantly overlaps with the other, nor does MS with UiP. Within our t-SNE visualization, shown in Figure 1, we identified a literal research gap representing a decrease in the density of works where the four communities would meet—where we expected and found UP and PAS works incorporating cross-disciplinary approaches. Interestingly, the vast majority of works in our codebook lie at the intersection of multiple communities, highlighting the importance of cross-disciplinary research needed to address this gap and encourage UPAS contributions.

Drawing from our comprehensive analysis, we propose the following recommendations to bridge the research gap between Usable Privacy (UP) and Privacy-Aware Sensing (PAS): (1) Reduce the siloing effect within each community by increasing the diversity of privacy design principles that support privacy mitigations (2) Conduct end-to-end evaluations that include both user-centric and technical assessments to ensure proposed privacy mitigations are both effective and user-friendly. (3) Engage with broader regulatory, ethical, and societal contexts to ensure that privacy solutions align with legal frameworks and address societal concerns. This holistic approach will facilitate the development of privacy-aware sensing technologies that are both practical and acceptable to users.

This paper makes the following contributions:

(1) We proposed a novel approach to finding relevant works for analysis as part of a Systematization of Knowledge contribution;
(2) Through those relevant works, we systematize two individual research domains: Privacy-Aware Sensing and Usable Privacy;
(3) Through a qualitative investigation of those communities' research, we identified four distinct themes that reveal a research gap between the two in Usable Privacy-Aware Sensing research;
(4) Through a quantitative analysis of those communities' research, we offered further evidence of gaps in the research domain;
(5) Grounded in our qualitative and quantitative findings, we provided recommendations for the community to address this research gap based on existing research in this domain.

2 Background

As the number of sensing devices has increased in our daily lives, people who interact with them, intentionally or unintentionally, have become increasingly aware of the privacy implications of the data collected by their onboard sensors, such as cameras and microphones. Furthermore, users have expressed complex privacy needs that are highly dependent on where (e.g., in public or at home) or what kind (e.g., speech/video vs. thermostat information) of data they collect. There is no "one-size-fits-all" solution to sensor privacy. Various academic communities have explored how to mitigate users' concerns but have taken different approaches derived from practices and norms within their academic community. For the purposes of this work, we contextualize our definition of *Usable Privacy-Aware Sensing* (UPAS) with two historical ubiquitous computing works.

2.1 *The Implications of Always-on Sensing*

Always-on sensing approaches are ones where a device in the environment, or more recently wearable on the body, continuously collects sensor information to perform a particular task. For example, a voice assistant (e.g., Amazon Echo) continuously processes microphone data to identify a wakeword such as "Alexa" to transmit captured audio information to a server for processing transcription and request tasks. Historically, always-on approaches have significant roots within ubiquitous computing. In *The Computer for the 21st Century* (1991) [124], Mark Weiser describes a future where computers disappear into the background yet are ever-present and ready to respond to our requests. He presents a fictional account to motivate the numerous ways ubiquitous computing can assist daily life. For these ubiquitous systems to be helpful, they must always remain on; they cannot proactively meet user needs if the user needs to turn them on. However, in the same work, Weiser highlights the central issue with always-on sensing:

"Perhaps key among them is privacy: hundreds of computers in every room, all capable of sensing people near them and linked by high-speed networks, have the potential to make totalitarianism up to now seem like sheerest anarchy." (pg. 9)

Thus, even at the dawn of ubiquitous computing, researchers looking to the future were aware of the significant privacy implications of always-on sensing. To wit, Weiser prescribes a possible remedy:

"Fortunately, cryptographic techniques already exist to secure messages from one ubiquitous computer to another and to safeguard private information stored in networked systems. If designed into systems from the outset, these techniques can ensure that private data does not become public." (pg. 10)

While data encryption has become near-ubiquitous, becoming invisible layer of protection, for encryption to improve privacy, it relies on the assumption that only the user has ownership and control of their private encrypted data and who can see and access it [68]. The binary concept of "private" data and "public" data that maps to either "only I or people I allow can see it" vs. "everyone can see it" no longer holds: Are the images a robot vacuum collects (including while someone is using the toilet [50]) that are annotated by humans to better train obstacle avoidance models "private" or "public"? How about the audio recordings stored by voice agents [80] or the information they pass on [59] to databrokers? Or the doorbell video feeds that can be accessed by law enforcement [29] without user consent or awareness? There is an emerging grey area where "private" data can be accessed by those who are not the individual who purchased the sensing device or those explicitly granted permission by that individual. Yet, in all the aforementioned situations, the data was "secure," encrypted through transmission, and accessed by "authorized" users; there was no data breach here, and data was never made public, so it should satisfy Weiser's argument. So why do all of these situations feel like an invasion of privacy?

Privacy scholars have criticized this binary concept of "private" and "public" data, as it does not account for these situations; for example, Nissenbaum's Privacy as Contextual Integrity [85] argues that the information flow has a significant role in user data privacy, where five critical parameters must be satisfied or otherwise can lead to ambiguities that can result in privacy issues. Specifically, in the abovementioned cases, is that "private" data was accessed by others deemed "authorized." Still, the full extent of who is "authorized" is often nebulous given how these ubiquitous systems are constructed. Thus, per Nissenbaum, both the recipient of the data and the transmission principle are left ambiguous, which has led to privacy issues. Furthermore, this "grey-area privacy" does not match the user's mental model that "their" data is "theirs": consumers often believe that since they own the device, they, by extension, own the data the device generates [62]. Furthermore, consumers have a perception that they are entitled to control the data it generates since they assume that they do own the data [62]. Thus, the academic community has looked towards developing additional ways to preserve user privacy through a more modern understanding of user needs while maintaining the usability and utility of these always-on systems.

2.2 *Usable Privacy-Aware Sensing*

Beyond the cryptographic approaches mentioned above, there is an extensive multi-disciplinary body of work exploring ways to improve privacy with specific concerns around these always-on sensing devices. While individual communities (e.g., HCI, S&P, UiP, MS) have focused on specific aspects of improving the privacy surrounding sensing devices, in this work, we guide our focus toward better understanding emerging UPAS approaches. As a lens to analyze this body of work, we use a design framework from Langheinrich's *Privacy by Design - Principles of Privacy-Aware Ubiquitous Systems* [68], which provides a comprehensive introduction to the privacy challenges these ubiquitous, always-on sensing systems present. Importantly, Langheinrich presents six principles for these systems, which have been well-adopted across multiple academic venues and remain highly relevant, that empower individual users to have a strong role in maintaining their privacy. To summarize, they are as follows:

(1) **Notice:** Systems should notify the user when collecting data.
(2) **Choice and Consent:** Systems should allow users to choose and consent to data collection.
(3) **Anonymity and Pseudonymity:** Systems should collect data in a way that cannot be traced back to the user.
(4) **Proximity and Locality:** Systems should confine their sensing ability to within a user-specified range or locations.
(5) **Adequate Security:** Systems should employ adequate measures to secure and protect a user's sensitive information.
(6) **Access and Resource:** Systems should only collect data sufficient for a well-defined purpose, which the user can audit.

These principles argue that although many of the operations performed by sensing systems may happen outside users' awareness, the mechanisms that protect their privacy should be easily accessible and, importantly, usable. These principles are particularly valuable to use as a lens to analyze Usable Privacy-Aware Sensing works since they cover many of the privacy-aware approaches preferred by the individual communities that contribute to this academic space. For example, many works that appear at USENIX Security and IEEE S&P contribute methods to improve user *Anonymity* and *Security*, while works that appear at CHI and SOUPS

often contribute to improved mechanisms surrounding *Notice* and *Choice* for users of sensing devices.

While each principle can strongly improve user privacy, they are complementary and, when combined, can holistically enhance user privacy. Ideally, designers should incorporate as many privacy design principles as possible within reason. Thus, we define *Usable Privacy-Aware Sensing* as any work that contributes to improving user privacy surrounding sensing devices by incorporating any of those principles. This includes identifying ways to interpret these principles as design recommendations for novel contexts or developing novel sensing approaches that incorporate these principles to safeguard privacy at the systems level.

3 Methodology

This section details our comprehensive methodology for selecting relevant works in the usable privacy-aware sensing domain. We first present a brief overview of the challenges we encountered using traditional keyword-based search approaches for cross-disciplinary work that spans a wide number of venues, which led us to explore an embedding vector-based approach to identify highly relevant works. We then provide a brief background on the difference between Large Language Models (LLMs), which are not used in this SoK, and embedding models, which we utilize to sort collected works. Finally, we provide technical details of our iterative sorting system, criteria for inclusion in the codebook via thematic questions, and our approach to coding the identified works.

3.1 Challenges with Traditional Search Engine Approaches for Identifying Relevant Works

We first utilized conventional digital research libraries (e.g., Google Scholar, DBLP, ACM DL, IEEE Xplore) and their respective keyword-based search to assist in identifying relevant works, similar to previous SoKs [104, 109, 116, 126]. However, we encountered challenges and limitations when using this approach to find relevant works.

First, individual sets of search keywords provided by each author of this work often produced results from their own home community. For example, using search keywords from an author who more frequently publishes within HCI venues generally returned results from HCI venues. Additionally, using the suggested keywords from an exemplar work previously known by the authors resulted in similar behavior, where the results often were from the same academic venue grouping of that work. It was apparent from this behavior that keyword-based search returned only a partial view of the complete academic space, highlighting the strong presence of community-specific jargon and the potential of a siloing effect within their respective venues.

Second, we also found search engines often omitted highly relevant works, possibly because the works did not explicitly include one or more of the keywords yet included semantically relevant terms. For example, the word "sensing" often does not appear in works about privacy related to smart IoT devices that perform sensing tasks in homes, yet it is a highly related term. Additionally, small variations in jargon across venues may lead to additional omissions; consumer devices, consumer off-the-shelf (CotS) devices, and IoT devices are terms that largely map to similar domains but yield very different search results.

Third, overly broad search terms yield thousands of loosely related results that would make finding relevant papers overly tedious for domain experts to individually review (e.g., searching using the terms "privacy sensing" in Google Scholar yields over 5M results). While the search space can be reduced by identifying specific venues that are more likely to yield relevant works, as our approach does, there is no effective "sorting" function to determine its semantic relevance compared to, for example, how often those keywords appear in the work. A work may be highly relevant, but if it does not have many references to that loosely related search term, it may be pushed far down the list of tens of thousands of search results. This is especially challenging when using multiple academic search engines and merging their results based on relevance. Thus, our proposed approach uses an Embedding-assisted search, organization, and visualization tool to be more inclusive of relevant works outside the authors' home communities and mitigate the limitations mentioned above, which may result from potential siloing effects from using community-specific keywords.

3.2 Background on Large Language Models versus Embedding Models

The emergence of Large Language Models (LLMs), most notably ChatGPT, has already significantly affected scientific research and writing. On the positive side, these AI tools can assist in many research tasks, such as summarizing articles and correcting grammatical issues in writing. On the negative side, these tools can sometimes fabricate sources (i.e., hallucinations) and revise passages to include plagiarised text convincingly, making it easy for researchers to rely on the LLM improperly. We explicitly state that LLMs were not used to perform qualitative tasks as part of this SoK; we never used an LLM to perform tasks requiring human judgment or synthesis, including labeling the work, determining inclusion in the codebook, and performing coding and thematic analysis.

We note that while ChatGPT can summarize texts and respond to basic queries (e.g., "Does this paper relate to privacy?"), it cannot replace human authors when performing judgment tasks. First and foremost, ChatGPT, in its current implementation (GPT-4.0), is non-deterministic and produces improper results for even simple judgment tasks (e.g., "Does this work relate to privacy *and* sensing?"); they can be incorrect (e.g., responds with "no" instead of the correct response, "yes"), inconsistent (e.g., responds "no" for a given work, but "yes" when asked a second time for the same or similar work), or indeterminate (e.g., responds with an irrelevant response). Furthermore, ChatGPT cannot provide consistent reasoning when pressed to explain its decision when given a judgment task. Thus, we reiterate that we did not use ChatGPT as part of our SoK and recommend that future authors avoid using ChatGPT or generative models in their current instantiation for judgment tasks.

Instead, we developed a tool based on the vector embedding representation of papers to perform a well-defined set of tasks to assist authors, described below in a further subsection, and specifically designed it such that it cannot produce an output that could be used to substitute human judgment. We use OpenAI's Ada model *not* ChatGPT or any generative language model. The main distinction is that the Ada model is a deterministic model that represents each

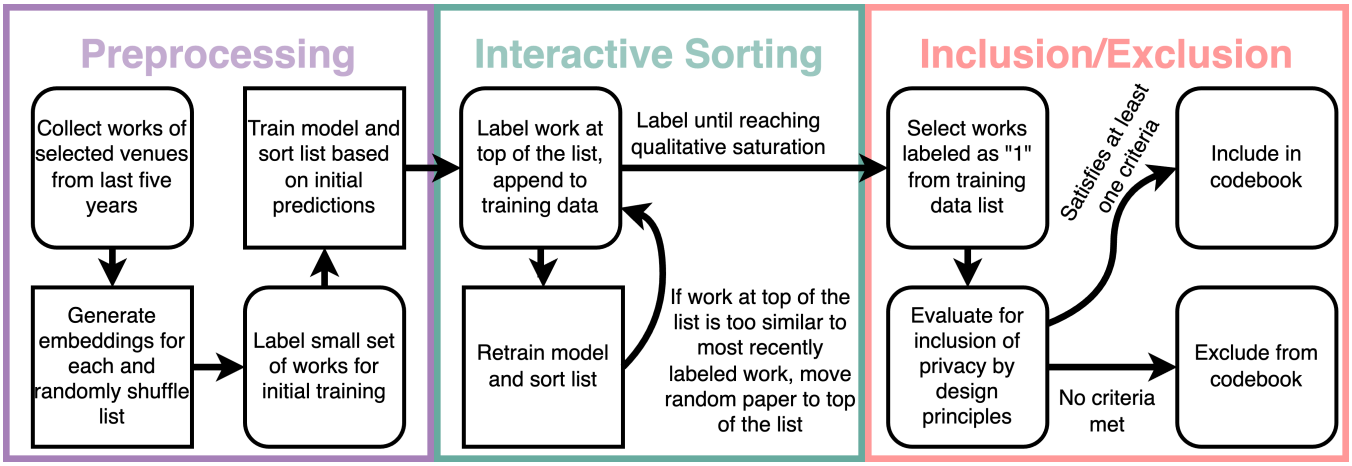

Figure 2: A flowchart highlighting the steps of our process to identify relevant works for our codebook. Steps that are performed manually by the authors are denoted by rounded squares. Note that the system does not assist the review process beyond sorting the list of works. The authors manually evaluate for inclusion/exclusion and code per the criteria described in Section 3.

work independently as a vector embedding, which we use to improve and accelerate the sorting of papers for the human authors to review, label, code, and analyze individually. In this regard, the Ada model presents a more effective search engine, similar to other prior SoK methodologies, but rather than optimizing results based on various aggregate relatedness metrics (e.g., keyword frequency, cited by, citing), the results are interactively optimized to what we, as authors, are seeking. Again, we reiterate that the embedding model was only used to improve the sorting ability of candidate works, but each work was individually evaluated by the authors. Furthermore, the authors could query candidate works manually for search keywords and other traditional sensemaking approaches to identify exemplar works. Thus, we are confident that our methodology is equally thorough and aligned with existing SoK works.

3.3 Embedding-assisted Sorting of Works

3.3.1 ***Collecting Works from Relevant Venues.*** We began by preprocessing our dataset, as depicted in the left panel of Figure 2. The authors identified 12 venues spanning four distinct shareholder communities: *Human-Computer Interaction* (CHI, IMWUT), *Mobile Systems* (MobiSys, MobiCom, SenSys), *Security & Privacy* (USENIX Security, IEEE S&P, CCS, NDSS), and *Usability in Privacy* (SOUPS, PETS, FAccT). We group these venues into communities based on factors such as their SIG relationships (e.g., SIGCHI, SIGMOBILE), overlap in their Program Committee, and topics listed in their Calls for Papers. These groupings were agreed upon by author consensus. With the aid of the DBLP bibliography library, the title and abstracts from works at these venues from 2018 to 2023 were compiled by a combination of programmatic or manual collection (i.e., copy-paste into a spreadsheet), resulting in 10,122 entries. Venues without their 2023 proceedings available prior to July 2023 were not included.

3.3.2 ***Representing Collected Works as Embedding Vectors.*** An advanced vector embedding model, such as the OpenAI Ada 002 model [87], allows us to build a system to search and organize relevant works that offer advantages over traditional keyword search. The first advantage is that rather than identifying relevant works based on whether the text contains individual search keywords or synonyms of the search keywords, we can represent text as its embedding vector and identify relevant works based on how similar they are to the embedding vectors of previously identified relevant works [88]. Functionally speaking, this means that rather than relying on the authors' ability to identify the correct set of keywords (which has the limitations described above and may differ across communities and involve community-specific jargon), the system more directly relies on the authors' ability to identify a set of relevant works for which the system can find other related works. While other open-source embedding models can be used for this task, we selected the Ada model as it was the highest-performing model available at the time and did not require us to purchase dedicated hardware or deploy costly large GPU cloud instances. For reference, while OpenAI charges $0.10 USD for 1M tokens [88], Amazon EC2 charges $4.56/hr for the slowest instance [105] that can run typical 7B embedding models (27 GB VRAM) and have comparable performance to the Ada model (e.g., e5-mistral-7b-instruct) [43].

This Ada model was used for each work to generate vector embedding for text containing only the work's title and abstract. This was done for two reasons, which were found through initial small-scale experiments. First, we found that generating a vector representation on text from additional portions of the work, such as the introduction, did not significantly increase the cosine similarity, and thus alignment, of the vectors of pairs of work known by the authors to be related, compared to vectors based on just the title and the abstract. However, the additional text processing comes with an increase in computational cost. Second, other meta-data, such as author or affiliation, influenced the embedding vectors, such that works by the same author or affiliation had an increase in their cosine similarity to each other compared to when the author or affiliation information was omitted. Furthermore, when including the author or affiliation, works often had higher cosine similarity to works by the same author or affiliation versus to other works

selected by authors that were known to be related and on similar topics—we did not want the system to have its sorting behavior influenced by such affiliations. Thus, we generated a 1536-long embedding vector for each work based on its title and abstract.

3.3.3 ***Iterative Sorting and Labeling Based on Author Expertise.*** The second advantage is that rather than utilizing these embeddings as part of a single-shot classifier to determine works as either relevant or not relevant, we can utilize these embeddings as part of an iterative regressor, which can determine the degree of relevance [88], and thus, continually sort the list of works to be reviewed, effectively creating a recommendation engine based on how the authors categorized known relevant works. These embedding vectors can encode the semantic distribution and clustering of relevant works [88], offering a visual representation of relevant works and providing the ability to explore the "neighborhood" of related works, which allows us to understand the coverage of our identified works relative to the broader academic landscape.

Leveraging these characteristics, we built a tool that utilizes a Random Forest (RF) Regressor (SciKitLearn, default parameters) model to interactively sort and assist in identifying relevant works for inclusion in the codebook, as described by the middle panel of Figure 2. To start, we defined the following criteria to label works with a score from 1-5 as follows:

(1) Related to *both* privacy and sensing.
(2) Related to *either* privacy or sensing *and* a related field (e.g., sensing and encryption)
(3) Related to *either* privacy or sensing
(4) Related to *only* a related field (e.g., encryption)
(5) Completely unrelated

The collected works are randomly shuffled and compiled into a "working list" (WL), and the work at the top of the list is presented to the authors to label based on the criteria above. Once labeled, the work is removed from the WL and placed in the "training list" (TL). This process is repeated until at least five "1" works and five "5" works are identified. The TL then trains the RF model to predict the label of each work in the WL. The WL is then sorted based on their label prediction in ascending order (i.e., works with predicted labels closer to 1 appear first). Then, as the authors label works recommended from the WL, the RF model is retrained with each new label, regenerates predictions, and sorts the whole WL.

To avoid getting "stuck" in a "neighborhood" of related works, if the cosine similarity of the next upcoming work is within a threshold of the most recently labeled work, the system presents a work randomly from the WL. This provided the added benefit that the model periodically received negative training examples of completely unrelated works and random sampling to find undiscovered areas of relevant works. That randomly selected work is similarly labeled, moved from the WL to the TL, and the model is retrained, generates new predictions, and sorts the WL.

This process was repeated until the works reached saturation, where high-relevance works were no longer being identified. At this point, the model's label predictions were 3s or higher and were always labeled by the authors as 3s or higher. In total, the authors manually labeled 600 papers, of which 117 were labeled with a "1" and, thus, highly likely candidates for inclusion in the codebook. Work from all but one venue was represented in this list, as none of the work from FAccT produced likely candidates for inclusion.

3.3.4 ***Model Performance and Validation.*** To validate our approach, we performed an evaluation to determine reproducibility, model consistency, and how well the model performs at identifying high-relevance works. In this evaluation, we used the list of 600 works with assigned values labeled by the authors and performed a stratified 10-fold cross-validation, whereby the dataset is split into 10 balanced train/test sets, and 10 independent models are trained and evaluated: these 10 models would be trained on works that have ground truth value assignments and tested on works that also have ground truth value assignments. This process ensured each subset was representative of the entire dataset. We utilized the following metrics to encode how well the models work at their assigned task—additionally, since it is a 10-fold cross-validation, we can measure the consistency across the 10 models:

(1) Average Error: The average error in ground truth value versus predicted value indicates the models' accuracy.
(2) Standard Deviation of Errors: The standard deviation in errors across the 10 folds signifies how consistently the models perform regardless of their training and would show that the models are reproducible despite being trained with different training sets with different training orders.
(3) Percentage of Significant Errors: What percentage of works have a predicted value with an error greater than 2, and, as a corollary, how many papers that are high relevance value (i.e., 1s) had a predicted value higher than 3, which signifies medium to low relevance? This last metric is particularly important as if a ground truth value 1 work has a predicted value of higher than 3, it might not have been suggested by our tool. However, if a ground true value 1 work has a predicted value of 3 or lower, the authors would have still come across this work as works were evaluated well into the 3s until saturation.

Through our 10-fold cross-validation, the average error for the RF model was 0.600, which is less than the step size of 1 for the label values. In practice, this indicates, on average, that the model will predict the vast majority of papers correctly or have a score that is off by one (e.g., predicted a 2 for a work that is a 1). We found 59.5% of works had their ground truth and predicted scores exactly match, 81.5% of works had a difference in scores of 1 or less, and 99.0% predicted works had a difference of 2 or less, meaning that regardless of how the models are trained, they still accurately can predict the relevance of the paper and recommend it for author review. The standard deviation of errors was 0.101, indicating that the 10 models performed consistently; thus, if they were trained with different training sets and different training orders—as was performed in this evaluation—the results would be similar.

For significant errors, across all models, only 6 works were categorized with an error greater than 2, resulting in a 1% significant error rate. All 6 of these works had a ground truth score of 4, which was predicted as a score of 1. In practice, this error would not affect the outcome of the codebook as authors would review the work that was predicted as a 1, determine that it is not of relevance, and not include it in the codebook. More importantly, of the works that had a ground truth value of 1, none were predicted to have a value greater than 3. This indicates that, regardless of how the models

were trained, the model would still have recommended reviewing all of the works that were included in the codebook (i.e., none of the works in the codebook would have been missed).

We additionally re-ran this evaluation using Support Vector Machines (SVM)—with both a linear kernel and an RBF kernel—and found similar results for errors and standard deviations of errors—0.595 (SD = 0.130) and 0.555 (SD = 0.124), respectively. Similarly, there were no significant errors that would cause work to have been missed with either SVM model. This further indicates that our approach remains robust even when using different ML approaches.

3.4 Inclusion Criteria via Thematic Questions

To determine which works would be included in the codebook, depicted in the right panel of Figure 2, we tailored a set of questions based on Langheinrich's six principles of Privacy by Design (PbD) [68] (denoted in bold):

(1) **Notice:** *Does the work incorporate a way or recommendation for person(s) to have noticed that a sensing device is collecting information?*
(2) **Choice and Consent:** *Does the work incorporate a way or recommendation for person(s) to consent to control the data being collected?*
(3) **Anonymity and Pseudonymity:** *Does the work incorporate a way or recommendation for person(s) to have their information not linked to their identity?*
(4) **Proximity and Locality:** *Does the work incorporate a way or recommendation for person(s) to confine the sensing to a specific location or proximity?*
(5) **Adequate Security:** *Does the work incorporate a way or recommendation for the system to secure sensitive information or prevent it from getting leaked in some manner?*
(6) **Access and Resource:** *Does the work incorporate a way or recommendation for person(s) to be able to audit the system and/or does the system restrict data collection to only relevant data (and not more)?*

While there are various alternative privacy by design and taxonomy frameworks, such as ones by Cavoukian [22] and Solove [107], we found Langheinrich's principles, which are situated in the context of ubiquitous computing and IoT devices, enabled us to create this set of thematic questions that offered wide coverage relative to the breadth of contributions while remaining focused in the area we wish to explore, which is usable privacy-aware sensing.

A work's contribution must have satisfied at least one of the above questions for inclusion in the codebook in addition to being a work that engaged both sensing and privacy. This was to exclude works that were too generalized or theoretical, such as a novel encryption method that is not sensor-specific or was not created based on user-driven recommendations. Conversely, an interview study about a sensor would be included if it included discussions and recommendations incorporating a PbD principle. Based on these inclusion criteria, of the 117 highly likely "1" candidates, only three works were excluded as they did not contribute privacy mitigation that engaged sensors and user privacy beyond simply not incorporating any of the six PbD principles. These questions are also used to identify core characteristics, where each question represents a column in the codebook.

3.5 Additional Systematization Characteristics

In addition to the six characteristics derived from the PbD principles, which were defined ahead of time, we constructed seven additional characteristics that emerged through the analysis based on the contribution's privacy approaches and evaluation approaches to code the works by, each forming a column in the codebook.

3.5.1 ***Privacy Approaches.*** For privacy approaches, we constructed three systematization characteristics, which capture the academic space in which the authors present the work, their mitigation approach, and the author's mindset surrounding how to address their presented threat model.

Primary Identity: In which domain does the work identify as its primary contribution, usable privacy or privacy-aware sensing, even though the work can involve both? This was determined by the work's CCS Concept tags, author-defined keywords, or listed contributions. In situations where there were elements of both UP and PAS contributions (though infrequent), we used our collective judgement to determine which was the predominant contribution in the work. This characteristic was added to code what the authors of the contribution consider their work to be contributing towards, therefore categorizing their perspective rather than what we, as authors of the SoK, believe the work contributes towards. We note that this characteristic is used to group the works into UP and PAS works for later sections and quantitative analysis.

Privacy Factor: What type of privacy mitigation is proposed in the work? For this characteristic, we defined five categories:

(1) *authentication*, which utilizes user identity to define system behavior;
(2) *granularity*, which reduces the resolution or range of the system sensing or data collection;
(3) *purpose*, whereby the system uses purpose-built sensors to restrict its ability to collect certain types of privacy-invasive data;
(4) *sanitization*, whereby the system actively removes or prevents the collection of privacy-invasive data;
(5) *control*, whereby the system presents user-accessible actions to determine sensor behavior.

This characteristic was added to code the general approach the suggested privacy mitigation takes and to what effect the mitigation has on system behavior and data collection.

Privacy Mindset: How does the work contextualize its proposed privacy mitigation? For this characteristic, we defined four categories:

(1) *all or nothing*, which presents a binary perspective on privacy and proposes an absolute privacy mitigation;
(2) *better than nothing*, which presents a gradient perspective on privacy and proposes a privacy mitigation with known and well-scoped limitations;
(3) *security is privacy (defensive)*, which presents defensive security measures (e.g., system hardening, encryption) as proposed privacy mitigations;
(4) *security is privacy (offensive)*, which presents offensive security measures (e.g., jamming/attacking privacy invasive systems) as proposed privacy mitigations.

This characteristic was added to code how the work contextualizes its contribution and approaches mitigating a presented privacy threat model.

3.5.2 ***Evaluation Approaches.*** For evaluation approaches, we constructed four systematization characteristics, which capture how the proposed privacy mitigation is evaluated for a given measure of success.

Technical Evaluation: Did the work perform a technical evaluation of the proposed privacy mitigation demonstrating increased privacy (e.g., the proposed method effectively sanitizes the data)?

User Evaluation: Did the work evaluate the proposed privacy mitigation with users for factors such as acceptance, usability, or comfort?

Feasibility Evaluation: Did the work evaluate the feasibility of the proposed privacy mitigation, such as through a prototype, research artifact, or operational system?

System Evaluation: Did the work evaluate a system incorporating the proposed privacy mitigation for performance and privacy-preserving ability?

The first two, technical and user, are very common in PAS and UP works, respectively, but we wanted to code whether PAS and UP works additionally evaluated technically or against users. A feasibility evaluation characteristic was added to the codebook to determine whether the authors of the work considered how the proposed privacy mitigation would manifest in real-world systems or whether privacy mitigation is more theoretical in nature. Finally, a system evaluation characteristic was added to code whether a real artifact was evaluated and whether it was effective in a real and physical experimental situation rather than, for example, solely in a technical simulation or Wizard-of-Oz'ed for users.

We compiled a spreadsheet where each row represents a work, and each column represents the abovementioned characteristics, similar to what is performed in prior SoKs. For coding, we followed a consensus coding process similar to that described by Richards et al. [99]. Initially, one author performed the coding, and two other authors joined to discuss and refine the codes to agreement. During these discussions, we applied a similar analysis approach to a prior SoK [126], applying a reflexive thematic analysis approach [21] where new themes emerged (beyond the initial 6 PbD-based columns), leading to updates in the codebook and recoding of the works. These new themes formed seven characteristics defined by their Privacy Approaches and Evaluation Approaches. The authors discussed and reached a consensus on each work, ensuring full agreement on the codes and themes presented. Regarding the experience of the coders, one is a Ph.D. Candidate with 10 years of sensing research experience in HCI and Mobile Systems, one is an Assistant Professor with research and Papers Committee experience in HCI, UiP, and S&P, and one is an Associate Professor with research and Papers Committee experience in HCI and Mobile Systems and serves as an Editor of an HCI journal.

4 Systematization

Through analysis of our codebook, four distinct themes emerged that highlight how the broader community has explored research questions surrounding usable privacy-aware sensing:

(1) *Control and awareness form the cornerstone of usable privacy*
(2) *"All or Nothing" vs. "Better than Nothing": balancing system functionality, user acceptance, and privacy*
(3) *Unvalidated assumptions of user needs in privacy-aware sensing contributions*
(4) *User-driven privacy mitigations often lack feasibility or system evaluations*

The following subsections will analyze each individual theme, how it relates to the broader body of usable privacy-aware sensing work, and key takeaways. Based on these themes, we discuss community recommendations in a later section (Section 6.1).

4.1 Control and Awareness Form the Cornerstone of Usable Privacy

We observed that works contributing to UP strongly emphasized user *control* and *awareness* as their Privacy Factor and Privacy Mindset in developing privacy mitigation mechanisms for sensing devices. Within this broader theme, we observed two subthemes: UP contributions very infrequently employ alternative Privacy Factors and Privacy Mindsets, and, when PAS contributions do employ *control* as a Privacy Factor, albeit rarely, they are paired with a different Privacy Mindset than *awareness*.

UP contributions very infrequently (only 13 [2, 19, 37, 52, 54, 66, 69, 78, 81–83, 119, 125] of 70) came from a different Privacy Factor and Privacy Mindset other than *control* and *awareness*, which represented the remaining 57 out of 70 UP works. However, these works still often recommended the design principles of *Notice* and *Choice & Consent* as part of their privacy mitigation mechanisms. This suggests that even if a UP contribution emphasizes a differing Privacy Factor or Privacy Mindset than *control* and *awareness*, they still play a significant role in that work; for UP contributions, regardless of Privacy Factor or Mindset, 78.5% incorporated *Notice* and 88.5% incorporated*Choice & Consent*.

PAS contributions, conversely, only infrequently (4 [33, 34, 84, 117] of 44) had the *control* privacy factor, with only one pairing it with *awareness* [84]. In works that employed an offensive mindset, control meant granting a user the ability to enable/disable a system or enable/disable an active defense (e.g., controlling microphone behavior [28, 47, 57, 67, 71, 110]), and awareness is implied but not a goal (i.e., the system doesn't seek to notify or remind the user of its state). Additionally, very few works incorporated the design principles of *Notice* (5 [26, 33, 34, 60, 84] of 44) and *Choice & Consent* (14 [23, 26, 33, 34, 47, 55, 57, 60, 71, 74, 84, 114, 128, 131] of 44).

From this theme, we identified a tension between user autonomy and user burden. Works categorized with *control* and *awareness*, (most often UP works) argue that the more the user can control elements of sensing devices that relate to privacy, the more they can ensure their own privacy. However, PAS works often prescribe privacy mitigations that do not engage the user and operate in the background, emphasizing automating privacy protections and reducing user burden. This gap presents an exciting avenue for each community to explore further. UP works could explore less user-dependent methods to improve privacy and, within that space, evaluate what behind-the-scene sensor privacy mitigations would be acceptable to users. PAS works could explore how the design principles of *Notice* and *Choice & Consent* could enhance existing

privacy mitigations by leveraging the user's ability to identify privacy concerns, work with the user to mitigate them, and improve user trust through active participation.

4.2 Balancing System Functionality, User Acceptance, and Privacy

Within this theme, we observed a divide in the approaches between two Privacy Mindsets, "All or Nothing" and "Better than Nothing." In the "All or Nothing" approach, contributions prescribed definitive, explicit safeguards for user privacy that involve completely disabling a sensing device (or its sensing capability) or leaving it to operate without modification [26, 33, 34, 72, 101, 103, 125, 131]. This ensures comprehensive privacy mitigation that can easily match a user's mental model, as the actual operation behaviors of a completely disabled device (i.e., no data collection) are well defined and should match the user's expectation (i.e., no data collection). Examples include UP contributions that tangible privacy-preserving methods for smart homes [125] and PAS contributions that fully disable sensors from collecting data by cutting off power [26].

In "Better than Nothing" approaches [4, 8, 9, 32, 37, 52, 60, 61, 69, 83, 98, 100, 119, 121, 132, 137], contributions identified that "All or Nothing" approaches come at the cost of long-term and continuous system performance (i.e., how effective can the system be in performing its duties if it is frequently disabled and re-enabled) in order to provide explicit guarantees of user comfort and privacy. Conversely, these works propose methods that balance these factors by utilizing sanitization (e.g., removing speech from acoustic information [60, 132]) or granularity (e.g., reducing the resolution of images to reduce sensitive content [100, 121]).

With respect to how these approaches relate to the broader body of work, we observe a difference in how these works present and evaluate their contribution. Since the "Better than Nothing" approaches aim to find a solution that balances system functionality with user acceptance and privacy, the contributions of the work often engage in the discussion and evaluation of how the contribution considers its tradeoffs, limitations, and how users perceive the system. "All or Nothing" approaches often placed a technical measure of privacy above all. While many works evaluated their contributions with users to determine their acceptance, they are more often evaluated for whether the user accepts the mitigation as a solution to the threat model they present rather than whether the user accepts the mitigation as usable [34, 125]. Furthermore, since the system's needs are secondary to the privacy need, whether the system remains usable or functional (or other downstream effects of the mitigation) is often not evaluated or discussed. For example, while offensive ultrasonic jamming [28, 57, 110] can manifest additionally as an "All or Nothing" protection against microphones for the wearer, these systems indiscriminately jam all microphones in the environment within a certain radius. A more benign side-effect is that a wearer of such a device effectively robs those in proximity of the convenience of using voice commands with virtual assistants. Worse, however, is that the wearer could unintentionally and significantly impede a nearby person's access to urgent necessities, such as a phone call to emergency services.

From this theme, there are competing factors of system functionality, user acceptance, and privacy. However, for *usable privacy-aware sensing*, all three factors are of immense importance. While "All or nothing" approaches can make strong privacy guarantees, those guarantees often come at the cost of usability or system functionality. Notable exceptions, however, are approaches that pair the "Purpose" Privacy Factor with the "All or Nothing" Privacy Mindset, where the system is designed for a specific purpose and has sensors only capable of collecting specific information (e.g., air quality) and is explicitly prevented from performing any other tasks or collecting other kinds of information. In this case, the privacy mitigation does not affect the usability or functionality of the device. Overall, while both approaches can improve sensor privacy, we recommend that all approaches consider their proposed privacy mitigations' downstream effects on usability and system functionality.

4.3 Unvalidated Assumptions of User Needs in Privacy-Aware Sensing Contributions

For PAS contributions, a significant number of identified works (35 of 44) did not validate their privacy mitigation against user needs and comfort but rather based their approach on assumptions or an interpretation of prior work. For example, many works often perform a technical analysis of their privacy mitigation and represent its effectiveness using some metric that maps to increased privacy (e.g., how effective speech was anonymized [9, 131, 132], was sensitive content removed from an image [100, 137]) but do not evaluate it with users or as part of a system to determine whether this mitigation is useful. The privacy mitigation may not meet user comfort needs, might be difficult to control/use, or may not be perceived as valuable, potentially reducing its acceptance by users. The privacy mitigation may also have unforeseen impacts when included as part of a system as a whole. For example, it may reduce the performance of the system (e.g., accuracy, responsiveness, battery life), which may also be an unacceptable tradeoff for users.

The main takeaway from this theme is that the consideration should not always be whether the privacy mitigation is performant or "correct," but rather if the mitigation is usable and would be accepted and, therefore, adopted, which were evaluated by only 9 PAS works [8, 23, 26, 28, 33, 34, 60, 74, 114]. Since evaluations with users are less prevalent within the PAS works, the onus falls on a subsequent researcher or developer to determine whether that approach is accepted by users within a given context. This barrier may reduce the willingness of others to utilize this research as part of a larger system, reducing its potential for impact. Given these factors, there is an immense opportunity for PAS contributions to close the loop and evaluate the privacy mitigation with users. This can help refine the contribution and offer an additional metric to highlight its significance, leading to greater impact.

4.4 User-driven Privacy Mitigations Often Lack Feasibility or System Evaluations

Related to the theme above, UP contributions often (which we quantify in the following section) do not perform a technical, feasibility, or system evaluation to determine whether the proposed privacy mitigation can be implemented as part of a system and, if so, what its quantified effects on user privacy are. For example,

while many useable privacy contributions identify valuable user privacy needs and use qualitative methods to develop foundational privacy theory and design recommendations, the works often do not provide concrete examples of how these mitigations would be implemented or a discussion of how existing systems could be adapted to support user needs. This makes it difficult for designers and future researchers to develop systems that build upon these works' findings. Many works that present concrete design recommendations but do not perform a feasibility evaluation present solutions that may not be feasible given current computing or hardware resource constraints (e.g., replacing all persons in images with a high-definition cartoon representation using only on-device hardware). In our coding of the feasibility characteristic, we took a very generous approach to what satisfied a feasibility evaluation: a discussion on how the usable privacy mitigation could be integrated into real-world situations would satisfy this characteristic. This had an effect of generally increasing the number of works that satisfied the feasibility characteristic (37 of 70), but of those works, 23 [2, 10–12, 20, 24, 39–41, 51, 64, 69, 75, 76, 78, 89, 102, 111, 122, 129, 133–135] included a discussion, 9 [38, 42, 54, 83, 94, 108, 115, 125, 130] presented an artifact (such as a paper prototype or visual example), and 5 [14, 19, 70, 73, 82] built a prototype (of which only 3 [19, 70, 82] included a system evaluation). If we consider feasibility in the narrowest sense (i.e., demonstration through prototype), only 7.1% of UP works satisfy this characteristic.

The key takeaway from this theme is that for UP contributions, feasibility evaluations, and discussions are critical to making their findings accessible and actionable for adoption by future system researchers and industry partners. The previous theme highlights that without user validation, the system may effectively protect user privacy, but users may not be comfortable with or willing to adopt the approach. This theme presents the inverse, where, without a feasibility evaluation, the proposed mitigation may be effective in improving user comfort and adoption, but system designers may not understand how to incorporate the design recommendations or implement them with the technology available.

5 Quantitative Findings

We also used quantitative approaches to investigate our qualitative findings and provide additional context to the methods employed by various stakeholder communities.

5.1 UP Validate through Users; PAS Validate through Technical Metrics

We utilized the codebook to also quantify how the two types of works evaluated their contributions as defined in Section 3. For reference, a technical evaluation involves evaluating the proposed method using an objective measure of increased privacy (e.g., 100% of speech content was removed from acoustic data collection); a user evaluation involves evaluating the proposed method with users for factors such as acceptance, comfort, or usability; a feasibility evaluation involves an evaluation of whether the proposed method can be integrated into a target situation/system (e.g., with an artifact, a prototype, a discussion of existing products or approaches); and a system evaluation involves evaluating the proposed method as part of a system and determining its effect on privacy preservation and/or system performance. The summary statistics for each type of work can be seen in Table 1.

Contribution	Technical	User	Feasibility	System
Usable Privacy	4.3%	92.9%	52.9%	4.3%
Privacy-Aware Sensing	93.2%	20.5%	90.9%	63.6%

Table 1: How UP and PAS works evaluate their contributions, broken down by percentage of works that performed technical, user, feasibility, and system evaluations.

Usable Privacy (70)	Notice	Choice & Consent	Anonymity & Pseudonymity	Proximity & Locality	Adequate Security	Access & Resource
Notice (55)	1	**0.909**	0.218	0.236	0.291	0.345
Choice & Consent (62)	**0.806**	1	0.210	0.242	0.403	0.419
Anonymity & Pseudonymity (17)	0.706	**0.765**	1	0.235	0.471	0.412
Proximity & Locality (15)	0.867	**1**	0.267	1	0.333	0.333
Adequate Security (27)	0.593	**0.926**	0.296	0.185	1	0.444
Access & Resource (28)	0.679	**0.929**	0.250	0.179	0.429	1

Table 2: The probability of any of the principles to be paired with another principle. The number of works that utilize that principle is denoted with parentheses. For Usable Privacy, Choice & Consent pairs strongly with all other approaches.

Privacy-Aware Sensing (44)	Notice	Choice & Consent	Anonymity & Pseudonymity	Proximity & Locality	Adequate Security	Access & Resource
Notice (5)	1	**1**	0.200	0	0.600	0.400
Choice & Consent (14)	0.357	1	0.357	0.357	**0.714**	0.357
Anonymity & Pseudonymity (27)	0.037	0.185	1	0.592	**0.481**	0.333
Proximity & Locality (19)	0	0.263	**0.842**	1	0.368	0.474
Adequate Security (26)	0.115	0.385	**0.500**	0.269	1	0.308
Access & Resource (15)	0.133	0.333	**0.600**	**0.600**	0.533	1

Table 3: The probability of any of the principles to be paired with another principle. The number of works that utilize that principle is denoted with parentheses. For PAS, approaches often included a pairing with Anonymity and/or Security. However, very few works included Notice and was an unlikely pairing with the other methods.

Overall, the quantitative results match the themes presented in the previous section. Usable Privacy (UP) contributions robustly evaluate their contributions with users. However, only a little over half of the contributions performed a feasibility evaluation of whether the proposed privacy mitigation can be incorporated or developed as part of existing systems or envisioned situations—or only 7.1% under a narrower definition of feasibility as described in the previous section. Furthermore, very few of these contributions perform technical or systems evaluations that provide a quantitative measure of how their privacy mitigation improves user privacy.

For Privacy-Aware Sensing (PAS) contributions, they conversely robustly evaluate the technical aspects of their contributions, but less than a quarter evaluate the contribution with users for usability or acceptance. While feasibility is often demonstrated through a prototype as part of the technical evaluation, a percentage of contributions did not evaluate the proposed mitigation as part of a system, denoting the lower System evaluation percentage.

5.2 UP and PAS Contributions Do Not Have Overlapping Privacy Priorities

We also analyzed our codebook to see if there are differences in how the different communities utilize privacy by design principles in their contributions. For each community, UP and PAS, we computed a pairing coefficient between each of the six privacy principles with another principle as a percentage of works that include that pairing. For example, for UP, of the 55 works that incorporate *Notice*, 90.9% of them also incorporated *Choice & Consent*. The results of these pairing coefficients can be seen in Tables 2 and 3.

For contributions in the UP domain, the majority of works employed *Notice* (55) and *Choice & Consent* (62). These principles also paired strongly with each other, where more than 80% of works that employed one employed the other. *Choice & Consent* also strongly paired with all other principles, being the top pairing for each. However, the remaining four principles had relatively weak pairings with each other and were infrequently utilized; for example, *Proximity & Locality* was used in less than a quarter of UP contributions.

Within the PAS contributions, *Anonymity & Pseudonymity* and *Adequate Security* were the most frequently used principles and had strong pairings with all other principles except for *Notice* and *Choice & Consent*, the two preferred principles for UP contributions. These two principles were the most infrequently used for PAS contributions, with only 5 works employing *Notice*. For the works that employed *Choice & Consent*, the more likely pairing was *Adequate Security*, rather than *Notice* as is the case for UP contributions.

Overall, we observe a distinct divide in the manner in which the two communities not only utilize privacy principles, but also how they pair and combine contributions. Table 4 presents a condensed version of our codebook with 37 unique combinations of privacy principles. While certain combinations are presented by both communities (e.g., Row 17 is represented in both communities), there remains a skew in the combinations. Of the 13 works represented by Row 13, 12 are UP contributions. Conversely, of the 7 presented in Row 22, 6 are PAS contributions. Overall, from the rest of the table, we can see that while there is a diverse distribution of combinations across the literature, each community is distinctly focused on a small subset of those combinations.

	Notice	Choice & Consent	Anonymity & Pseudonymity	Proximity & Locality	Adequate Security	Access & Resource	# Usable Privacy	# Privacy-Aware Sensing	# Total
1	●	●	●	○	●	●	0	1	1
2	○	○	●	●	○	○	0	5*	5
3	●	●	●	●	○	○	1	0	1
4	○	○	●	●	●	●	0	3*	3
5	○	●	○	●	●	●	1	1	2
6	○	○	●	●	○	●	0	4*	4
7	○	○	●	○	○	○	1	3*	4
8	●	●	○	○	●	○	4*	2	6
9	●	○	○	○	○	○	4*	0	4
10	○	○	○	●	○	●	0	1	1
11	○	●	●	○	●	●	0	1	1
12	●	●	●	●	○	●	1	0	1
13	●	●	○	○	○	○	12*	1	13
14	○	○	○	○	●	○	0	5*	5
15	○	●	●	●	○	○	0	2*	2
16	●	●	○	○	●	●	4*	0	4
17	○	●	○	○	●	○	3	2	5
18	●	●	●	●	●	●	1	0	1
19	●	●	○	●	○	○	7*	0	7
20	○	●	●	●	○	●	1	0	1
21	○	●	●	○	●	○	1	0	1
22	○	○	●	○	●	○	1	6*	7
23	●	●	●	○	○	●	2*	0	2
24	○	○	●	●	●	○	0	1	1
25	●	●	●	○	○	○	2*	0	2
26	●	●	○	○	○	●	9	1	10
27	○	○	○	○	○	●	0	1	1
28	●	●	●	○	●	○	4*	0	4
29	○	●	○	○	●	●	4*	1	5
30	○	●	○	○	○	●	2*	0	2
31	○	●	○	●	●	○	0	1	1
32	●	●	○	●	●	○	2*	0	2
33	●	●	○	●	●	●	1	0	1
34	●	○	●	○	○	●	1	0	1
35	○	●	●	●	●	○	0	1	1
36	○	○	○	○	●	●	0	1	1
37	○	○	●	○	●	●	1	0	1

Table 4: The 37 unique combinations found in our codebook. A star denotes a combination of more than two works in which a lopsided composition exists across communities.

	Human-Computer Interaction	Mobile Systems	Security & Privacy	Usability in Privacy
Human-Computer Interaction	1	0.516	**0.299**	0.565
Mobile Systems	0.516	1	0.480	**0.256**
Security & Privacy	**0.299**	0.480	1	0.673
Usability in Privacy	0.565	**0.256**	0.673	1

Table 5: The BC values for each community pairing. A value of 1 indicates identical distributions; 0 indicates no overlap. The pairing with the least overlap is in bold.

5.3 Direct Embeddings Analysis Shows Distinct Divide in How Contributions Talk About Usable Privacy-Aware Sensing

Beyond the works included in our codebook, we revisited the generated embeddings for all 10,122 works collected across the 12 venues spanning four distinct shareholder communities: Human-Computer Interaction (CHI, IMWUT), Mobile Systems (MobiSys, MobiCom, SenSys), Security & Privacy (USENIX Security, IEEE S&P, CCS, NDSS), and Usability in Privacy (SOUPS, PETS, FAccT). We generated a t-SNE projection to reduce the dimensionality of the embedding and assist in visualization (see Figure 1). Then, using the reduced dataset, we fit a 2-D Gaussian ellipse to each community for visualization purposes and compute the Bhattacharyya coefficient (BC), which represents the similarity between two distributions for each pairwise combination, as seen in Table 5. In this case, the BC distance provides an effective general-purpose metric for overlap in two normally distributed distributions and provides an intuitive measure of overlap. A BC of 1 indicates identical distributions, while a BC of 0 indicates completely non-overlapping distributions. For a point of comparison, Cohen's d, a commonly used metric for comparing distributions, does not take into account the variance and shapes of the distribution but rather the magnitude of mean differences [90]. Given that the shapes of the distributions are also of interest, we chose the BC distance metric.

While these results are based on our entire embeddings dataset, not just the works selected for the codebook, they reveal characteristics that reflect themes identified from the qualitative analysis of our codebook. The BC values for Usability in Privacy venues suggest a significant overlap with Security (0.673) and HCI (.565) venues. Similarly, the BC values for Mobile Systems venues suggest a significant overlap with Security (.480) and HCI (.516). This is unsurprising given the historical relationship and development of these communities and how they situate themselves (e.g., usability in privacy engages both HCI and security topics, and mobile systems engages both HCI and security topics).

However, the BC values also suggest minimal overlap between Mobile Systems and Usability in Privacy (0.256) as well as minimal overlap between HCI and Security (0.299). As seen in Figure 1, these are also the two combinations where their fitted ellipses do not have any overlap. This indicates that few works engage these two combinations of communities, despite these communities having significant overlap with a shared community: Mobile Systems and Usability in Privacy both have significant overlap with HCI and Security individually, but not with each other, while HCI and Security both have significant overlap with Mobile Systems and Usability in Privacy, but not with each other.

Given the strong preference in the manner that UP and PAS works evaluate their contributions and opt for different privacy design principles, shown in the previous subsections, these patterns of significant or minimized overlap across communities offer additional evidence that the communities have distinct "cultures"

that influence how privacy mitigations for sensors are ideated, designed, and evaluated. Furthermore, the minimal overlap between certain combinations of communities further suggests a lack of cross-pollination, highlighted by the low density of works at the center of the t-SNE visualization, which may partially influence the minimal overlap in the methodologies described above. We see this as a significant growth opportunity for cross-disciplinary Usable Privacy-Aware Sensing research.

6 Discussion

In this section, we discuss a series of recommendations based on our findings that can help address the gap between Usable Privacy (UP) and Privacy-Aware Sensing (PAS) research approaches and foster greater collaboration. We then reflect on our SoK approach, its limitations, and future work. Finally, we philosophize on opportunities for cross-disciplinary academics to influence IoT design and policy through usable privacy-aware sensing (UPAS) research.

6.1 Reflections and Recommendations for the Research Community

We propose two recommendations based on our findings in this work. The first recommends increasing the diversity and breadth of privacy principles (i.e., using some of the other less common Privacy by Design principles) within both UP and PAS contributions through increased engagement across disciplines. The second recommends both contribution types evaluate their work end-to-end, such as by. For including both a user component *and* a feasibility evaluation to allow researchers from any community to find value in the contribution that overlaps with their priorities.

6.1.1 Increasing the Diversity of Privacy Principles in both Usable Privacy and Privacy-Aware Sensing. Given the strong affinity of UP and PAS contributions towards specific privacy principles, we recommend that UP and PAS researchers increase the variety of approaches and explore each others' preferred approaches. While there are many opportunities to increase the overlap between the two contribution types, we will primarily provide examples based on the lesser utilized approaches in each community: *Notice* for PAS contributions and *Proximity & Locality* and *Anonymity & Pseudonymity* for UP contributions.

PAS contributions could incorporate the *Notice* principle more frequently and find novel ways sensor hardware can support this opportunity. This could be as simple as adding an LED that lights up whenever the device collects sensor information. This relatively simple and easy addition could be incorporated in many situations. However, as a caveat, PAS designers must also consider whether that notice is adequate given diverse communities or accessibility needs (e.g., non-visual cues for blind and low-vision users).

Conversely, UP contributions can explore ways to make some of the lesser-used principles have greater usability or engage the user in a manner that does not increase their burden. *Proximity & Locality* and *Anonymity & Pseudonymity* were relatively underutilized privacy principles in UP works, which can enormously benefit from a greater diversity of usable interventions. For example, Usable Privacy contributions can explore how to inform users that the sensing radius of a device can be defined and, with that information, better understand the tradeoffs when utilizing such an intervention. Overall, we envision a boon of research contributions as UP and PAS overlap more significantly and contribute towards cross-disciplinary UPAS research. Ultimately, we believe a greater overlap in the approaches of both groups would proactively guard against academic siloing.

6.1.2 End-to-End Evaluations for Both Communities. We also recommend that both contribution types strive to evaluate their approaches from end-to-end when applicable, meaning that PAS contributions could be evaluated with users (or provide evidence of user acceptance based on the literature) and UP contributions could manifest their recommendations through prototypes (e.g., paper prototypes, Wizard of Oz prototypes) or, at the very least, provide discussions of the feasibility of their recommendations and how they can be incorporated into existing sensing systems.

For PAS contributions, user evaluations present an honest signal of whether or not users are comfortable with a given privacy mitigation and, therefore, improve the acceptance of the sensing system. For designers, when deciding between varying approaches, including a user evaluation is an important additional factor that helps contextualize where and when the mitigation is appropriate, as there is no one solution for all problems. Furthermore, to determine how inclusive a privacy solution is (i.e., across various demographics, different accessibility needs, and expectations of differing communities), user research must be included in all stages of the design process. Thus, we strongly recommend these evaluations to increase the ultimate impact of PAS contributions.

For UP contributions, evaluating their feasibility helps designers and PAS researchers gauge how well the recommendations could be incorporated into systems. Furthermore, these feasibility evaluations, such as through a paper prototype, help present how the authors see these recommendations in practice, as many theoretical recommendations can have various interpretations, as seen in many diverse works we identified that engage the same privacy principle.

We caveat the recommendations above with the understanding that different communities have varying methodologies for presenting and evaluating privacy mitigations; there will be contributions where it does not necessarily make sense to do technical, user, feasibility, or systems evaluations. However, to support emerging UPAS research, the following recommendations will make it easier for PAS and UP researchers individually to understand and appreciate each others' contributions.

6.2 Limitations

Within this work, we detail two sources of limitations of our SoK: the limitations of our embedding-based sorting tool and the limitations inherent to the qualitative methods used in SoKs.

6.2.1 Limitations of Our Embedding-based Sorting. Our embedding-based sorting approach has limitations, the first being that it partially inherits the limitations of other SoK works that narrow the focus of the search space by selecting a set of venues. While we utilized our author expertise to select 12 relevant venues, we do not expect it to be an exhaustive list of where relevant works could be found. Like other SoKs, the cost of collecting, organizing, and formatting works presents a similar human-labor bottleneck to collect works that can be sorted by our tool. While collecting titles

and abstracts and properly formatting them, sometimes manually, requires the most time and effort, generating and processing their embeddings also takes time and computational resources.

The second limitation is specific to the embedding model used to sort the works. As mentioned in an earlier section, the Ada model is not an open-source model and inherits all of the limitations that come from a closed-source model. In particular, OpenAI may replace the Ada model with a different offering or quietly fine-tune the model, making newly generated embeddings potentially incompatible with previously generated ones. Alternatively, robust open-source models would not face these issues, as the model can be checkpointed, shared, and reused indefinitely. However, the large computational requirements may make this approach inaccessible to many researchers. Based on the effectiveness of the sorting that allowed us to find a diverse set of works, we believe this embedding-based approach was well worth the effort, especially for SoK works that look at cross-disciplinary research. Future work will explore preparing our tool to be open-source, evaluate smaller open-source embedding models, and improve its usability to make it easy for researchers to help find connections across communities.

6.2.2 Limitations of Author Expertise-based SoK Work. We acknowledge that, irrespective of how effective our sorting tool is, our findings are still limited by author expertise. The tool can only help sort relevant works; it cannot aid in deciding whether a work should be included in a codebook or how to code a work. We still relied on our expertise to determine whether we reached qualitative exhaustion, as the tool could not make that determination and was explicitly designed not to do so. Thus, while the tool was incredibly effective in its narrow scope, it cannot replace human-effort SoK work. We also acknowledge that how we group communities (e.g., HCI, S&P, MS, UiP) is based on author expertise. However, regardless of how a work is grouped, its embeddings are not influenced by the grouping—the group label, nor the venue the work is from, is not included as part of the embedding. Thus, regardless of how we grouped the works, we required our qualitative expertise to contextualize the quantitative results that present a research gap represented within the embeddings.

Our work diverged from more common SoK approaches by shifting where, in the process, author expertise is applied. In more common SoK approaches, author expertise is applied at the keyword crafting stage, whereby authors precisely and iteratively craft search terms to find the right set of papers [104, 109, 126]. To bolster these results, authors rely on their expertise and prior knowledge of relevant papers to compile a well-rounded and representative codebook [104, 116]. However, relying on expert keywords also imposes an unrealistic expectation in cross-disciplinary exploration, where the authors must have expertise in various disciplines and their respective jargon and community-specific phrases.

To avoid this challenge, rather than knowing the proper jargon to use as search terms for those communities, our sorting tool allowed us to focus our author expertise on identifying UP and PAS contributions and whether they merited inclusion in our codebook. While our embeddings tool allowed us to find a diverse set of UP and PAS works dispersed across a wide range of venues—works that may have been buried under thousands of results if more generic terms such as "privacy" and "sensing" were used—we retain the same limitation as other SoKs where author expertise is still required to determine whether a work is deemed for inclusion or exclusion and, subsequently, how to code the work.

6.3 An Emerging Opportunity for Academia to Influence IoT Design and Policy

This work details, qualitatively and quantitatively, a research gap between UP and PAS contributions. This research gap causes contributions from either side to lack context and information that is valuable for affecting real-world IoT design and device privacy policies. As mentioned above, PAS contributions may not be adopted because it is unclear how usable their privacy mitigation is or whether or not they would be adopted by users. UP contributions may never see mainstream adoption by PAS researchers or designers because the recommendations may be too theoretical to put into practice and may have recommendations that cannot be implemented with today's technologies. Essentially, for an IoT device or policy to widely adopt a user privacy safeguard, it needs to be usable and real-world functional. Thus, this research gap hinders the emergence of more mainstream UPAS systems today.

Whether academia can influence IoT design and policy to improve user privacy requires us to consider whether the established stakeholder communities (e.g., S&P, HCI, MS, UiP) can, on their own, house and foster cross-disciplinary research that incorporates the best elements of each community. Based on our findings in this work, it does not appear that any single community can, on its own, support this type of work. Neither would any two combinations of communities; for example, HCI+MS contribute a significant number of PAS works, but very few UP works. However, despite the divide between UP and PAS contributions and the lack of overlap between certain community pairs, we see evidence of works at the center of all four stakeholder communities, albeit at a significantly lower density and occurrence. We see this emerging community as a way for academia to improve IoT user privacy.

7 Conclusion

To conclude, this work presented a Systematization of Knowledge (SoK) whereby 114 selected works were analyzed and qualitatively categorized into four distinct themes. Through these themes, we found strong community preferences that influence how researchers approach the challenges of designing Usable Privacy-Aware Sensing systems, revealing a research gap. We used these themes to highlight significant opportunities for cross-community collaboration to close this research gap. Various quantitative analyses provided additional evidence of this gap, and a significant divide between UP and PAS-style contributions is quantified. These results also showed a quantifiable lack of overlap in the venues where emerging UPAS research appears, and greater cross-pollination could increase support for this type of research. We concluded with recommendations, including how a greater diversity of design principles could be incorporated within UP and PAS works and formal recommendations for how stakeholder communities can incentivize new research in the cross-disciplinary UPAS domain.

References

[1] 2024. SenSys Sensors S&P 2023. https://sensorssp.github.io/sensorssp23/. Accessed: January 30, 2024.

[2] Noura Abdi, Kopo M. Ramokapane, and Jose M. Such. 2019. More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. USENIX Association, Santa Clara, CA, 451–466. https://www.usenix.org/conference/soups2019/presentation/abdi

[3] Desiree Abrokwa, Shruti Das, Omer Akgul, and Michelle L. Mazurek. 2021. Comparing Security and Privacy Attitudes Among U.S. Users of Different Smartphone and Smart-Speaker Platforms. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, 139–158. https://www.usenix.org/conference/soups2021/presentation/abrokwa

[4] Rebecca Adaimi, Howard Yong, and Edison Thomaz. 2021. Ok Google, What Am I Doing? Acoustic Activity Recognition Bounded by Conversational Assistant Interactions. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 2 (mar 2021), 24 pages. https://doi.org/10.1145/3448090

[5] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan C. Semaan. 2020. "I am uncomfortable sharing what I can't see": Privacy Concerns of the Visually Impaired with Camera Based Assistive Applications. In *29th USENIX Security Symposium, USENIX Security 2020, August 12-14, 2020*, Srdjan Capkun and Franziska Roesner (Eds.). USENIX Association, 1929–1948. https://www.usenix.org/conference/usenixsecurity20/presentation/akter

[6] Wael Albayaydh and Ivan Flechais. 2023. Examining Power Dynamics and User Privacy in Smart Technology Use Among Jordanian Households. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 4643–4659. https://www.usenix.org/conference/usenixsecurity23/presentation/albayaydh

[7] Wael S Albayaydh and Ivan Flechais. 2022. Exploring Bystanders' Privacy Concerns with Smart Homes in Jordan. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 446, 24 pages. https://doi.org/10.1145/3491102.3502097

[8] Rawan Alharbi, Mariam Tolba, Lucia C. Petito, Josiah Hester, and Nabil Alshurafa. 2019. To Mask or Not to Mask? Balancing Privacy with Visual Confirmation Utility in Activity-Oriented Wearable Cameras. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 72 (sep 2019), 29 pages. https://doi.org/10.1145/3351230

[9] Ranya Aloufi, Hamed Haddadi, and David Boyle. 2019. Privacy preserving speech analysis using emotion filtering at the edge: poster abstract. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems* (New York, New York) *(SenSys '19)*. Association for Computing Machinery, New York, NY, USA, 426–427. https://doi.org/10.1145/3356250.3361947

[10] Abdulmajeed Alqhatani and Heather Richter Lipford. 2019. "There is nothing that I need to keep secret": Sharing Practices and Concerns of Wearable Fitness Data. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. USENIX Association, Santa Clara, CA, 421—434. https://www.usenix.org/conference/soups2019/presentation/alqhatani

[11] Ahmed Alshehri, Eugin Pahk, Joseph Spielman, Jacob T Parker, Benjamin Gilbert, and Chuan Yue. 2023. Exploring the Negotiation Behaviors of Owners and Bystanders over Data Practices of Smart Home Devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 67, 27 pages. https://doi.org/10.1145/3544548.3581360

[12] Ahmed Alshehri, Joseph Spielman, Amiya Prasad, and Chuan Yue. 2022. Exploring the Privacy Concerns of Bystanders in Smart Homes from the Perspectives of Both Owners and Bystanders. *Proc. Priv. Enhancing Technol.* 2022, 3 (2022), 99–119. https://doi.org/10.56553/POPETS-2022-0064

[13] Noah J. Apthorpe, Danny Yuxing Huang, Dillon Reisman, Arvind Narayanan, and Nick Feamster. 2019. Keeping the Smart Home Private with Smart(er) IoT Traffic Shaping. *Proc. Priv. Enhancing Technol.* 2019, 3 (2019), 128–148. https://doi.org/10.2478/POPETS-2019-0040

[14] Natã Miccael Barbosa, Joon S. Park, Yaxing Yao, and Yang Wang. 2019. "What if?" Predicting Individual Users' Smart Home Privacy Preferences and Their Changes. *Proc. Priv. Enhancing Technol.* 2019, 4 (2019), 211–231. https://doi.org/10.2478/POPETS-2019-0066

[15] Natã M. Barbosa, Zhuohao Zhang, and Yang Wang. 2020. Do Privacy and Security Matter to Everyone? Quantifying and Clustering User-Centric Considerations About Smart Home Device Adoption. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. USENIX Association, 417–435. https://www.usenix.org/conference/soups2020/presentation/barbosa

[16] Ludovic Barman, Alexandre Dumur, Apostolos Pyrgelis, and Jean-Pierre Hubaux. 2021. Every Byte Matters: Traffic Analysis of Bluetooth Wearable Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2, Article 54 (jun 2021), 45 pages. https://doi.org/10.1145/3463512

[17] Benjamin Baron and Mirco Musolesi. 2020. Where You Go Matters: A Study on the Privacy Implications of Continuous Location Tracking. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 169 (dec 2020), 32 pages. https://doi.org/10.1145/3432699

[18] Julia Bernd, Ruba Abu-Salma, Junghyun Choy, and Alisa Frik. 2022. Balancing Power Dynamics in Smart Homes: Nannies' Perspectives on How Cameras Reflect and Affect Relationships. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. USENIX Association, Boston, MA, 687–706. https://www.usenix.org/conference/soups2022/presentation/bernd

[19] Yohan Beugin, Quinn Burke, Blaine Hoak, Ryan Sheatsley, Eric Pauley, Gang Tan, Syed Rafiul Hussain, and Patrick D. McDaniel. 2022. Building a Privacy-Preserving Smart Camera System. *Proc. Priv. Enhancing Technol.* 2022, 2 (2022), 25–46. https://doi.org/10.2478/POPETS-2022-0034

[20] Patrick Bombik, Tom Wenzel, Jens Grossklags, and Sameer Patil. 2022. A Multi-Region Investigation of the Perceptions and Use of Smart Home Devices. *Proc. Priv. Enhancing Technol.* 2022, 3 (2022), 6–32. https://doi.org/10.56553/POPETS-2022-0060

[21] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[22] Ann Cavoukian. 2012. Privacy by design: origins, meaning, and prospects for assuring privacy and trust in the information era. In *Privacy protection measures and technologies in business organizations: aspects and standards*. IGI Global, 170–208.

[23] Jason Ceci, Jonah Stegman, and Hassan Khan. 2023. No Privacy in the Electronics Repair Industry. In *2023 IEEE Symposium on Security and Privacy (SP)*. 3347–3364. https://doi.org/10.1109/SP46215.2023.10179413

[24] George Chalhoub, Ivan Flechais, Norbert Nthala, and Ruba Abu-Salma. 2020. Innovation Inaction or In Action? The Role of User Experience in the Security and Privacy Design of Smart Home Cameras. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. USENIX Association, 185–204. https://www.usenix.org/conference/soups2020/presentation/chalhoub

[25] George Chalhoub, Martin J Kraemer, Norbert Nthala, and Ivan Flechais. 2021. "It did not give me an option to decline": A Longitudinal Analysis of the User Experience of Security and Privacy in Smart Home Products. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 555, 16 pages. https://doi.org/10.1145/3411764.3445691

[26] Varun Chandrasekaran, Suman Banerjee, Bilge Mutlu, and Kassem Fawaz. 2021. PowerCut and Obfuscator: An Exploration of the Design Space for Privacy-Preserving Interventions for Smart Speakers. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, 535–552. https://www.usenix.org/conference/soups2021/presentation/chandrasekaran

[27] Chun-Yu (Daniel) Chen, Bo-Yao Lin, Junding Wang, and Kang G. Shin. 2019. Keep Others from Peeking at Your Mobile Device Screen!. In *The 25th Annual International Conference on Mobile Computing and Networking* (Los Cabos, Mexico) *(MobiCom '19)*. Association for Computing Machinery, New York, NY, USA, Article 22, 16 pages. https://doi.org/10.1145/3300061.3300119

[28] Yuxin Chen, Huiying Li, Shan-Yuan Teng, Steven Nagels, Zhijing Li, Pedro Lopes, Ben Y. Zhao, and Haitao Zheng. 2020. Wearable Microphone Jamming. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376304

[29] Mitchell Clark. 2022. "Google, like Amazon, may let police see your video without a warrant". Website. Retrieved July 27, 2022 from https://www.theverge.com/2022/7/26/23279562/arlo-apple-wyze-eufy-google-ring-security-camera-foortage-warrant.

[30] Camille Cobb, Sruti Bhagavatula, Kalil Anderson Garrett, Alison Hoffman, Varun Rao, and Lujo Bauer. 2021. "I would have to evaluate their objections": Privacy tensions between smart home device owners and incidental users. *Proc. Priv. Enhancing Technol.* 2021, 4 (2021), 54–75. https://doi.org/10.2478/POPETS-2021-0060

[31] Ivan De Oliveira Nunes, Seoyeon Hwang, Sashidhar Jakkamsetti, and Gene Tsudik. 2022. Privacy-from-Birth: Protecting Sensed Data from Malicious Sensors with VERSA. In *2022 IEEE Symposium on Security and Privacy (SP)*. 2413–2429. https://doi.org/10.1109/SP46214.2022.9833737

[32] Karel Dhondt, Victor Le Pochat, Alexios Voulimeneas, Wouter Joosen, and Stijn Volckaert. 2022. A Run a Day Won't Keep the Hacker Away: Inference Attacks on Endpoint Privacy Zones in Fitness Tracking Social Networks. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security* (Los Angeles, CA, USA) *(CCS '22)*. Association for Computing Machinery, New York, NY, USA, 801–814. https://doi.org/10.1145/3548606.3560616

[33] Youngwook Do, Nivedita Arora, Ali Mirzazadeh, Injoo Moon, Eryue Xu, Zhihan Zhang, Gregory D. Abowd, and Sauvik Das. 2023. Powering for Privacy: Improving User Trust in Smart Speaker Microphones with Intentional Powering and Perceptible Assurance. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 2473–2490. https://www.usenix.org/conference/usenixsecurity23/presentation/do

[34] Youngwook Do, Jung Wook Park, Yuxi Wu, Avinandan Basu, Dingtian Zhang, Gregory D. Abowd, and Sauvik Das. 2022. Smart Webcam Cover: Exploring the Design of an Intelligent Webcam Cover to Improve Usability and Trust. *Proc.*

ACM Interact. Mob. Wearable Ubiquitous Technol. 5, 4, Article 154 (dec 2022), 21 pages. https://doi.org/10.1145/3494983

[35] Yiwen Dong, Yuyan Wu, and Hae Young Noh. 2021. Social Distancing Compliance Monitoring for COVID-19 Recovery Through Footstep-Induced Floor Vibrations. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems* (Coimbra, Portugal) *(SenSys '21)*. Association for Computing Machinery, New York, NY, USA, 399–400. https://doi.org/10.1145/3485730.3492893

[36] Julia C. Dunbar, Emily Bascom, Ashley Boone, and Alexis Hiniker. 2021. Is Someone Listening? Audio-Related Privacy Perceptions and Design Recommendations from Guardians, Pragmatists, and Cynics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 98 (sep 2021), 23 pages. https://doi.org/10.1145/3478091

[37] Passant Elagroudy, Mohamed Khamis, Florian Mathis, Diana Irmscher, Ekta Sood, Andreas Bulling, and Albrecht Schmidt. 2023. Impact of Privacy Protection Methods of Lifelogs on Remembered Memories. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 508, 10 pages. https://doi.org/10.1145/3544548.3581565

[38] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the Experts: What Should Be on an IoT Privacy and Security Label?. In *2020 IEEE Symposium on Security and Privacy (SP)*. 447–464. https://doi.org/10.1109/SP40000.2020.00043

[39] Pardis Emami-Naeini, Joseph Breda, Wei Dai, Tadayoshi Kohno, Kim Laine, Shwetak Patel, and Franziska Roesner. 2023. Understanding People's Concerns and Attitudes Toward Smart Cities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 71, 24 pages. https://doi.org/10.1145/3544548.3581558

[40] Pardis Emami-Naeini, Janarth Dheenadhayalan, Yuvraj Agarwal, and Lorrie Faith Cranor. 2021. Which Privacy and Security Attributes Most Impact Consumers' Risk Perception and Willingness to Purchase IoT Devices?. In *2021 IEEE Symposium on Security and Privacy (SP)*. 519–536. https://doi.org/10.1109/SP40001.2021.00112

[41] Pardis Emami-Naeini, Janarth Dheenadhayalan, Yuvraj Agarwal, and Lorrie Faith Cranor. 2023. Are Consumers Willing to Pay for Security and Privacy of IoT Devices?. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 1505–1522. https://www.usenix.org/conference/usenixsecurity23/presentation/emami-naeini

[42] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. 2019. Exploring How Privacy and Security Factor into IoT Device Purchase Behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300764

[43] Hugging Face. 2024. MTEB Leaderboard. https://huggingface.co/spaces/mteb/leaderboard. Accessed: 2024-05-31.

[44] Yuanyuan Feng, Yaxing Yao, and Norman Sadeh. 2021. A Design Space for Privacy Choices: Towards Meaningful Privacy Control in the Internet of Things. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 64, 16 pages. https://doi.org/10.1145/3411764.3445148

[45] Sarah Frier. 2019. "Facebook Has Been Paying Contractors to Transcribe Users' Facebook Messenger Voice Chats". Website. Retrieved July 27, 2022 from https://time.com/5651395/facebook-contractors-transcribe-conversations-audio-files/.

[46] Andrea Gallardo, Chris Choy, Jaideep Juneja, Efe Bozkir, Camille Cobb, Lujo Bauer, and Lorrie Cranor. 2023. Speculative Privacy Concerns about AR Glasses Data Collection. *Proc. Priv. Enhancing Technol.* 2023, 4 (2023), 416–435. https://doi.org/10.56553/POPETS-2023-0117

[47] Chuhan Gao, Kassem Fawaz, Sanjib Sur, and Suman Banerjee. 2019. Privacy Protection for Audio Sensing Against Multi-Microphone Adversaries. *Proc. Priv. Enhancing Technol.* 2019, 2 (2019), 146–165. https://doi.org/10.2478/POPETS-2019-0024

[48] Christine Geeng and Franziska Roesner. 2019. Who's In Control? Interactions In Multi-User Smart Homes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300498

[49] Nina Gerber, Benjamin Reinheimer, and Melanie Volkamer. 2019. Investigating People's Privacy Risk Perception. *Proc. Priv. Enhancing Technol.* 2019, 3 (2019), 267–288. https://doi.org/10.2478/POPETS-2019-0047

[50] Eileen Guo. 2022. "A Roomba recorded a woman on the toilet. How did screenshots end up on Facebook?". Website. Retrieved January 2, 2023 from https://www.technologyreview.com/2022/12/19/1065306/roomba-irobot-robot-vacuums-artificial-intelligence-training-data-privacy/.

[51] Julie Haney, Yasemin Acar, and Susanne Furman. 2021. "It's the Company, the Government, You and I": User Perceptions of Responsibility for Smart Home Privacy and Security. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, 411–428. https://www.usenix.org/conference/usenixsecurity21/presentation/haney

[52] Rakibul Hasan, Yifang Li, Eman Hassan, Kelly Caine, David J. Crandall, Roberto Hoyle, and Apu Kapadia. 2019. Can Privacy Be Satisfying? On Improving Viewer Satisfaction for Privacy-Enhanced Photos Using Aesthetic Transforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300597

[53] Hussein Hazazi and Mohamed Shehab. 2023. Exploring the Usability, Security, and Privacy of Smart Locks from the Perspective of the End User. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. USENIX Association, Anaheim, CA, 559–577. https://www.usenix.org/conference/soups2023/presentation/hazazi

[54] Jinhan Hu, Andrei Iosifescu, and Robert LiKamWa. 2021. LensCap: split-process framework for fine-grained visual privacy control for augmented reality apps. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services* (Virtual Event, Wisconsin) *(MobiSys '21)*. Association for Computing Machinery, New York, NY, USA, 14–27. https://doi.org/10.1145/3458864.3467676

[55] Pengfei Hu, Hui Zhuang, Panneer Selvam Santhalingam, Riccardo Spolaor, Parth Pathak, Guoming Zhang, and Xiuzhen Cheng. 2022. AccEar: Accelerometer Acoustic Eavesdropping with Unconstrained Vocabulary. In *2022 IEEE Symposium on Security and Privacy (SP)*. 1757–1773. https://doi.org/10.1109/SP46214.2022.9833716

[56] Long Huang and Chen Wang. 2021. Notification privacy protection via unobtrusive gripping hand verification using media sounds. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking* (New Orleans, Louisiana) *(MobiCom '21)*. Association for Computing Machinery, New York, NY, USA, 491–504. https://doi.org/10.1145/3447993.3483277

[57] Peng Huang, Yao Wei, Peng Cheng, Zhongjie Ba, Li Lu, Feng Lin, Fan Zhang, and Kui Ren. 2023. InfoMasker: Preventing Eavesdropping Using Phoneme-Based Noise. In *30th Annual Network and Distributed System Security Symposium, NDSS 2023, San Diego, California, USA, February 27 - March 3, 2023*. The Internet Society. https://www.ndss-symposium.org/ndss-paper/infomasker-preventing-eavesdropping-using-phoneme-based-noise/

[58] Yue Huang, Borke Obada-Obieh, and Konstantin (Kosta) Beznosov. 2020. Amazon vs. My Brother: How Users of Shared Smart Speakers Perceive and Cope with Privacy Risks. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376529

[59] Umar Iqbal, Pouneh Nikkhah Bahrami, Rahmadi Trimananda, Hao Cui, Alexander Gamero-Garrido, Daniel Dubois, David Choffnes, Athina Markopoulou, Franziska Roesner, and Zubair Shafiq. 2022. Your Echos are Heard: Tracking, Profiling, and Ad Targeting in the Amazon Smart Speaker Ecosystem. https://doi.org/10.48550/ARXIV.2204.10920

[60] Yasha Iravantchi, Karan Ahuja, Mayank Goel, Chris Harrison, and Alanson Sample. 2021. PrivacyMic: Utilizing Inaudible Frequencies for Privacy Preserving Daily Activity Recognition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 198, 13 pages. https://doi.org/10.1145/3411764.3445169

[61] Yasha Iravantchi, Yi Zhao, Kenrick Kin, and Alanson P. Sample. 2023. SAWSense: Using Surface Acoustic Waves for Surface-bound Event Recognition. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 422, 18 pages. https://doi.org/10.1145/3544548.3580991

[62] Václav Janeček. 2018. Ownership of personal data in the Internet of Things. *Computer Law & Security Review* 34, 5 (2018), 1039–1052. https://doi.org/10.1016/j.clsr.2018.04.007

[63] Sandjar Kozubaev, Fernando Rochaix, Carl DiSalvo, and Christopher A. Le Dantec. 2019. Spaces and Traces: Implications of Smart Technology in Public Housing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300669

[64] Jacob Leon Kröger, Leon Gellrich, Sebastian Pape, Saba Rebecca Brause, and Stefan Ullrich. 2022. Personal information inference from voice recordings: User awareness and privacy concerns. *Proc. Priv. Enhancing Technol.* 2022, 1 (2022), 6–27. https://doi.org/10.2478/POPETS-2022-0002

[65] Deepak Kumar, Kelly Shen, Benton Case, Deepali Garg, Galina Alperovich, Dmitry Kuznetsov, Rajarshi Gupta, and Zakir Durumeric. 2019. All Things Considered: An Analysis of IoT Devices on Home Networks. In *28th USENIX Security Symposium (USENIX Security 19)*. USENIX Association, Santa Clara, CA, 1169–1185. https://www.usenix.org/conference/usenixsecurity19/presentation/kumar-deepak

[66] Lorenz Kustosch, Carlos Gañán, Mattis van 't Schip, Michel van Eeten, and Simon Parkin. 2023. Measuring Up to (Reasonable) Consumer Expectations: Providing an Empirical Basis for Holding IoT Manufacturers Legally Responsible. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 1487–1504. https://www.usenix.org/conference/usenixsecurity23/

presentation/kustosch
[67] Andrew Kwong, Wenyuan Xu, and Kevin Fu. 2019. Hard Drive of Hearing: Disks that Eavesdrop with a Synthesized Microphone. In *2019 IEEE Symposium on Security and Privacy (SP)*. 905–919. https://doi.org/10.1109/SP.2019.00008
[68] Marc Langheinrich. 2001. Privacy by Design — Principles of Privacy-Aware Ubiquitous Systems. In *Ubicomp 2001: Ubiquitous Computing*, Gregory D. Abowd, Barry Brumitt, and Steven Shafer (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 273–291.
[69] Hyunsoo Lee, Soowon Kang, and Uichin Lee. 2022. Understanding Privacy Risks and Perceived Benefits in Open Dataset Collection for Mobile Affective Computing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 61 (jul 2022), 26 pages. https://doi.org/10.1145/3534623
[70] Kyungjun Lee, Daisuke Sato, Saki Asakawa, Hernisa Kacorri, and Chieko Asakawa. 2020. Pedestrian Detection with Wearable Cameras for the Blind: A Two-way Perspective. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376398
[71] Lingkun Li, Manni Liu, Yuguang Yao, Fan Dang, Zhichao Cao, and Yunhao Liu. 2020. Patronus: preventing unauthorized speech recordings with support for selective unscrambling. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems* (Virtual Event, Japan) *(SenSys '20)*. Association for Computing Machinery, New York, NY, USA, 245–257. https://doi.org/10.1145/3384419.3430713
[72] Zhengxiong Li, Aditya Singh Rathore, Baicheng Chen, Chen Song, Zhuolin Yang, and Wenyao Xu. 2019. SpecEye: Towards Pervasive and Privacy-Preserving Screen Exposure Detection in Daily Life. (2019), 103–116. https://doi.org/10.1145/3307334.3326076
[73] Gary Liu and Nathan Malkin. 2022. Effects of Privacy Permissions on User Choices in Voice Assistant App Stores. *Proc. Priv. Enhancing Technol.* 2022, 4 (2022), 421–439. https://doi.org/10.56553/POPETS-2022-0116
[74] Yuchen Liu, Ziyu Xiang, Eun Ji Seong, Apu Kapadia, and Donald S. Williamson. 2021. Defending Against Microphone-Based Attacks with Personalized Noise. *Proc. Priv. Enhancing Technol.* 2021, 2 (2021), 130–150. https://doi.org/10.2478/POPETS-2021-0021
[75] Nathan Malkin, Joe Deatrick, Allen Tong, Primal Wijesekera, Serge Egelman, and David A. Wagner. 2019. Privacy Attitudes of Smart Speaker Users. *Proc. Priv. Enhancing Technol.* 2019, 4 (2019), 250–271. https://doi.org/10.2478/POPETS-2019-0068
[76] Sunil Manandhar, Kaushal Kafle, Benjamin Andow, Kapil Singh, and Adwait Nadkarni. 2022. Smart Home Privacy Policies Demystified: A Study of Availability, Content, and Coverage. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 3521–3538. https://www.usenix.org/conference/usenixsecurity22/presentation/manandhar
[77] Anna Maria Mandalari, Hamed Haddadi, Daniel J. Dubois, and David Choffnes. 2023. Protected or Porous: A Comparative Analysis of Threat Detection Capability of IoT Safeguards. In *2023 IEEE Symposium on Security and Privacy (SP)*. 3061–3078. https://doi.org/10.1109/SP46215.2023.10179282
[78] Shrirang Mare, Franziska Roesner, and Tadayoshi Kohno. 2020. Smart Devices in Airbnbs: Considering Privacy and Security for both Guests and Hosts. *Proc. Priv. Enhancing Technol.* 2020, 2 (2020), 436–458. https://doi.org/10.2478/POPETS-2020-0035
[79] Karola Marky, Nina Gerber, Michelle Gabriela Pelzer, Mohamed Khamis, and Max Mühlhäuser. 2022. "You offer privacy like you offer tea": Investigating Mechanisms for Improving Guest Privacy in IoT-Equipped Households. *Proc. Priv. Enhancing Technol.* 2022, 4 (2022), 400–420. https://doi.org/10.56553/POPETS-2022-0115
[80] Natalia Drozdiak Matt Day, Giles Turner. 2019. "Thousands of Amazon Workers Listen to Alexa Users' Conversations". Website. Retrieved July 27, 2022 from https://time.com/5568815/amazon-workers-listen-to-alexa/.
[81] Rachel McAmis and Tadayoshi Kohno. 2023. The Writing on the Wall and 3D Digital Twins: Personal Information in (not so) Private Real Estate. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 2169–2186. https://www.usenix.org/conference/usenixsecurity23/presentation/mcamis
[82] Abraham H. Mhaidli, Manikandan Kandadai Venkatesh, Yixin Zou, and Florian Schaub. 2020. Listen Only When Spoken To: Interpersonal Communication Cues as Smart Speaker Privacy Controls. *Proc. Priv. Enhancing Technol.* 2020, 2 (2020), 251–270. https://doi.org/10.2478/POPETS-2020-0026
[83] Jaron Mink, Amanda Rose Yuile, Uma Pal, Adam J Aviv, and Adam Bates. 2022. Users Can Deduce Sensitive Locations Protected by Privacy Zones on Fitness Tracking Apps. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 448, 21 pages. https://doi.org/10.1145/3491102.3502136
[84] Yuhong Nan, Xueqiang Wang, Luyi Xing, Xiaojing Liao, Ruoyu Wu, Jianliang Wu, Yifan Zhang, and XiaoFeng Wang. 2023. Are You Spying on Me? Large-Scale Analysis on IoT Data Exposure through Companion Apps. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 6665–6682. https://www.usenix.org/conference/usenixsecurity23/presentation/nan
[85] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.
[86] Joseph O'Hagan, Pejman Saeghe, Jan Gugenheimer, Daniel Medeiros, Karola Marky, Mohamed Khamis, and Mark McGill. 2023. Privacy-Enhancing Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 177 (jan 2023), 35 pages. https://doi.org/10.1145/3569501
[87] OpenAI. 2024. New and Improved Embedding Model. https://openai.com/blog/new-and-improved-embedding-model. Accessed: January 30, 2024.
[88] OpenAI. 2024. OpenAI Embeddings Use Cases. https://platform.openai.com/docs/guides/embeddings/use-cases. Accessed: January 30, 2024.
[89] Sunyup Park, Anna Lenhart, Michael Zimmer, and Jessica Vitak. 2023. "Nobody's Happy": Design Insights from Privacy-Conscious Smart Home Power Users on Enhancing Data Transparency, Visibility, and Control. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. USENIX Association, Anaheim, CA, 543–558. https://www.usenix.org/conference/soups2023/presentation/park
[90] Massimiliano Pastore and Antonio Calcagnì. 2019. Measuring Distribution Similarities Between Samples: A Distribution-Free Overlapping Index. *Frontiers in Psychology* 10 (2019). https://doi.org/10.3389/fpsyg.2019.01089
[91] Kavit Patel, Kyle Massa, Nithin Raghunathan, Heng Zhang, Ananth Iyer, and Saurabh Bagchi. 2020. Proactive privacy-preserving proximity prevention through bluetooth transceivers: poster abstract. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems* (Virtual Event, Japan) *(SenSys '20)*. Association for Computing Machinery, New York, NY, USA, 778–779. https://doi.org/10.1145/3384419.3430608
[92] Rishabh Poddar, Ganesh Ananthanarayanan, Srinath Setty, Stavros Volos, and Raluca Ada Popa. 2020. Visor: Privacy-Preserving Video Analytics as a Cloud Service. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, 1039–1056. https://www.usenix.org/conference/usenixsecurity20/presentation/poddar
[93] Sarah Prange, Ahmed Shams, Robin Piening, Yomna Abdelrahman, and Florian Alt. 2021. PriView– Exploring Visualisations to Support Users' Privacy Awareness. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 69, 18 pages. https://doi.org/10.1145/3411764.3445067
[94] Feng Qian and Bin Li. 2022. Boosting remote multi-user AR privacy through a magic rope. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services* (Portland, Oregon) *(MobiSys '22)*. Association for Computing Machinery, New York, NY, USA, 583–584. https://doi.org/10.1145/3498361.3538795
[95] Jianwei Qian, Haohua Du, Jiahui Hou, Linlin Chen, Taeho Jung, and Xiang-Yang Li. 2018. Hidebehind: Enjoy Voice Input with Voiceprint Unclonability and Anonymity. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems* (Shenzhen, China) *(SenSys '18)*. Association for Computing Machinery, New York, NY, USA, 82–94. https://doi.org/10.1145/3274783.3274855
[96] Emilee Rader. 2022. Normative and Non-Social Beliefs about Sensor Data: Implications for Collective Privacy Management. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. USENIX Association, Boston, MA, 653–670. https://www.usenix.org/conference/soups2022/presentation/rader
[97] Kopo Marvin Ramokapane, Caroline Bird, Awais Rashid, and Ruzanna Chitchyan. 2022. Privacy Design Strategies for Home Energy Management Systems (HEMS). In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 405, 15 pages. https://doi.org/10.1145/3491102.3517515
[98] Nisarg Raval, Ashwin Machanavajjhala, and Jerry Pan. 2019. Olympus: Sensor Privacy through Utility Aware Obfuscation. *Proc. Priv. Enhancing Technol.* 2019, 1 (2019), 5–25. https://doi.org/10.2478/POPETS-2019-0002
[99] K Andrew R Richards and Michael A Hemphill. 2018. A practical guide to collaborative qualitative data analysis. *Journal of Teaching in Physical education* 37, 2 (2018), 225–231.
[100] Mikko Rinta-Homi, Naser Hossein Motlagh, Agustin Zuniga, Huber Flores, and Petteri Nurmi. 2021. How Low Can You Go? Performance Trade-offs in Low-Resolution Thermal Sensors for Occupancy Detection: A Systematic Evaluation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 126 (sep 2021), 22 pages. https://doi.org/10.1145/3478104
[101] Hamada Rizk, Yuma Okochi, and Hirozumi Yamaguchi. 2022. Demonstrating hitonavi-u: a novel wearable LiDAR for human activity recognition. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking* (Sydney, NSW, Australia) *(MobiCom '22)*. Association for Computing Machinery, New York, NY, USA, 756–757. https://doi.org/10.1145/3495243.3558744
[102] Aafaq Sabir, Evan Lafontaine, and Anupam Das. 2022. Hey Alexa, Who Am I Talking to?: Analyzing Users' Perception and Awareness Regarding

Third-party Alexa Skills. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 447, 15 pages. https://doi.org/10.1145/3491102.3517510

[103] Nurani Saoda, Md Fazlay Rabbi Masum Billah, Victor Ariel Leal Sobral, and Bradford Campbell. 2023. SolarWalk Dataset: Occupant Identification Using Indoor Photovoltaic Harvester Output Voltage. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems* (Boston, Massachusetts) *(SenSys '22)*. Association for Computing Machinery, New York, NY, USA, 1031–1034. https://doi.org/10.1145/3560905.3567773

[104] Sarah Scheffler and Jonathan R. Mayer. 2023. SoK: Content Moderation for End-to-End Encryption. *Proc. Priv. Enhancing Technol.* 2023, 2 (2023), 403–429. https://doi.org/10.56553/POPETS-2023-0060

[105] Amazon Web Services. 2024. Amazon EC2 G3 Instances. https://aws.amazon.com/ec2/instance-types/g3/. Accessed: 2024-05-31.

[106] William Seymour, Martin J. Kraemer, Reuben Binns, and Max Van Kleek. 2020. Informing the Design of Privacy-Empowering Tools for the Connected Home. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376264

[107] Daniel J Solove. 2005. A taxonomy of privacy. *U. Pa. l. Rev.* 154 (2005), 477.

[108] Yunpeng Song, Yun Huang, Zhongmin Cai, and Jason I. Hong. 2020. I'm All Eyes and Ears: Exploring Effective Locators for Privacy Awareness in IoT Scenarios. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376585

[109] Sophie Stephenson, Bijeeta Pal, Stephen Fan, Earlence Fernandes, Yuhang Zhao, and Rahul Chatterjee. 2022. SoK: Authentication in Augmented and Virtual Reality. In *2022 IEEE Symposium on Security and Privacy (SP)*. 267–284. https://doi.org/10.1109/SP46214.2022.9833742

[110] Ke Sun, Chen Chen, and Xinyu Zhang. 2020. "Alexa, stop spying on me!": speech privacy protection against voice assistants. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems* (Virtual Event, Japan) *(SenSys '20)*. Association for Computing Machinery, New York, NY, USA, 298–311. https://doi.org/10.1145/3384419.3430727

[111] Madiha Tabassum, Tomasz Kosinski, and Heather Richter Lipford. 2019. "I don't own the data": End User Perceptions of Smart Home Device Data Practices and Risks. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. USENIX Association, Santa Clara, CA, 435–450. https://www.usenix.org/conference/soups2019/presentation/tabassum

[112] Madiha Tabassum, Jess Kropczynski, Pamela Wisniewski, and Heather Richter Lipford. 2020. Smart Home Beyond the Home: A Case for Community-Based Access Control. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376255

[113] Madiha Tabassum and Heather Lipford. 2023. Exploring privacy implications of awareness and control mechanisms in smart home devices. *Proc. Priv. Enhancing Technol.* 2023, 1 (2023), 571–588. https://doi.org/10.56553/POPETS-2023-0033

[114] Brian Jay Tang and Kang G. Shin. 2023. Eye-Shield: Real-Time Protection of Mobile Device Screen Information from Shoulder Surfing. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 5449–5466. https://www.usenix.org/conference/usenixsecurity23/presentation/tang

[115] Parth Kirankumar Thakkar, Shijing He, Shiyu Xu, Danny Yuxing Huang, and Yaxing Yao. 2022. "It would probably turn into a social faux-pas": Users' and Bystanders' Preferences of Privacy Awareness Mechanisms in Smart Homes. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 404, 13 pages. https://doi.org/10.1145/3491102.3502137

[116] Kurt Thomas, Devdatta Akhawe, Michael Bailey, Dan Boneh, Elie Bursztein, Sunny Consolvo, Nicola Dell, Zakir Durumeric, Patrick Gage Kelley, Deepak Kumar, Damon McCoy, Sarah Meiklejohn, Thomas Ristenpart, and Gianluca Stringhini. 2021. SoK: Hate, Harassment, and the Changing Landscape of Online Abuse. In *2021 IEEE Symposium on Security and Privacy (SP)*. 247–267. https://doi.org/10.1109/SP40001.2021.00028

[117] Piet De Vaere and Adrian Perrig. 2023. Hey Kimya, Is My Smart Speaker Spying on Me? Taking Control of Sensor Privacy Through Isolation and Amnesia. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 2401–2418. https://www.usenix.org/conference/usenixsecurity23/presentation/de-vaere

[118] Dirk van der Linden, Matthew Edwards, Irit Hadar, and Anna Zamansky. 2020. Pets without PETs: on pet owners' under-estimation of privacy concerns in pet wearables. *Proc. Priv. Enhancing Technol.* 2020, 1 (2020), 143–164. https://doi.org/10.2478/POPETS-2020-0009

[119] Lev Velykoivanenko, Kavous Salehzadeh Niksirat, Noé Zufferey, Mathias Humbert, Kévin Huguenin, and Mauro Cherubini. 2022. Are Those Steps Worth Your Privacy? Fitness-Tracker Users' Perceptions of Privacy and Utility. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 181 (dec 2022), 41 pages. https://doi.org/10.1145/3494960

[120] Swaathi Vetrivel, Veerle van Harten, Carlos H. Ganan, Michel van Eeten, and Simon Parkin. 2023. Examining Consumer Reviews to Understand Security and Privacy Issues in the Market of Smart Home Devices. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 1523–1540. https://www.usenix.org/conference/usenixsecurity23/presentation/vetrivel

[121] Yuntao Wang, Zirui Cheng, Xin Yi, Yan Kong, Xueyang Wang, Xuhai Xu, Yukang Yan, Chun Yu, Shwetak Patel, and Yuanchun Shi. 2023. Modeling the Trade-off of Privacy Preservation and Activity Recognition on Low-Resolution Images. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 589, 15 pages. https://doi.org/10.1145/3544548.3581425

[122] Zixin Wang, Danny Yuxing Huang, and Yaxing Yao. 2023. Exploring Tenants' Preferences of Privacy Negotiation in Airbnb. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 535–551. https://www.usenix.org/conference/usenixsecurity23/presentation/wang-zixin

[123] Zhiwei Wang, Yihui Yan, Yueli Yan, Huangxun Chen, and Zhice Yang. 2022. CamShield: Securing Smart Cameras through Physical Replication and Isolation. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 3467–3484. https://www.usenix.org/conference/usenixsecurity22/presentation/wang-zhiwei

[124] Mark Weiser. 1991. The Computer for the 21st Century. *Scientific american* 265, 3 (1991), 94–105.

[125] Maximiliane Windl, Albrecht Schmidt, and Sebastian S. Feger. 2023. Investigating Tangible Privacy-Preserving Mechanisms for Future Smart Homes. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 70, 16 pages. https://doi.org/10.1145/3544548.3581167

[126] Yuxi Wu, W. Keith Edwards, and Sauvik Das. 2022. SoK: Social Cybersecurity. In *2022 IEEE Symposium on Security and Privacy (SP)*. 1863–1879. https://doi.org/10.1109/SP46214.2022.9833757

[127] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang. 2019. Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300428

[128] Hyunwoo Yu, Jaemin Lim, Kiyeon Kim, and Suk-Bok Lee. 2018. Pinto: Enabling Video Privacy for Commodity IoT Cameras. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security* (Toronto, Canada) *(CCS '18)*. Association for Computing Machinery, New York, NY, USA, 1089–1101. https://doi.org/10.1145/3243734.3243830

[129] Igor Zavalyshyn, Axel Legay, Annanda Rath, and Etienne Rivière. 2022. SoK: Privacy-enhancing Smart Home Hubs. *Proc. Priv. Enhancing Technol.* 2022, 4 (2022), 24–43. https://doi.org/10.56553/POPETS-2022-0097

[130] Eric Zeng and Franziska Roesner. 2019. Understanding and Improving Security and Privacy in Multi-User Smart Homes: A Design Exploration and In-Home User Study. In *28th USENIX Security Symposium (USENIX Security 19)*. USENIX Association, Santa Clara, CA, 159–176. https://www.usenix.org/conference/usenixsecurity19/presentation/zeng

[131] Hanbin Zhang, Chen Song, Aosen Wang, Chenhan Xu, Dongmei Li, and Wenyao Xu. 2019. PDVocal: Towards Privacy-preserving Parkinson's Disease Detection using Non-speech Body Sounds. In *The 25th Annual International Conference on Mobile Computing and Networking* (Los Cabos, Mexico) *(MobiCom '19)*. Association for Computing Machinery, New York, NY, USA, Article 16, 16 pages. https://doi.org/10.1145/3300061.3300125

[132] Ruidong Zhang, Ke Li, Yihong Hao, Yufan Wang, Zhengnan Lai, François Guimbretière, and Cheng Zhang. 2023. EchoSpeech: Continuous Silent Speech Recognition on Minimally-obtrusive Eyewear Powered by Acoustic Sensing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 852, 18 pages. https://doi.org/10.1145/3544548.3580801

[133] Shikun Zhang, Yuanyuan Feng, Lujo Bauer, Lorrie Faith Cranor, Anupam Das, and Norman M. Sadeh. 2021. "Did you know this camera tracks your mood?": Understanding Privacy Expectations and Preferences in the Age of Video Analytics. *Proc. Priv. Enhancing Technol.* 2021, 2 (2021), 282–304. https://doi.org/10.2478/POPETS-2021-0028

[134] Shikun Zhang, Yuanyuan Feng, Yaxing Yao, Lorrie Faith Cranor, and Norman Sadeh. 2022. How Usable Are iOS App Privacy Labels? *Proc. Priv. Enhancing Technol.* 2022, 4 (2022), 204–228. https://doi.org/10.56553/POPETS-2022-0106

[135] Yuhang Zhao, Yaxing Yao, Jiaru Fu, and Nihan Zhou. 2023. "If sighted people know, I should be able to know:" Privacy Perceptions of Bystanders with Visual Impairments around Camera-based Technology. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 4661–4678. https://www.usenix.org/conference/usenixsecurity23/presentation/zhao-yuhang

[136] Tongqing Zhou, Zhiping Cai, and Fang Liu. 2021. The Crowd Wisdom for Location Privacy of Crowdsensing Photos: Spear or Shield? *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 142 (sep 2021), 23 pages. https:

//doi.org/10.1145/3478106

[137] Shuai Zhu, Thiemo Voigt, Daniel F. Perez-Ramirez, and Joakim Eriksson. 2021. A Low-resolution infrared thermal dataset and potential privacy-preserving applications. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems* (Coimbra, Portugal) *(SenSys '21)*. Association for Computing Machinery, New York, NY, USA, 552–555. https://doi.org/10.1145/3485730.3493692

[138] M. Zurko and J. Haney. 2023. Usable Security and Privacy for Security and Privacy Workers. *IEEE Security & Privacy* 21, 01 (jan 2023), 8–10. https://doi.org/10.1109/MSEC.2022.3221855

A Appendix

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

A.1 Codebook

Below is the entire codebook, developed using the methodology described in Section 3. We utilized this codebook to uncover the themes described in Section 4 and quantify our findings detailed in Section 5.

	Venue	Year	Citation	Title	Primary Identity	P1: Notice	P2: Choice and Consent	P3: Anonymity and Pseudonymity	P4: Proximity and Locality	P5: Adequate Security	P6: Access and Resource	Privacy and security factors:	Privacy mindset:	Technical Evaluation	User Evaluation	Feasibility Evaluation	System Evaluation
1	CHI	2021	[60]	PrivacyMic: Utilizing Inaudible Frequencies fo...	Privacy-Aware Sensing	●	●	●	○	●	●	sanitization	better than nothing	yes	yes	built system	both
2	SenSys	2021	[35]	Social Distancing Compliance Monitoring for CO...	Privacy-Aware Sensing	○	○	●	●	○	○	purpose	defensive	assumes	no	built system	performance
3	IMWUT	Vol. 6	[86]	Privacy-Enhancing Technology and Everyday Augm...	Usable Privacy	●	●	●	●	○	○	control	awareness	no	yes	no	no
4	IMWUT	Vol. 3	[8]	To Mask or Not to Mask?: Balancing Privacy wit...	Privacy-Aware Sensing	○	○	●	●	○	○	sanitization	better than nothing	yes	yes	artifact representation	no
5	MobiSys	2019	[72]	SpecEye: Towards Pervasive and Privacy-Preserv...	Privacy-Aware Sensing	○	○	●	●	●	●	purpose	all or nothing	assumes	no	built system	performance
6	MobiCom	2019	[131]	PDVocal: Towards Privacy-preserving Parkinson'...	Privacy-Aware Sensing	○	●	○	●	●	●	sanitization	all or nothing	yes	no	built system	both
7	IMWUT	Vol. 5	[100]	How Low Can You Go?: Performance Trade-offs in...	Privacy-Aware Sensing	○	○	●	●	○	●	granularity	better than nothing	assumes	no	artifact representation	no
8	CHI	2023	[37]	Impact of Privacy Protection Methods of Lifelo...	Usable Privacy	○	○	●	○	○	○	sanitization	better than nothing	no	yes	no	no
9	MobiCom	2022	[101]	Demonstrating hitonavi-μ: a novel wearable LiD...	Privacy-Aware Sensing	○	○	●	●	○	●	purpose	all or nothing	assumes	no	no	no
10	CHI	2020	[28]	Wearable Microphone Jamming	Privacy-Aware Sensing	○	○	●	●	○	○	purpose	offensive	yes	yes	built system	both
11	IMWUT	Vol. 5	[34]	Smart Webcam Cover: Exploring the Design of an...	Privacy-Aware Sensing	●	●	○	○	●	○	control	all or nothing	assumes	yes	built system	no
12	CHI	2023	[125]	Investigating Tangible Privacy-Preserving Mech...	Usable Privacy	●	●	○	○	●	○	sanitization	all or nothing	assumes	yes	artifact representation	no
13	SenSys	2019	[9]	Privacy preserving speech analysis using emoti...	Privacy-Aware Sensing	○	○	●	○	○	○	sanitization	better than nothing	yes	no	artifact representation	no
14	CHI	2020	[108]	I'm All Eyes and Ears: Exploring Effective Loc...	Usable Privacy	●	○	○	○	○	○	control	awareness	no	yes	artifact representation	no
15	CHI	2023	[132]	EchoSpeech: Continuous Silent Speech Recogniti...	Privacy-Aware Sensing	○	○	○	●	○	●	sanitization	better than nothing	assumes	no	built system	no
16	CCS	2018	[128]	Pinto: Enabling Video Privacy for Commodity Io...	Privacy-Aware Sensing	○	●	●	○	●	●	authentication	defensive	yes	no	built system	both
17	MobiSys	2022	[94]	Boosting remote multi-user AR privacy through ...	Usable Privacy	●	●	●	●	○	●	control	defensive	no	no	artifact representation	no
18	SenSys	2018	[95]	Hidebehind: Enjoy Voice Input with Voiceprint ...	Privacy-Aware Sensing	○	○	●	○	○	○	sanitization	defensive	yes	no	built system	both
19	CHI	2023	[121]	Modeling the Trade-off of Privacy Preservation...	Privacy-Aware Sensing	○	○	●	●	○	●	granularity	better than nothing	assumes	no	built system	no
20	SenSys	2020	[110]	"Alexa, stop spying on me!": speech privacy pr...	Privacy-Aware Sensing	○	○	●	●	○	○	sanitization	offensive	yes	no	built system	both
21	CHI	2022	[102]	Hey Alexa, Who Am I Talking to?: Analyzing Use...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	discussion	no
22	MobiCom	2021	[56]	Notification privacy protection via unobtrusiv...	Privacy-Aware Sensing	○	○	○	○	●	○	authentication	defensive	yes	no	built system	both
23	CHI	2021	[93]	PriView– Exploring Visualisations to Support U...	Usable Privacy	●	○	○	○	○	○	control	awareness	no	yes	no	no
24	IMWUT	Vol. 5	[36]	Is Someone Listening?: Audio-Related Privacy P...	Usable Privacy	●	○	○	○	○	○	control	awareness	no	yes	no	no
25	SenSys	2020	[71]	Patronus: preventing unauthorized speech recor...	Privacy-Aware Sensing	○	●	●	●	○	○	sanitization	offensive	yes	no	built system	both
26	IMWUT	Vol. 6	[69]	Understanding Privacy Risks and Perceived Bene...	Usable Privacy	●	●	○	○	●	●	sanitization	better than nothing	no	yes	discussion - author expertise	no
27	SenSys	2021	[137]	A Low-resolution infrared thermal dataset and ...	Privacy-Aware Sensing	○	○	●	●	○	●	granularity	better than nothing	assumes	no	no	no
28	CHI	2021	[25]	"It did not give me an option to decline": A L...	Usable Privacy	●	●	○	○	●	○	control	awareness	no	yes	no	no
29	CHI	2019	[52]	Can Privacy Be Satisfying?: On Improving Viewe...	Usable Privacy	○	●	○	○	●	○	sanitization	better than nothing	no	yes	no	no
30	CHI	2022	[97]	Privacy Design Strategies for Home Energy Mana...	Usable Privacy	●	●	○	○	●	●	control	awareness	no	yes	no	no
31	CHI	2023	[39]	Understanding People's Concerns and Attitudes ...	Usable Privacy	●	●	●	●	●	●	control	awareness	no	yes	discussion - author expertise	no
32	IMWUT	Vol. 4	[17]	Where You Go Matters: A Study on the Privacy I...	Usable Privacy	●	●	○	●	○	○	control	awareness	no	yes	no	no
33	IMWUT	Vol. 5	[119]	Are Those Steps Worth Your Privacy?: Fitness-T...	Usable Privacy	○	●	●	●	○	●	granularity	better than nothing	no	yes	no	no
34	CHI	2022	[83]	Users Can Deduce Sensitive Locations Protected...	Usable Privacy	○	●	●	○	●	○	sanitization	better than nothing	no	yes	artifact representation	no
35	CCS	2022	[32]	A Run a Day Won't Keep the Hacker Away: Infere...	Privacy-Aware Sensing	○	○	●	○	●	○	granularity	better than nothing	yes	no	built system	both
36	CHI	2023	[11]	Exploring the Negotiation Behaviors of Owners ...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	acknowledge author expertise needed	no
37	MobiSys	2021	[54]	LensCap: split-process framework for fine-grai...	Usable Privacy	○	●	○	●	●	●	authentication	defensive	no	yes	artifact representation	no
38	IMWUT	Vol. 5	[136]	The Crowd Wisdom for Location Privacy of Crowd...	Privacy-Aware Sensing	○	○	●	●	○	○	granularity	defensive	yes	no	built system	both
39	IMWUT	Vol. 5	[16]	Every Byte Matters: Traffic Analysis of Blueto...	Privacy-Aware Sensing	○	○	●	○	●	○	authentication	defensive	yes	no	built system	both
40	CHI	2019	[42]	Exploring How Privacy and Security Factor into...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	artifact representation	no
41	SenSys	2020	[91]	Proactive privacy-preserving proximity prevent...	Privacy-Aware Sensing	○	○	●	○	○	○	purpose	defensive	assumes	no	built system	no
42	CHI	2019	[63]	Spaces and Traces: Implications of Smart Techn...	Usable Privacy	●	●	●	○	○	●	control	awareness	no	yes	no	no
43	MobiCom	2019	[27]	Keep Others from Peeking at Your Mobile Device...	Privacy-Aware Sensing	○	○	●	●	●	○	authentication	defensive	yes	no	built system	both
44	CHI	2019	[48]	Who's In Control?: Interactions In Multi-User ...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	no	no
45	CHI	2020	[112]	Smart Home Beyond the Home: A Case for Communi...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	no	no
46	CHI	2022	[7]	Exploring Bystanders' Privacy Concerns with Sm...	Usable Privacy	●	●	●	○	○	○	control	awareness	no	yes	no	no
47	CHI	2021	[44]	A Design Space for Privacy Choices: Towards Me...	Usable Privacy	●	●	○	●	○	○	control	awareness	no	no	no	no
48	CHI	2022	[115]	"It would probably turn into a social faux-pas...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	artifact representation	no
49	CHI	2020	[58]	Amazon vs. My Brother: How Users of Shared Sma...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	no	no
50	CHI	2019	[127]	Defending My Castle: A Co-Design Study of Priv...	Usable Privacy	●	●	○	○	●	○	control	awareness	no	yes	no	no
51	CHI	2020	[106]	Informing the Design of Privacy-Empowering Too...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	no	no
52	CHI	2020	[70]	Pedestrian Detection with Wearable Cameras for...	Usable Privacy	●	●	○	●	○	○	control	awareness	no	yes	built system	performance
53	IMWUT	Vol. 5	[4]	Ok Google, What Am I Doing?: Acoustic Activity...	Privacy-Aware Sensing	○	○	○	○	○	●	granularity	better than nothing	no	no	built system	performance
54	SenSys	2022	[103]	SolarWalkDataset: Occupant Identification Usin...	Privacy-Aware Sensing	○	○	●	●	●	●	purpose	all or nothing	assumes	no	built system	performance
55	CHI	2023	[61]	SAWSense: Using Surface Acoustic Waves for Sur...	Privacy-Aware Sensing	○	○	●	●	●	●	purpose	better than nothing	yes	no	built system	both
56	PETS	2019	[47]	Privacy Protection for Audio Sensing Against M...	Privacy-Aware Sensing	○	●	●	●	○	○	authentication	offensive	yes	no	built system	both
57	PETS	2020	[118]	Pets without PETs: on pet owners' under-estima...	Usable Privacy	●	●	○	●	○	○	control	awareness	no	yes	no	no
58	USENIX Security	2019	[130]	Understanding and Improving Security and Priva...	Usable Privacy	●	●	○	●	○	○	control	awareness	no	yes	artifact representation	no
59	USENIX Security	2023	[66]	Measuring Up to (Reasonable) Consumer Expectat...	Usable Privacy	○	●	○	○	●	○	authentication	awareness	no	yes	no	no
60	SOUPS	2019	[2]	More than Smart Speakers: Security and Privacy...	Usable Privacy	●	●	○	○	●	●	authentication	defensive	no	yes	discussion	no
61	PETS	2022	[64]	Personal information inference from voice reco...	Usable Privacy	●	●	●	○	●	○	control	awareness	no	yes	discussion	no
62	PETS	2022	[129]	SoK: Privacy-enhancing Smart Home Hubs	Usable Privacy	○	●	○	○	●	●	control	awareness	no	no	discussion	no
63	PETS	2021	[133]	"Did you know this camera tracks your mood?": ...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	discussion	no
64	SOUPS	2022	[18]	Balancing Power Dynamics in Smart Homes: Nanni...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	no	no
65	USENIX Security	2023	[33]	Powering for Privacy: Improving User Trust in ...	Privacy-Aware Sensing	●	●	○	○	○	○	control	all or nothing	no	yes	built system	no
66	PETS	2019	[13]	Keeping the Smart Home Private with Smart(er) ...	Privacy-Aware Sensing	○	○	●	○	●	○	authentication	defensive	yes	no	built system	both
67	SOUPS	2019	[10]	"There is nothing that I need to keep secret":...	Usable Privacy	●	●	○	●	○	○	control	awareness	no	yes	discussion	no
68	SOUPS	2022	[96]	Normative and Non-Social Beliefs about Sensor ...	Usable Privacy	○	●	○	○	○	●	control	awareness	no	yes	no	no
69	PETS	2020	[82]	Listen Only When Spoken To: Interpersonal Comm...	Usable Privacy	●	●	○	●	○	○	authentication	awareness	no	yes	built system	both
70	PETS	2019	[98]	Olympus: Sensor Privacy through Utility Aware ...	Privacy-Aware Sensing	○	○	●	○	●	○	sanitization	better than nothing	yes	no	built system	both
71	USENIX Security	2020	[92]	Visor: Privacy-Preserving Video Analytics as a...	Privacy-Aware Sensing	○	○	●	○	●	○	sanitization	defensive	yes	no	built system	both
72	USENIX Security	2023	[114]	Eye-Shield: Real-Time Protection of Mobile Dev...	Privacy-Aware Sensing	○	●	○	●	●	○	authentication	defensive	yes	yes	built system	both
73	USENIX Security	2023	[117]	Hey Kimya, Is My Smart Speaker Spying on Me? T...	Privacy-Aware Sensing	○	○	○	○	●	○	control	defensive	yes	no	built system	both
74	SOUPS	2021	[26]	PowerCut and Obfuscator: An Exploration of the...	Privacy-Aware Sensing	●	●	○	○	●	○	authentication	all or nothing	no	yes	built system	no
75	PETS	2022	[19]	Building a Privacy-Preserving Smart Camera System	Usable Privacy	○	●	○	○	●	●	authentication	defensive	yes	yes	built system	yes
76	PETS	2022	[20]	A Multi-Region Investigation of the Perception...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	discussion	no
77	USENIX Security	2022	[123]	CamShield: Securing Smart Cameras through Phys...	Privacy-Aware Sensing	○	○	●	○	●	○	authentication	defensive	yes	no	built system	yes
78	USENIX Security	2022	[76]	Smart Home Privacy Policies Demystified: A Stu...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	no	discussion	no
79	SOUPS	2020	[15]	Do Privacy and Security Matter to Everyone? Qu...	Usable Privacy	●	●	●	○	●	○	control	awareness	no	yes	no	no
80	SOUPS	2023	[89]	Nobody's Happy: Design Insights from Privacy-C...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	discussion	no
81	PETS	2021	[30]	"I would have to evaluate their objections": P...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	no	no
82	PETS	2022	[79]	"You offer privacy like you offer tea": Invest...	Usable Privacy	●	●	●	○	○	●	control	awareness	no	yes	no	no
83	SOUPS	2020	[24]	Innovation Inaction or In Action? The Role of ...	Usable Privacy	●	●	○	●	●	○	control	awareness	no	yes	discussion	no
84	SOUPS	2019	[111]	I don't own the data: End User Perceptions of ...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	discussion	no
85	SOUPS	2023	[53]	Exploring the Usability, Security, and Privacy...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	no	no
86	PETS	2023	[113]	Exploring privacy implications of awareness an...	Usable Privacy	●	●	○	●	●	●	control	awareness	no	yes	no	no
87	PETS	2021	[74]	Defending Against Microphone-Based Attacks wit...	Privacy-Aware Sensing	○	●	○	○	●	○	purpose	defensive	yes	yes	built system	yes
88	PETS	2019	[14]	"What if?" Predicting Individual Users' Smart ...	Usable Privacy	●	●	○	○	●	○	control	awareness	no	yes	built system	no
89	USENIX Security	2021	[51]	"It's the Company, the Government, You and I":...	Usable Privacy	○	●	○	○	●	●	control	awareness	no	yes	discussion	no
90	PETS	2020	[78]	Smart Devices in Airbnbs: Considering Privacy ...	Usable Privacy	●	○	●	○	○	●	purpose	awareness	no	yes	discussion	no
91	USENIX Security	2023	[120]	Examining Consumer Reviews to Understand Secur...	Usable Privacy	○	●	○	○	●	○	control	awareness	no	yes	no	no
92	IEEE S&P	2021	[40]	Which Privacy and Security Attributes Most Imp...	Usable Privacy	●	●	●	○	●	○	control	awareness	no	yes	discussion	no
93	PETS	2019	[75]	Privacy Attitudes of Smart Speaker Users	Usable Privacy	○	●	○	○	○	●	control	awareness	no	yes	discussion	no
94	PETS	2022	[12]	Exploring the Privacy Concerns of Bystanders i...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	discussion	no
95	IEEE S&P	2020	[38]	Ask the Experts: What Should Be on an IoT Priv...	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	artifact representation	no
96	NDSS	2023	[57]	InfoMasker: Preventing Eavesdropping Using Pho...	Privacy-Aware Sensing	○	●	●	●	●	○	authentication	offensive	yes	no	built system	both
97	USENIX Security	2023	[41]	Are Consumers Willing to Pay for Security and ...	Usable Privacy	●	●	●	○	●	○	control	awareness	no	yes	discussion	no
98	IEEE S&P	2022	[31]	Privacy-from-Birth: Protecting Sensed Data fro...	Privacy-Aware Sensing	○	○	○	○	●	○	authentication	defensive	yes	no	built system	both
99	PETS	2019	[49]	Investigating People's Privacy Risk Perception	Usable Privacy	●	●	○	○	○	○	control	awareness	no	yes	no	no
100	IEEE S&P	2019	[67]	Hard Drive of Hearing: Disks that Eavesdrop wi...	Privacy-Aware Sensing	○	○	○	○	●	○	granularity	offensive	flaws	no	built system	no
101	PETS	2023	[46]	Speculative Privacy Concerns about AR Glasses ...	Usable Privacy	●	●	○	○	●	●	control	awareness	no	yes	no	no
102	USENIX Security	2023	[135]	"If sighted people know, I should be able to k...	Usable Privacy	●	●	○	●	●	○	control	awareness	no	yes	discussion	no
103	USENIX Security	2019	[65]	All Things Considered: An Analysis of IoT Devi...	Privacy-Aware Sensing	○	○	○	○	●	●	authentication	defensive	flaws	no	discussion	no
104	USENIX Security	2023	[84]	Are You Spying on Me? Large-Scale Analysis on ...	Privacy-Aware Sensing	●	●	○	○	○	●	control	awareness	flaws	no	no	no
105	SOUPS	2021	[3]	Comparing Security and Privacy Attitudes Among...	Usable Privacy	●	○	○	○	○	○	control	awareness	no	yes	no	no
106	USENIX Security	2023	[81]	The Writing on the Wall and 3D Digital Twins: ...	Usable Privacy	○	○	●	○	●	●	sanitization	defensive	flaws	no	no	no
107	USENIX Security	2023	[6]	Examining Power Dynamics and User Privacy in S...	Usable Privacy	●	●	●	○	○	○	control	awareness	no	yes	no	no
108	IEEE S&P	2022	[55]	AccEar: Accelerometer Acoustic Eavesdropping w...	Privacy-Aware Sensing	○	●	○	○	●	●	granularity	defensive	yes	no	built system	both
109	IEEE S&P	2023	[23]	No Privacy in the Electronics Repair Industry	Privacy-Aware Sensing	○	●	○	○	●	○	authentication	defensive	flaws	yes	discussion	no
110	IEEE S&P	2023	[77]	Protected or Porous: A Comparative Analysis of...	Privacy-Aware Sensing	○	○	○	○	●	○	authentication	defensive	flaws	no	no	no
111	USENIX Security	2023	[122]	Exploring Tenants' Preferences of Privacy Nego...	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	discussion	no
112	PETS	2022	[73]	Effects of Privacy Permissions on User Choices...	Usable Privacy	○	●	○	○	●	●	control	awareness	no	yes	built system	no
113	PETS	2022	[134]	How Usable Are iOS App Privacy Labels?	Usable Privacy	●	●	○	○	○	●	control	awareness	no	yes	discussion	no
114	USENIX Security	2020	[5]	"I am uncomfortable sharing what I can't see":...	Usable Privacy	○	○	●	○	●	○	control	awareness	no	yes	no	no

Table 6: The codebook we generated to perform our qualitative and quantitative analysis.