SoK: A Framework and Guide for Human-Centered Threat Modeling in Security and Privacy Research

Warda Usman
Department of Computer Science
Brigham Young University

Daniel Zappala
Department of Computer Science
Brigham Young University

***Abstract*—Human-centered threat modeling is a practice that researchers use to identify security and privacy threats to people, as well as ways to mitigate those threats. Often this may be the first step toward understanding the security and privacy needs, perspectives, experiences, and practices of a group or community, so that researchers can learn how to better improve their overall safety. However, research in this area is relatively ad hoc as compared to the more well-developed field of threat modeling for systems, leading to a fragmented and incomplete understanding of how researchers should engage in this endeavor. The goal of this work is to systematize the practice of human-centered threat modeling, identifying the core components of a human-centered threat modeling exercise by studying the practices of researchers in the area. We gathered a corpus of 78 papers in this area, using qualitative analysis to understand the practices used by researchers to elicit a threat model. Our results include a framework for human-centered threat modeling, a guide for using the framework that is grounded in best practices, and a description of how human-centered threat modeling differs from systems threat modeling. Our work can be used to guide new and experienced researchers in the field as they work to center human safety in their practices.**

1. Introduction

A growing body of work within the security and privacy research community focuses on studying threat models of people. This body of research encompasses a range of approaches. Some studies draw from systems threat modeling—commonly used in both industry and academia—where employees or researchers systematically identify vulnerabilities in a *system* and then seek to design or modify the system to mitigate or eliminate those vulnerabilities. Other studies draw from rich traditions within Human-Computer Interaction (HCI), focusing on how people interact with technology and how these interactions may lead to potential harms or vulnerabilities. The common focus in this work is on *people*, rather than systems.

To distinguish this work, we have coined the term *human-centered threat modeling*, a field in which researchers seek to understand how people may be vulnerable to a variety of threats or harms, as well as the practices people use to keep themselves safe. This breadth of approaches has led to an array of research involving researchers from systems, security, HCI, and the social sciences, aiming to understand technology-related harms and practices. The resulting research spans a wide variety of topics, including studies of at-risk users [25], [55], [77], [45], [18], risks from a particular technology [43], [57], [26], particular experiences or harms [93], [86], [81], and risks to populations in particular locations [70], [3], [94]. One lesson of all of this work is that threat modeling is needed in a wide range of situations because people themselves are highly diverse in terms of their goals, needs, perspectives, experiences, and practices. This diversity has led to a rich field of research that seeks to understand the impact of technology on people's lives and how we can better help people achieve security and privacy.

However, human-centered threat modeling has been pursued in a relatively ad hoc fashion as compared to the more well-established field of threat modeling for systems. Many researchers do not even use the term "threat model", yet are clearly engaged in a similar process of assessing risks and understanding the protective decisions taken by participants in their studies [56], [43], [50], [22]. Conversely, some researchers borrow threat modeling from the security field, but use conflicting terminology, with many defining it as narrowly consisting of only a listing of threats or potential harms [85], [86], [77], while others include defenses or other related items [38], [90].

These disparate approaches and terminology have led to a fragmented and incomplete understanding of the process of human-centered threat modeling. This presents several challenges. First, it is difficult for someone new to the field to understand how to conduct a threat modeling exercise with people. Second, it may be difficult to build a complete understanding of a given population and their needs, as the methods and definitions used are not uniform. In contrast, systems threat modeling has matured through clear definitions and structured approaches [75], [24], [36] which has allowed the field to standardize best practices and effectively train new practitioners. A lack of systematization for human-centered threat modeling makes it difficult to find relevant literature and hinders the sharing of knowledge within the field. Overall, this inconsistency leads to confusion and limits our understanding of how threat modeling should be applied to individuals or communities and their unique situations.

To bridge this gap, we studied the human-centered threat modeling process used by security and privacy researchers, defining and systematizing human-centered threat modeling practices. Our corpus consists of 78 papers from conferences spanning the security, privacy, and human-computer interaction (HCI) communities. We broadly include all papers that study threats *to people, as perceived by people*, so that we can generalize across diverse practices in the field. We include both security and privacy papers in this corpus, because both the general public and researchers often do not clearly differentiate between the two concepts, and because threats people face may be intertwined among security and privacy concerns. We use a qualitative, inductive analysis, along with techniques borrowed from concept mapping, to ground our analysis in how researchers conduct this work. Our research was initially guided by the following research questions:

RQ1: How do researchers define threat modeling when they use it in a human context?
RQ2: How do researchers identify or elicit a human-centered threat model from their subjects?
RQ3: What are the key components typically included in a human-centered threat modeling exercise?

Our systematization makes the following contributions:

- We present a framework for human-centered threat modeling that researchers can use to better address the diverse risks that people face. The framework consists of a definition of human-centered threat modeling and a systematic listing of components that researchers have used in their study. Each component includes a set of factors for researchers to consider in their approach, along with citations of papers that use this factor. The framework derives directly from researcher practices in our corpus and includes clarifying examples and citations to point researchers to papers that illustrate how each component can contribute to a human-centered threat modeling exercise.
- We provide a guide for researchers on how to use the framework, so that researchers who are new to the field can find important lessons from prior work. We explain how the framework can help a researcher take a structured approach to human-centered threat modeling, list basic principles, and explain important considerations that are grounded in our corpus. Chief among these lessons is that each population has its own context and threats, so the researchers should customize the framework to fit the people or community with whom they are engaging.
- We compare human-centered threat modeling with systems threat modeling so that researchers can see what translates over and how this method differs in key areas.
- We conclude by highlighting examples of how threat modeling can lead to centering human safety in our research.

2. Background and Motivation

We begin by discussing security and privacy threat modeling within the context of hardware and software systems. We then introduce the concept of human-centered threat modeling and argue for the need to systematize its process.

2.1. Systems Threat Modeling

Threat modeling is the process of analyzing a hardware or software system to look for security vulnerabilities and ways to prevent or mitigate those threats.[1] A broad definition of threat modeling includes four basic questions [75]: *What are you building?*, *What can go wrong?*, *What should you do about those things that can go wrong?*, and *Did you do a decent job of analysis?*

Threat modeling has classically focused on identifying and patching security vulnerabilities, though it can also be used throughout the software lifecycle to ensure a system is built with a secure design. As indicated by the questions above, the process usually starts by first understanding the system that one is building or maintaining, because threats will be specific to the system design. A threat modeling team then seeks to identify threats to the system. This could be done with a framework that identifies common security properties, such as STRIDE (spoofing, tampering, repudiation, information disclosure, denial of service, elevation of privilege) [17]; an attack tree [68]; or a library of threats like the OWASP top ten web vulnerabilities[2]. Threat modeling at this stage often focuses on identifying adversaries [69] and the vulnerabilities they may attack. Once threats are identified, a team can prioritize threats and decide what approach to take. An organization may use a formal method to assess the risk of a threat [24], [36] and to prioritize which threats to address. Finally, testing can be used to verify that threats have been mitigated, such as through penetration testing or a quality assurance process. Some tools like the OWASP Threat Dragon[3] can help to automate parts of the threat modeling process.

Privacy threat modeling is an emergent area that shares similarities with security threat modeling. Privacy threats are focused on information about a person, such as found in Solove's taxonomy [79]: information collection, information processing, information dissemination, and invasion. These threats are often best understood in context, an approach taken by Nissenbaum's contextual integrity framework [59], [49]. The framework is intended to help understand why a particular information flow may be appropriate or problematic, based on societal norms. Understanding the context includes examining the type of data being held, the owner of that information, the sender and receiver of the information, and the transmission principle, which covers any constraints on the information being exchanged. Formal approaches

1. https://owasp.org/www-community/Threat_Modeling
2. https://owasp.org/www-project-top-ten/
3. https://owasp.org/www-project-threat-dragon/

to privacy threat modeling include frameworks that identify common privacy issues, such as LINDDUN (linkability, identifiability, non-repudiation, detectability, disclosure of information, unawareness, non-compliance) [20] and MITRE PANOPTIC [74].

2.2. Human-Centered Threat Modeling

A distinguishing feature of human-centered threat modeling is that it focuses on *threats to people* rather than threats to a system. We use the term "human-centered threat modeling" because work in this space is centered on people—their goals, needs, perspectives, experiences, and practices.

Human-centered threat modeling is akin to participatory threat modeling. As Slupska et al. describe when introducing the method, *Rather than dictating what threats citizens should be worrying about, this project develops a method ... for eliciting and listening to citizens' concerns to expand the scope of threats considered in cybersecurity* [78]. This method directly involves participants as research partners, such as through a workshop, so that they are truly centered in the research [77]. We use the term human-centered threat model to encompass a broader set of practices that similarly center people's concerns.

The span of human-centered threat modeling research is vast. Research in this area has examined threats to at-risk users, such as older adults [25], people with Parkinson's [55], migrant domestic workers [77], the transgender community [45], or political activists [18]. Likewise, these methods have been used to study risks to people from a particular technology, such as smart speakers [43], mobile loan apps [57], and augmented reality glasses [26]. Still other papers use these methods to study particular experiences and protective practices surrounding those experiences, such as hate and harassment [93], removing data from people search websites [86], or IoT-enabled intimate partner abuse [81]. Finally, a variety of papers examine risks for populations in particular locations, such as South Asia [70], Bangladesh [3], or the Caribbean [94]. Because of the overlapping nature of security and privacy risks, papers often study problems at the intersection of these different types of context (e.g., at-risk users in a particular location or working with a particular technology).

Warford et al. [91] systematize at-risk user research, identifying contextual factors that have been studied for various at-risk populations, the types of protective practices these users adopt, and types of barriers that hinder adoption of these practices. Our work focuses on threat modeling as a practice among researchers, with a broader focus that includes other contexts in addition to at-risk users. Our framework includes risk factors, protective practices, and barriers, situating them within threat modeling practice alongside other components. Bellini et al. [12] systematize digital safety risks that occur in research that involves at-risk users, identifying practices researchers use to mitigate risk and developing guidance to researchers. We point to this paper in our work as a source of best practices for researchers who conduct threat modeling work with at-risk users. Thomas et al. [87] present a taxonomy for reasoning about hate and harassment, identifying classes of attacks and surveying users to identify their prevalence. This is a useful resource for researchers who are using threat modeling as an approach to understand how to help users stay safe from these kinds of attacks.

3. Methods

We conducted a broad, qualitative study of human-centered threat modeling literature to understand how researchers define threat modeling, what processes they use to identify or elicit a threat model from their subjects, and what language they use when describing their work.

3.1. Corpus Collection

We manually collected papers from SOUPS, CHI, CSCW, USENIX Security, IEEE Security & Privacy, and PETS, for the years 2018–2023. This set of conferences includes work from the security, privacy, and human-computer interaction (HCI) communities. We also examined NDSS and CCS, but they did not have relevant papers for this time period.

We used a manual method of reading individual papers because we found this to be much more effective than keyword searching in a database. The language used to describe human-centered threat modeling is highly varied, perhaps due to this being a nascent field, so using keywords may have excluded a significant portion of relevant literature. Manually reading and selecting papers enabled us to collect a comprehensive and contextually relevant selection of papers for our study. Moreover, focusing on top conferences meant we were more likely to use papers that represented the best practices in the field.

3.1.1. Phase I. During phase I, we systematically gathered the paper titles, abstracts, publication date, and the link to the publication from each conference's respective website, resulting in a comprehensive collection of 9200 papers. We only chose full-length papers, not works in progress, invited or keynote speeches, or poster presentations.

3.1.2. Phase II. In phase II, we refined this list by having two researchers meet and carefully read the titles and abstracts for each paper, determining their relevance to our research questions. Any disagreements were addressed through on-the-spot discussions, ensuring a consensus on the selections. At this step, we maintained a broad interpretation of relevance, considering any paper that even remotely aligned with the scope of our study, meaning any paper whose title and abstract mentioned user perspectives on security and privacy and used language that aligned with identifying or mitigating harms, threats, or concerns.

We narrowed our corpus to 165 papers at this step.

3.1.3. Phase III. To more tightly focus our corpus on threat modeling, we developed an **inclusion criteria** to identify papers whose purpose is to discover threats that users perceive and their strategies for mitigating those threats:

1) The research questions or goals state that one of the *primary purposes* of the paper is to study a human-centered security or privacy threat model. Because many papers don't explicitly use this terminology, we also include papers that state their purpose is to explore security and privacy concerns, awareness or knowledge of harms, threats, challenges, and/or risks. Among these papers authors often also identified strategies that users employ to avoid harm or mitigate threats. These are of interest to us but do not solely qualify a paper for inclusion. For example, we excluded [53] because it focuses on how people adapt privacy settings in the context of cognitive challenges, without identifying threats.
2) The paper examines *threats to people*, not to technologies such as devices or software. For example, we included [83] but excluded [13].
3) The paper studies threats as *perceived by people*. Some papers identify threats to users by watching their behavior (e.g. in a lab study) or identifying misconceptions (e.g. about app permissions [39]). These papers don't meet our criteria because they are not engaged in modeling modeling threats from the user's perspective, but instead are about identifying gaps in understanding or behaviors. While such papers are important when studying threats that people do not perceive and can complement the process of human-centered threat modeling, our framework is focused on eliciting threats as they are perceived by individuals. We similarly excluded papers that presented researcher-designed threat models. For such models to accurately reflect the community they are designed for, they would need to be based on prior research involving direct engagement with the community to understand their needs. Therefore, we focused on papers that directly involved working with the community and capturing their perceptions, rather than those deriving their threat models indirectly from other research.

If the wording in the research questions implied any of the concepts listed above but did not explicitly mention them or was unclear, we briefly looked at their contributions and findings to decide whether the paper met our criteria.

Because our goal was to broadly categorize human-centered threat modeling, we were generous in including papers in our corpus, to avoid leaving out any key concepts. Ultimately, we excluded a few papers as not relevant after we had read them fully during our analysis. We likewise excluded a few papers during our analysis that were purely quantitative, such as comparison studies. This was done because we found them less helpful for defining the threat modeling process, such as being more narrowly concerned with a single piece, like prioritizing threats.

Ultimately, we analyzed 78 papers. [4]

3.2. Analysis

During phase III, we marked papers that most closely aligned with our research questions. These were papers that explicitly mentioned threat modeling or presented findings extending beyond merely listing concerns or harms. We developed our initial codebook based on this set of 40 papers. This approach ensured that our initial codebook was rich and comprehensive.

To develop our codebook, two researchers collaboratively read and coded each paper, discussing codes and reaching agreement at all steps of the process. Our process was both iterative and inductive, looking for any information in each paper that was relevant to the threat modeling process. For each paper, we started by coding the research questions, if listed (n=43). For papers without explicit research questions, we coded their stated goals or contributions, since these enabled us to understand the main purpose of the study. We then coded the related work section to understand the context and background of their research, specifically searching for any frameworks or theories they used to see how these align with threat modeling. We then systematically examined and coded the methods of each paper in detail to understand exactly how the study was conducted. Where available (n=53), we looked for study instruments to see how the researchers elicited the threat models from their participants. We coded the findings or results of each paper to identify the components the authors chose to include in a threat model.

As we developed our codebook, we also diagrammed a flowchart for each paper to visualize the components of the threat model according to the researchers and to identify patterns, similarities, and differences in the approaches taken. We borrowed techniques from concept mapping, a method used to visually organize and represent knowledge, to enhance this process [37]. Concept mapping involves creating diagrams that show relationships between concepts, making it easier to identify overarching concepts and see how they are connected.

We also took numerous, detailed coding memos and had extensive discussions for every paper, which helped refine our codebook. As a result of these discussions, we grouped some codes into categories and occasionally broke them down into more granular, elaborate pieces. Our discussions facilitated continuous updates to the codebook throughout the process.

We dedicated a part of our codebook specifically to language used in the papers, due to the diversity of terms used in the field. This focused primarily on language to describe threats (concerns, risks, abuse) and protective practices (mitigations, behaviors, responses).

Once we had developed our codebook for the initial set of papers, one researcher read and coded the rest of the papers, an accepted practice [54], and similar to prior

4. Our full corpus and codebook is available at this link.

work [12]. This resulted in only a few more granular codes being added, without any new themes emerging. The two researchers who did the coding continued to meet and discuss themes across all papers, revising and updating them as they finalized the framework and guide.

3.3. Limitations

Our corpus is not intended to be a complete representation of all papers that have identified a threat model. We covered the most recent 6 years from a representative set of conferences, giving a broad look at current practices. Though our process omits some papers published in other years or venues, our methods resulted in a robust and representative framework and guide.

As with all qualitative work, our resulting framework and guide are based on coding that includes the interpretation and perspective of the researchers involved in the process. While the corpus represents the output of many experts in the field, future work could incorporate direct input from experts on how they conceive of the threat modeling exercise and their practices to prepare in advance, conduct the exercise, and analyze results.

Because identifying threats is the most basic practice of threat modeling, we did not include papers that examine only practices or solutions, such as papers that study protective behaviors exclusively. A different study could examine practices exclusively and develop a more comprehensive picture of that component, similarly to the work by Warford et al. [91] that systematizes protective practices of at-risk populations.

4. State of the Field

We first describe how researchers define threat modeling, then cover the methods they use to elicit a threat model.

4.1. Defining Human-Centered Threat Modeling

Our diagramming exercise helped us to identify three broad approaches to threat modeling in our corpus:

- *explicit*: the authors directly state that they are identifying a threat model [18], [86], [77], [45], [25]
- *implicit*: the paper essentially takes a threat modeling approach, for example identifying threats and protective practices, but the authors do not state that they are identifying a threat model [56], [43], [50], [22]
- *concerns only*: the paper is limited to exploring concerns or challenges [44], [100], [60], [97]

Only 22 of the papers in our corpus are in the first category, explicitly mentioning they are engaged in threat modeling. None of the papers that mention threat modeling qualified it as "human-centered" or discussed how it might be similar to or different from systems threat modeling. The vast majority of papers in our larger corpus don't use the term "threat modeling", but clearly are engaged in the same process of identifying threats/harms/risks and defensive practices.

TABLE 1. NUMBER OF PAPERS THAT INCLUDE OR DISCUSS A COMPONENT OF A THREAT MODEL

Component	Included	Discussed Separately
Threats/Harms/Risks	22	0
Adversaries	13	0
Defenses	5	11
Assets	1	0
Efficacy	1	1
Risk Factors	1	0
Goals	1	1

As a first pass to identify components of a threat model, we coded any references to threat modeling in these 22 papers. We also coded definitions of the term "threat model", which we found in 8 of these papers. These definitions are presented in Table 5 in the appendix. Table 1 lists the components in the threat models of these 22 papers, with counts of the number of papers that included a component in the threat model or discussed it as separate from the threat model.

The main takeaway from this exercise is that both explicit and implicit definitions of human-centered threat modeling are infrequent, inconsistent, and limited. While all papers included threats, harms, or risks in a threat modeling exercise, most considered defenses as related to but outside of the definition of a threat model. A majority of papers defining threat modeling considered adversaries, but there were also a variety of other factors (assets, efficacy of defenses, risk factors, and goals) that only a few papers included.

This motivates our discipline's need for a more comprehensive definition of and framework for human-centered threat modeling.

4.2. Eliciting a Threat Model

Most papers in our corpus use interviews to elicit threat models from participants, almost always using direct questions about threats and protective practices. For example, Samermit et al. interviewed online content creators [72], directly asking about concerns, protective practices, and advice or resources they would offer to others. Likewise, Steinbrink et al. asked asylum seekers in Germany about their privacy experiences during their migration [80]. Occasionally a paper's methods mentioned trying to avoid priming for security and privacy concerns by asking generally about technology and related concerns [14]. Interviews with children, in particular, use an indirect approach to understanding their threat models [99], as they may not have the words to describe security and privacy concerns [42].

A few papers use a mental modeling exercise as part of an interview; how a person thinks about a system can impact their perception of threats. For example, Meng [56]

used a drawing exercise and a free listing exercise to elicit mental models of smart speakers, with follow-up questions about privacy threats, benefits, or mitigation strategies if the participants brought them up.

Other methods used by a few papers each include focus groups and diary studies. A focus group provides an opportunity for participants to talk among themselves, thus potentially eliciting more natural expressions of a threat model. For example, Cobb et al. [15] conducted focus groups to discuss concerns about incidental users of devices in smart homes, followed by smaller group discussions to explore mitigations. A diary study enables participants to share context or in-the-moment interactions. For example, Lau et al. [43] had participants keep a diary about their interactions with smart speakers, providing context for a subsequent interview where they asked about their privacy perceptions, concerns, and behaviors.

We call particular attention to a paper by Słupska et al. [78], which used a participatory threat modeling workshop to understand lived experiences of domestic migrant workers related to security and privacy, identifying threats, harms, and sources of safety. Using this format enabled them to help participants from a marginalized community feel comfortable sharing their thoughts in open discussions.

Other methods used include a retrospective case study [11], an open-ended elicitation survey [52], and repertory grid analysis [7]. Several papers used mixed methods [58], [9], [64].

5. Threat Modeling Framework

To better understand the human-centered threat modeling process used by researchers, we used a qualitative analysis of our entire set of 78 papers, analysis steps described in Section 3.2. We looked broadly at all components included in the threat modeling process, irrespective of the language used to describe the process.

As a result of this analysis, we systematized a human-centered threat model process into the framework shown in Table 2. Our framework presents four interconnected components—context, threats, protective strategies, and reflection. We further break each component into a set of factors that provide additional detail on how researchers use this component to formulate a threat model. We provide definitions for each component and factor, along with examples of papers that illustrate this factor being used in a threat modeling exercise. For each component and factor, we sought to use general language that is representative of our corpus. Often, the language used to describe the components of a threat model is customized to fit the context of a paper. For example, Sambasivan et al., studying gender and digital abuse in South Asia, use the terms abuse, harms, and coping practices [71]. Lau et al., exploring smart speakers, use the terms privacy perceptions, concerns and privacy-seeking behaviors [43].

We also present a broad definition of human-centered threat modeling that is grounded in this data:

> *Human-centered threat modeling is a process to identify how people perceive and respond to risks to themselves. The process is centered on a person or community and grounded in understanding the context in which they operate. It includes the entire process of managing risks, including identifying them, responding to them, and reflecting on how existing behaviors and societal structures can be adapted to better manage risks in the future.*

In this section we describe each component and how it impacts the threat modeling process. Because context and reflection are less recognized as components of a human-centered threat modeling exercise, in these sections we highlight papers that exemplify how a given factor helped elicit a human-centered threat model. These summaries are in our own words, but styled similarly to quotes in a paper that covers qualitative data, providing support for how these factors fit in the framework.

5.1. Context

We define context as the circumstances, environment, and situational factors that play a role in shaping the threat model of an individual or a population. Contextualizing a threat model is crucial because it enables the researchers to gain a comprehensive understanding of the specific situations and constraints faced by participants. As a researcher studies context, they better understand what is unique about a person or community that intersects their threat model.

Context heavily influences the threat models that people perceive. This in turn helps researchers prioritize risks as perceived by the users, understand user behavior more holistically, respect cultural differences, adapt to evolving circumstances, and ultimately, better protect individuals and communities from harm. Researchers conducting threat modeling use context broadly, incorporating any facet of a person or their situation that could influence how they perceive threats or how they take protective measures. Context may be provided in a background section if it is well explored, but because it so heavily influences a threat model many papers uncover context in their findings.

While many papers situate a threat model within the context of a population or community, our results focus on those that include it in their findings, as opposed to those that cover context in a background or related work section. This was because in our corpus, when context was presented only as background, it often served to situate the study from the researchers' perspective, while context integrated into the findings reflected the perspectives and lived experiences of participants, which was crucial for accurately modeling the threats they perceived. For example, Daffalla et al. show how context is inextricably intertwined with a threat model when they study technology use by political activists during the sudanese revolution [18]. Other examples include (a) Redmiles explores the impact of cultural context on threat models for account security incidents [65], (b) Guberek et al. explore technology use by undocumented immigrants as part of understanding their threat models [28], (c) Lerner et

TABLE 2. THREAT MODELING FRAMEWORK

Component	Factor	Description	Examples
Context Circumstances, environment, and situational factors of an individual or population	Risk factors	Unique circumstances that augment or heighten an person's probability of being attacked and/or suffering disproportionate harm	[88], [51], [72], [8], [45]
	Unique challenges	The circumstances that influence a person's ability to effectively manage and respond to threats	[34], [76], [29], [45], [8]
	Culture	Common practices, beliefs, values, institutions, social norms, and daily activities of a population	[71], [5], [65], [7], [70]
	Personal identity and perception	Identity factors such as age, gender, location, as well as emotional states, self-perceptions, offline experiences, and personal definitions of safety, security, and privacy	[32], [8], [67], [6]
	Technology background and usage	Level of familiarity with technology, usage patterns, and mental models	[67], [90], [92], [83]
Threats A perceived event that could adversely impact safety	Concerns	Any perception a person has toward something that may negatively impact their security or privacy	[46], [70], [29], [7], [4], [63], [61], [83], [77]
	Harms	The negative consequences or impacts resulting from a threat, affecting safety, security, and privacy; may include physical, emotional, relational, or financial damage	[35], [70], [16], [77], [52]
	Actors	The individuals or entities involved in a threat scenario, including those who could cause harm and those who are the target, along with their motivations and capabilities	[16], [62], [92], [26]
Considerations	Trade-offs	The balancing decisions made between different security and privacy measures and other factors, such as convenience, cost, or usability	[21], [48], [38], [100]
	Threat appraisal	The perceived likelihood and severity of a threat	[94], [7], [8], [77]
Protective Strategies Measures taken in order to protect digital safety, privacy, and security	Recovery	Actions taken after an attack to restore normalcy or seek justice, including documentation, reporting, and any corrective measures	[35], [65], [22]
	Information seeking	Efforts to find out more about incidents and how to respond, including understanding sources of safety information; can be reactive or proactive	[65], [15], [25], [42]
	Defenses	Specific actions taken to protect against a person's perceived threats and/or to reduce harm; can be technical (technological solutions), behavioral (changes in personal behavior), and offline (physical actions)	[98], [70], [63], [83], [77], [46]
Considerations	Mapping	Aligning specific defenses with corresponding threats or harms	[38], [70], [7]
	Perceived cost	Resources required to implement protective measures, including time, effort, and financial expenses, as assessed by the individuals	[85], [45]
	Coping appraisal	Perceived ability to manage and respond to a threat, including the perceived efficacy of the protective strategy (response efficacy) and the confidence in one's capability to execute the strategy (self-efficacy)	[64], [50], [70], [52]
Reflection Assessment of protective strategies or broader societal changes needed to achieve security and privacy goals	Efficacy	The perceived overall effectiveness of protective strategies in addressing security and privacy concerns.	[22], [7]
	Barriers	Obstacles limiting or preventing the effective adoption of protective strategies, or reasons behind lack of protective actions, including conscious decisions that hinder the achievement of security and privacy goals	[48], [27], [44], [7]
	Advice	Guidance offered to others in similar situations, drawing on personal experiences	[77], [88], [86]
	Visioning	Desired actions or changes, both technological and societal, that could help enhance safety, security, and privacy	[26], [27], [77], [40], [10]

al. study how goals and values affect the threat models of members of the transgender community [45], and (d) Sun et al. explore how a variety of contextual factors influence parents as they consider smart home risks [84].

In our corpus, we identified five key factors, listed below, that both influence and situate participants' threat models. While this list is not exhaustive, it provides a basis for understanding what informs the threat model of participants. We also note that these factors are often intertwined, both among themselves and with other components of the threat model. Finally, researchers should not expect to explore every aspect of a person's context and should customize their threat modeling process to the people they are studying.

5.1.1. Risk Factors. Risk factors represent unique circumstances that augment or heighten an individual's probability of being digitally attacked and/or suffering disproportionate harm, consequently putting them at risk [72], [91]. Understanding these risk factors and their unique intersections helps discern the underlying causes of perceived threats among individuals. Some of these issues are resolvable, such as those related to knowledge gaps or limited access to resources. In such cases, it is imperative to understand these risk factors to better protect the population. Other

risk factors are inherent to the identity of a population and may not be easily solvable by technology. These factors are deeply ingrained and may be cultural, historical, or societal. In this case, it is important to surface these factors, so the solutions and interventions can be tailored to be more sensitive and responsive to the unique needs and challenges of the population. When studying at-risk populations, researchers should use the systematization of risk factors for at-risk populations by Warford et al., which categorized them as societal factors, relationship factors, and personal circumstances. [91].

Risk factors can be elicited through a participatory approach with people to understand their perceptions of the factors contributing to their vulnerability [72]. This method not only allows for the identification of risk factors but also facilitates the exploration of perceived barriers to safety, a topic discussed further in Section 5.4. Alternatively, some studies have implicitly extracted risk factors from data to discern the underlying factors posing risks to the population [88]. This approach may be helpful if participants can't readily identify the factors that put them at risk.

Privacy and activism in the transgender community [45]

Transgender activists can be vulnerable to a number of harms due to their identity, their visibility, and random bad luck. These risk factors influence the threats they experience and the protective actions they may (fail to) take. For example, the perceived risk factor of identity may lead an individual to believe they are unable to stop attacks because they are rooted in immutable characteristics of who they are.

5.1.2. Unique Challenges. It is important to understand the unique situations that are faced by the population being studied since their circumstances also help situate their threat models and can help specifically illustrate their barriers to safety. Sometimes these unique challenges overlap with their risk factors. These are cases where their challenges are exactly what is also putting them at more risk, such as in the case of transgender activists [45] and Muslim-American women [2]. But in some cases, there may be additional challenges that do not necessarily put them at more risk, but shape their privacy perceptions or practices, or put others at risk. For example, because IRB lack of regulations for digital data outside research settings, they may struggle to help protect research subjects [34]. Likewise, foster parents may struggle to regulate technology for foster teens if they do not respect their authority [8].

"Why wouldn't someone think of democracy as a target?" Security practices and challenges of people involved with U.S. political campaigns [16]

The work culture in a political campaign led to security not being prioritized. These factors include a focus on winning, transience, busy workers, tight budgets, amorphous boundaries, and lack of security knowledge.

5.1.3. Culture. Culture includes the common practices, beliefs, values, institutions, social norms, and daily activities of a population. Culture shapes a person's threat model, perceptions of capabilities as well as actual capabilities, and limitations. A nuanced understanding of these factors acts as a grounding point for safe technology designs. Culture becomes particularly significant when studying privacy, given its intricate, multifaceted, and culturally contingent nature [95]. For example, there is no specific term for the word "privacy" in the Dari, Pashtu, or Urdu languages [80], but those who lack a word for "privacy" may still have privacy perceptions and motivated behaviors [71]. Understanding these nuances is paramount, as privacy ideals diverge significantly across different cultural contexts, and social norms often define appropriate information flows [59]. Threat perceptions and practices can vary with a range of dimensions of culture, including language [80], values such as collectivism versus individualism [65], social norms, religious beliefs, historical context, and socioeconomic factors.

"Should I worry?" A cross-cultural examination of account security incident response [65]

Cultural context influences the threat model a person has regarding account compromise. People from a more collectivist culture worried more about someone close to them compromising an account, resulting in a feeling of violation from an account security incident.

5.1.4. Personal Identity and Perception. Understanding threat models requires a deep appreciation of how identity factors—such as age and gender—shape individuals' perceptions and experiences. These demographic details are often not reported in usable security and privacy and HCI research [30], [47], yet they are crucial for contextualizing a person's threat model. Additionally, factors like a person's emotional state [8], self-perceptions of (dis)abilities [32], and offline experiences [83] significantly impact their perceived threats and protective behaviors. These personal characteristics can lead to considerable variation in security and privacy practices.

Furthermore, participants' definitions of safety, security, and privacy play a pivotal role in shaping their threat models. Individuals describe threats based on their personal understanding of what constitutes safety, security, and privacy, which can vary widely. Asking questions about these definitions is essential, as it uncovers the differences in how people perceive and prioritize threats. Similarly, research often includes privacy perceptions about data collection, such as specific locations or social situations, who can use or access the data [67], address [60] and location data, and health data [26].

"They see you're a girl if you pick a pink robot with a skirt": A qualitative study of how children conceptualize data processing and digital privacy risks [83]

Children borrow perceptions of threats from the physical world and translate them into the digital world. In addition, age predicts a child's mental models of data processing, online risk perceptions, and self-protection behaviors.

5.1.5. Technology Background and Usage. We found several technology-related factors that inform threat models, including personal experiences with technology, knowledge of security practices, and self-perceptions of their technological capabilities. Additionally, understanding what individuals are trying to do with a particular technology is also important, as different activities can influence their perception of threats. These experiences determine people's comfort level with technology, their ability to identify potential threats, and their overall approach to security. Similarly, understanding a person's security and privacy needs or goals provides important context for what they want to achieve. These goals often include desires for specific protections, such as having control over their personal information [51], [43], ensuring data confidentiality [18], avoiding judgement [2], and maintaining anonymity [90].

A common area of research is to understand a person's mental model (or internal representation) of a process or system. Mental models influence a person's understanding of the capabilities and limitations of technology and their perception of potential threats [65].

Amazon vs. my brother: How users of shared smart speakers perceive and cope with privacy risks [33]

Participants with more sophisticated mental models of shared smart speakers were able to take specific actions to counter concerns, such as avoiding linking private information to the device. In contrast, participants with less sophisticated models were worried about risks that could be easily prevented.

5.2. Threats

Threats are the most prevalent and central component of a threat modeling exercise. We define threats as any perceived circumstance, event, or danger that could potentially adversely impact safety, including the consequences of the event, the actors involved, and their motivations.

5.2.1. Concerns. We refer to concerns as any perception an individual has toward something that may negatively impact their security or privacy. Under this umbrella we include both potential and experienced adverse events. The most predominant language used in our corpus is "threats", "concerns", and "risks", but we use concerns to cover the wide range of language across our corpus, as shown in Table 3. We do not advocate for uniform terminology; rather, each researcher should use language appropriate to the people they are studying.

TABLE 3. LANGUAGE USED TO DESCRIBE THREATS

abuse, abuse attacks, abuse risks, adversary, attack, attacker, attitudes, concerns, consequences, digital safety concerns, digital safety experiences, digital safety risks, digital safety threats, harms, impact of the harm, issues, negative experiences, perceived risks, perceived safety, privacy and security concerns, privacy challenges, privacy concerns, privacy risks, risk perceptions, risks, safety concerns, security challenges, security concerns, security incidents, security risks, threat, worries.

Notably, participants often do not differentiate between online and offline threats, sometimes employing online mitigations for real-life threats and vice versa, leading them to merge the two together. Studying both types of threats for a population provides a more extensive view.

In our corpus, we found that participants describe threats in three distinct ways based on their proximity to personal experience: purely speculative concerns—hypothetical scenarios they fear might happen [26], [40]; second-hand concerns—stories they had heard from others or read on the internet [3], [96], [84]; and first-hand personal experiences—threats that they had directly encountered [70], [88], [27]. Speculative concerns were more prominent with privacy risks, due to participants lacking knowledge or transparency of how data is collected and used.

5.2.2. Harms. We define harms as the negative consequences that materialize from perceived or experienced risks. In our corpus, we observed that some papers use the terms harms and threats interchangeably. However, we argue that these are separate concepts and that researchers can add depth to their findings by explicitly exploring harms if they not already mentioned by participants. For example, McDonald et al. [52] identify specific consequences resulting from negative experiences with phone number recycling. Alternatively, Samermit et al. centered harms faced by content creators, rather than connecting them to specific risks, due to the wide number of factors that can influence perception of harm [72].

As delineated by Scheuerman et al. [73], harms can be categorized into four types: physical, emotional, relational, and financial. This nuanced categorization enables a more granular analysis of the impacts threats can have on the individuals. Similarly, these harms can be analyzed through the dimension of direct harms, which encompass immediate negative impacts, and indirect harms, which include long-term or secondary consequences as well as collateral damage and harms to others.

5.2.3. Actors. Actors include both the attacker (or adversary) and the target of the attack. This involves understanding a person's perceptions of who might cause harm and who would be harmed. Attackers can be categorized into known and unknown individuals. Known attackers might include acquaintances [63], significant others [62], or bystanders [62], while unknown attackers could be strangers

TABLE 4. LANGUAGE USED TO DESCRIBE PROTECTIVE STRATEGIES

precautions, privacy considerations, privacy management, privacy practices, protective behaviors approaches, behavior changes, coping practices, coping strategies, defenses, mediating risks, mitigation strategies, mitigations, perceived security risk mitigations, privacy strategies, privacy-preserving strategies, privacy-seeking behavior, protection mechanisms, protective measures, protective behaviors, protective practices, safety strategies, security and privacy strategies, security incident response, strategies, technical defenses, responses, risk coping, risk management, safety considerations, safety practice, sources of safety

[2], anonymous hackers [1], or entities with no direct personal connection to the target [92]. Attacker motivations can also vary, ranging from malicious, where the actor is aware of their actions causing intentional harm [1], [62], to benign, where the intent might not be to harm but still results in unintended negative consequences [50].

In the context of privacy research, attackers are often conceptualized as "data collectors" and "data receivers" [26]. Given the contextual nature of privacy, a person's insights into about both of these entities can help clarify their norms about data privacy.

5.2.4. Considerations. When evaluating the harms in human-centered threat modeling, several considerations must be taken into account.

Trade-offs. Particularly when discussing privacy concerns, researchers consider trade-offs arising from the use of technology. For example, users of smart devices trade-off privacy for cost or convenience [43], [100], and explicitly recognize both benefits and harms to the technology [15]. Users of mobile loan apps in Kenya likewise perceive trade-offs that affect their perception of threats [57].

Threat appraisal. Threat appraisal includes understanding how people perceive the likelihood and severity of a particular threat [94]. This concept is used in protection motivation theory [66] to understand what motivates people to adopt new behaviors or change their attitudes. The perceived likelihood and severity of harms determine how participants prioritize these harms, allocate resources, and shape their management strategies. Additionally, understanding these factors can help uncover seemingly paradoxical decisions, such as perceiving the consequences of a harm to be extremely dangerous but still not taking action to mitigate it.

5.3. Protective Strategies

We define protective strategies as the comprehensive measures participants take in order to protect their digital safety, privacy, and security. As with threats, people often do not discriminate between offline and online worlds, meaning they'll often have offline mitigations for an online threat or vice versa. Table 4 shows the varied language used across our corpus to describe protective strategies.

5.3.1. Recovery. Recovery involves decisive actions taken after an attack to restore normalcy and seek justice. These actions are often corrective and reactive, involving developing an understanding of the immediate harm and mitigating its impact. Depending on the incident, recovery actions can range from seeking legal recourse [81] to engaging in community support [77].

5.3.2. Information seeking. Information seeking involves finding out more about an incident and how to respond to it. It also encompasses understanding where participants may obtain their safety advice and information. Participants may search for resources reactively after an incident [65] or proactively before an incident to prevent it [15]. By understanding the sources they rely on, researchers can gauge the accuracy and reliability of the information participants use to make decisions about their digital safety. Knowing where people turn for help also offers opportunities for experts to provide resources in those avenues.

Participants may turn to a variety of sources for information, including general online searches, specific online platforms and forums [72], news media [25], official websites, family and friends [32], and in some cases, advocates and activists [27]. We found that information-seeking is highly context-dependent, varying based on who is seeking advice and the specific issue at hand. For example, children often consulted parents [42], content creators turned to supporters or allies for help [88], and teachers referred to various resources collected over time and general online searchers [50].

5.3.3. Defenses. Defenses are the specific actions participants take to protect against their perceived threats and/or to reduce harm. While participants often do not differentiate between offline and online defenses, researchers sometimes categorize them separately when presenting their results, so we follow this approach by categorizing these practices into technical, behavioral, and offline defenses.

Technical defenses. Technical defenses involve implementing technological solutions to enhance digital safety and privacy. Examples include using alternative identifications [19], blurring profile photos and adjusting privacy settings on social media [7], using firewalls or VPNs [85], and changing passcodes frequently [62].

Behavioral defenses. Behavioral defenses refer to changes in personal behavior when individuals interact with technology, to improve safety. Examples include self-censorship [85], [2], only using apps with self-delete timers on photos [7], using trusted devices only [25], and refraining from using certain systems altogether [62], [56].

Offline defenses. Offline defenses encompass physical actions and real-world measures taken to protect against digital threats. These actions occur in the physical world to mitigate risks that have digital implications. Examples include putting tape over the front camera [7], changing physical locations [62], and reporting incidents to relevant authorities or platforms [22].

5.3.4. Considerations. When studying protective strategies, several considerations must be taken into account.

Mapping. Mapping involves identifying which defense is a response to which threat, rather than separately listing perceived threats and defenses. Without mapping, the relationship between threats and defenses can appear ambiguous or even contradictory. For example, Sambasivan et al. [70], in a study of gender and digital abuse in South Asia, provide a mapping that links types of abuse to mechanisms, harms, and coping practices. This enables subsequent researchers or supporters to identify how they can target resources to help with particular forms of abuse.

Perceived cost. People typically consider the cost of implementing protective measures along such dimensions as time, effort, and financial costs. This understanding provides insights into their decision-making processes and why they may choose to adopt or not adopt certain defenses [85].

Coping appraisal. Coping appraisal, drawn from protection motivation theory [66], includes response efficacy, the belief that a protective strategy will effectively mitigate the threat, and self-efficacy, an individual's confidence in their ability to execute the protective measure. For instance, Guberek et al. found in their study of undocumented immigrants that participants often experienced a sense of resignation, believing that government authorities already possessed extensive information about them irrespective of their technological choices [28]. This perception of low response efficacy resulted in a diminished likelihood of adopting conscious technical defenses. Likewise, in Ray's study of working age adults, they expressed similar sentiment about third-parties and companies selling their private information [64].

5.4. Reflection

We define reflection as the component of threat modeling where participants assess their protective strategies or broader societal changes that are needed to obtain security and privacy goals. Incorporating this perspective from participants helps researchers understand what works (or doesn't work) well as people try to protect themselves, as well as how society can help. Reflection is focused on understanding practices (both personal and societal), whereas context is focused on understanding the person.

We note that it is common for the papers in our corpus to conclude with a set of recommendations that focus on ways to better mitigate harms for the subjects they studied. These recommendations are typically grounded in data but coming from the perspective of the researchers. Here we focus on papers that work directly with participants to identify recommendations for their situation.

5.4.1. Efficacy. Often participants naturally share perspectives on the effectiveness of their practices when researchers ask about what they do to protect themselves. For example, when discussing whether to report digital safety experiences, youth and parents were skeptical that platforms took these reports seriously, and were concerned about notifying schools for fear that this could lead to negative consequences [22]. Likewise, young adults in Pakistan shared opinions about which of their current practices were effective in meeting their online concerns about issues such as cyberstalking and fake profiles [7]. This could also include asking people to prioritize among threats to identify those they consider most pressing.

5.4.2. Barriers. Researchers may surface barriers to safety and reasons why participants fail to take any protective actions. Barriers may include lack of lack of accessible resources, cognitive burdens, ineffectiveness of platforms, negative reactions from family members, poor usability, or lack of guidance. Reasons for not taking a protective action are closely related, including issues such as lack of trust, low perceived likelihood of success, low self-efficacy, low response efficacy, procrastination, lack of knowledge, high cost, marginal risk, resignation, or acceptance of trade-offs. Kapoor et al. frame this as a set of decisions that labor organizers make as they consider privacy practices [38].

"It's the equivalent of feeling like you're in jail": Lessons from firsthand and secondhand accounts of IoT-enabled intimate partner abuse [81]
Victim-survivors report a large number of barriers, which the authors link to identifying abuse, mitigating abuse, providing actionable advice, and taking legal action. Having barriers directly linked to context and to specific needs or goals helps clarify unique challenges victim-survivors face and where they need additional help.

5.4.3. Advice. Researchers may ask participants to share advice they would offer to others in their same situation. Since researchers are not typically in the same situation as the subjects of their study, this line of questioning enables the researcher to better ground recommendations in participant experiences of what is likely to work well or what needs improvement.

"It's common and a part of being a content creator": Understanding how creators experience and cope with hate and harassment online [88]
Creators gave advice to other creators who are dealing with hate and harassment, listing a variety of items (don't engage, moderation is crucial) that provide depth and insight into creator experiences. These perspectives led directly to concrete recommendations from the authors to improve tools for creators.

5.4.4. Visioning. Researchers may conduct a visioning exercise, where researchers ask participants about what they would like to see done to help keep them safe. This likewise helps ground researcher recommendations in participant experiences and perspectives. For example, this can include improvements from a social media platform to make it easier to gather evidence of harassment [27], data privacy

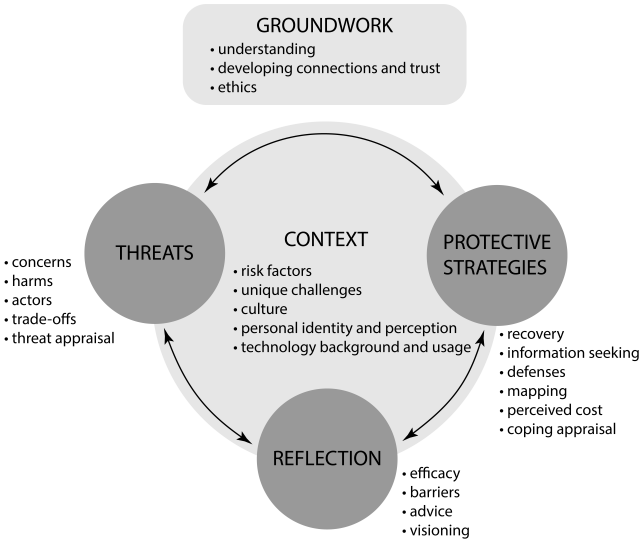

Figure 1. Visualization of the threat modeling framework

regulations in countries that lack them [57], privacy recommendations for smart home devices that collect data on incidental users [15], legal reforms to protect migrant domestic workers [78], or desired safety features for social virtual reality [19].

6. A Guide to Using the Framework

We believe our framework can be used to guide research in a variety of ways. By providing a structured approach to human-centered threat modeling, our framework can help researchers: (a) **identify new research directions** by helping formulate research questions, focusing on un- or under-explored areas within specific communities; (b) **prepare for a study** by guiding the development of interview guides, focus group materials, or participatory threat modeling exercises, helping incorporate all components from the framework to ensure comprehensive understanding of threat models; (c) **conduct more thorough threat modeling** ensuring that the full spectrum of user experiences, threats, and protective strategies is covered, allowing researchers to address relevant aspects that may have been previously unstudied or overlooked; and (d) **compare literature and identify gaps**, enabling the identification of unanswered questions and incomplete areas in human-centered threat modeling

We also provide a guide for human-centered threat modeling based on best practices we have learned from the corpus.

6.1. Basic Principles

Figure 1 visualizes how the different components of the threat model interact and illustrates some basic principles of the threat modeling process.

Start with groundwork. While groundwork is not directly a part of the threat modeling exercise, it is a crucial step and should precede any direct interactions with the participants. Groundwork can involve learning about the population, including studying prior research on the population, developing trust with the population [32], understanding how harm to the population can be minimized, building connections within the community, and ensuring ethical and culturally aware research practices. When working with at-risk populations, consult guidelines for safer research with at-risk populations [12].

Understand the centrality of context. Context influences all aspects of the threat modeling process, including the identification of threats, the practices adopted to mitigate them, and the barriers encountered. The threat models of individuals are dependent on and valid in that context. Context is discovered from both preliminary groundwork and from direct interactions with the participants during the threat modeling exercise.

Threat modeling is a broad process of discovery. Although some papers consider protective practices and coping strategies to be separate from the threat model, they are inextricably tied to it. Protective practices directly influence and are influenced by the perceived threats, making it essential to study them in conjunction with threats to understand the full spectrum of participants' responses. Similarly, context and reflection are integral to the process as they provide the necessary background and actionable insights that make threat modeling practical and effective. Therefore, threat modeling can be a broad discovery process, seeking a comprehensive understanding of the participants' security and privacy landscape to inform more effective interventions.

Threat modeling is a non-linear and connected process. Researchers should not expect to proceed from groundwork to context to threats, then protective strategies and reflection. These are not discrete steps in a linear process, and the components may not neatly fit into boxes. Participants often jump between discussing the different components such as threats and protective practices, since the way they experience threats is fluid. We have included double-sided arrows to illustrate the flexibility that researchers should use. Researchers should seek to develop an integrated model of a person. Lerner et al. [45], for example, present *risk models* of participants that incorporate their perceived risk factors and illustrate how those perceptions influence the threats they perceive and the strategies they use to mitigate those threats. This holistic approach ensures a broad discovery process that goes beyond separately listing individual components. Such connections across components allow for broader implications and deeper understanding of how their threat models affect their safety behavior.

6.2. Considerations

In reading the variety of approaches to threat modeling in our corpus, we developed the following insights for researchers.

Customize the framework. The framework we present is not intended to be a one-size-fits-all or prescriptive approach. The scope of work in this field is vast, spanning topics from privacy concerns with shared smart devices to mitigating harms for victim-survivors of interpersonal violence. Each person or community will have different goals, needs, perspectives, experiences, and practices, and thus different threat models. In fact, a given population may hold *multiple* threat models, and researchers should avoid shoehorning people into the same model for convenience. Likewise, there is no set process that each researcher should follow; rather they should work as partners in a community to identify an approach that works best for their situation. Researchers should start with groundwork, and that can help guide further customization. There are a wide variety of approaches in the literature to draw inspiration from.

Consider offline threats and non-technical solutions. A number of the papers in our corpus cover both online and offline threats. In a variety of contexts, these threats are intertwined, and people would not naturally distinguish between them when discussing their situation [28], [94] Researchers should be careful to avoid focusing on only online concerns where this may be manifest. Likewise, mitigations of threats may be non-technical in nature, so researchers should avoid thinking only in terms of technical solutions.

Work toward a complete threat model. Our framework can help researchers identify less explored components—in their own research or in the literature—for a particular population. For any given paper, researchers can use the framework to identify an important aspect of a threat model that they may have left unexplored, for example whether their participants prioritize responding to a particular threat or have any barriers that inhibit safety. For a larger research project, researchers can use the framework to guide multiple studies that examine different aspects of a threat model, working systematically to develop a broader picture.

Use the framework at different stages of the research process. Our framework is primarily meant to guide a researcher during the planning process, helping them check whether they have considered each component, and various factors within each component, as they develop an interview guide or other methods. We have linked the framework to exemplary papers so that researchers can identify other work that considers similar threat modeling questions. Researchers could also use the framework when analyzing data, to see how different aspects of a threat model may be present in qualitative data they have collected. In some cases, researchers may not be directly studying a threat model, but aspects of a threat model may emerge during research that surfaces risks, threats, harms, or protective behaviors. In these cases, the framework can help guide researchers toward subsequent research that can explore these themes in more depth.

6.3. Potential Gaps

In reviewing our corpus, we identified several gaps that future work in human-centered threat modeling should address. These include:

- More thoroughly reporting on groundwork. This is a crucial step in the threat modeling process, yet many papers don't report on steps taken in this phase of the research.
- Considering other factors of context, such as socioeconomic status, religion, or urban/suburban/rural environments. There are likely many areas of context that are unexplored.
- Considering the breadth of threat modeling to go beyond identifying harms. Many papers stopped at eliciting concerns or harms, without considering other aspects of a threat model. It's possible that this appears in subsequent work that is not in our corpus, but it may also be an area ripe for exploration in many of the problems studied.
- Incorporating reflection. Most papers use researcher-derived recommendations, without engaging the people they study in identifying recommendations or advice or co-designing interventions. Engaging with participants in seeking or designing solutions is a form of respect and may result in greater success at preventing or mitigating harm.
- Incorporating a broader set of methods to elicit threat models. Most papers use interviews, which could mean that there are elements of threat models that are not well explored.

7. What Translates from Systems Threat Modeling?

At first glance, there are some parallels between Shostack's four questions for systems threat modeling and the four components of our human-centered threat modeling framework. Both involve examining threats and mitigations, both involve some kind of reflection on the process, and both are set in a context (a system or people). However, human-centered threat modeling has some important distinctions that researchers engaged in this process should be aware of.

First, with systems threat modeling, the type of system a team is building or maintaining certainly influences the threat modeling process. But with human-centered threat modeling, the human context is highly varied and touches every aspect of threat modeling with people in myriad ways. People may have risk factors that make them susceptible to certain threats, or unique challenges they face. People have culture and customs that influence the types of threats they perceive and the types of actions that are available to them. People have goals, values, emotions, and perspectives that must be considered to understand their threats and behaviors.

Second, threats to systems and corresponding mitigating practices are fairly well known and generalize-able; there are formal exercises, card games, frameworks, and a variety of tools based on identifying well-known threats. In contrast, the landscape of threats to humans is vast—with harms that span physical, emotional, relational, and financial—and specific to particular contexts. There are still threats to a variety of people and contexts that may not be well studied. Moreover, people may not be able to eliminate or significantly mitigate some threats. In fact, technology may make some threats worse! Emotions play a role in how someone protects themselves; people may feel trapped or resigned, without the ability to take a protective action.

Third, with systems threat modeling, reflection is primarily a matter of identifying how well the threat modeling process has worked. With human-centered threat-modeling, reflection is much broader, assessing not just the effectiveness of protective practices, but also barriers, advice, and how technology and society itself may need to be changed to address a threat.

Finally, with systems threat modeling, those conducting the exercise are usually experts in knowing how the system works, with the focus on identifying and mitigating threats. With human-centered threat modeling, the people at the center of the exercise are the experts in their particular context. The researcher works as a partner with people, focusing on understanding them and their situation. Eliciting threats and protective practices are core parts of the researcher's process, but ultimately the goal is to learn from and to help people.

8. Centering Human Safety

Researcher engagement with the people they study raises a host of ethical questions. Do researchers have an obligation to help find solutions to threats they surface, or provide advice? Should researchers correct misconceptions? Should researchers tell participants about their unrecognized threats? What if telling them causes more anxiety or worry, or there is no good mitigation strategy, or there are barriers to them enacting a particular practice? What if societal changes are needed? We point readers to a recent paper that discusses ethical frameworks that can help analyze these kinds of questions [41].

Here, we wish to highlight a few papers where researchers are deeply engaged with the community they are studying, as examples of how our field can be ethically engaged in the practice of computer science. Hayes et al. describe attending monthly gatherings and volunteering with an organization of people with visual impairments, citing this as an effective way to build trust with a community prior to running a study [32]. Slupska et al. describe creating a digital privacy and security guide for domestic migrant workers that they distributed for free to interested organizations and the public [78], a result of the participatory design workshops they held. Researchers at Cornell Tech have established a clinical approach to computer security [31], founding a consultation service to help victim-survivors of intimate partner violence. This has led to a growing body of research [11], [23], [89] that is centered on safety as a basic human need [82].

A primary takeaway from our reading is that researchers should *not* be using threat modeling as a means to exploit a group of people for another paper in our field. Researchers should seek to partner with the people they are studying, for example by working with an organization that is already engaged in the field. They should likewise partner with other researchers who specialize in complementary areas, such as sociology or social work, to ensure that they are following best practices. Partnerships are essential to developing the understanding needed to conduct the research and to centering human safety. Our hope is that through the use of our framework and guide, researchers can conduct meaningful and impactful threat modeling studies that not only advance academic knowledge but also contribute to the safety and well-being of the communities they study.

9. Conclusion

Human-centered threat modeling in security and privacy research is underdefined, often borrowing from software threat modeling, which is inherently designed for systems rather than the complex nature of human behavior. This results in fragmented and incomplete understandings of user threat models, hindering their safety. To address this, we analyzed 78 papers and developed a framework and guide for human-centered threat modeling. Our framework includes four essential components: context, threats, protective strategies, and reflection, and highlights their complex interplay and the non-linear nature of threat modeling. We also developed a guide, based on best practices from the papers in our corpus, and differentiated human-centered threat modeling from systems threat modeling. By using our framework and guide, researchers can better understand a person's multifaceted experiences, identify barriers to effective threat mitigation, and propose more tailored interventions. Our hope is that these help the field continue its move toward centering human safety.

Acknowledgements

We thank Elissa Redmiles for inspiring discussions that led to this research, Noel Warford for his help with research methods, and Saba Iqbal for assisting with the final paper counts. We also thank the anonymous reviewers and our shepherd for their valuable feedback and guidance.

References

[1] N. Abdi, K. M. Ramokapane, and J. M. Such. More than smart speakers: Security and privacy perceptions of smart home personal assistants. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, pages 451–466, Santa Clara, CA, Aug. 2019. USENIX Association.

[2] T. Afnan, Y. Zou, M. Mustafa, M. Naseem, and F. Schaub. Aunties, strangers, and the {FBI}: Online privacy concerns and experiences of {Muslim-American} women. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, pages 387–406, 2022.

[3] S. I. Ahmed, M. R. Haque, J. Chen, and N. Dell. Digital privacy challenges with shared mobile phone use in bangladesh. *Proceedings of the ACM on Human-computer Interaction*, 1(CSCW):1–20, 2017.

[4] T. Akter, B. Dosono, T. Ahmed, A. Kapadia, and B. Semaan. "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 20)*, pages 1929–1948, 2020.

[5] M. N. Al-Ameen, H. Kocabas, S. Nandy, and T. Tamanna. "we, three brothers have always known everything of each other": A cross-cultural study of sharing digital devices and online accounts. *Proceedings on Privacy Enhancing Technologies*, 2021.

[6] W. S. Albayaydh and I. Flechais. Exploring bystanders' privacy concerns with smart homes in jordan. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–24, 2022.

[7] A. Ashraf, C. J. König, M. Javed, M. Mustafa, et al. "stalking is immoral but not illegal": Understanding security, cyber crimes and threats in pakistan. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*, pages 37–56, 2023.

[8] K. Badillo-Urquiola, X. Page, and P. J. Wisniewski. Risk vs. restriction: The tension between providing a sense of normalcy and keeping foster teens safe online. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

[9] D. G. Balash, D. Kim, D. Shaibekova, R. A. Fainchtein, M. Sherr, and A. J. Aviv. Examining the examiners: Students' privacy and security perceptions of online proctoring services. In *Seventeenth symposium on usable privacy and security (SOUPS 2021)*, pages 633–652, 2021.

[10] D. G. Balash, X. Wu, M. Grant, I. Reyes, and A. J. Aviv. Security and privacy perceptions of Third-Party application access for google accounts. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 3397–3414, Boston, MA, Aug. 2022. USENIX Association.

[11] R. Bellini. Paying the price: When intimate partners use technology for financial harm. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2023.

[12] R. Bellini, E. Tseng, N. Warford, A. Daffalla, T. Matthews, S. Consolvo, J. P. Woelfer, P. G. Kelley, M. L. Mazurek, D. Cuomo, et al. Sok: Safer digital-safety research involving at-risk users. *IEEE Security & Privacy*, 2024.

[13] L. Bieringer, K. Grosse, M. Backes, B. Biggio, and K. Krombholz. Industrial practitioners' mental models of adversarial machine learning. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, pages 97–116, 2022.

[14] C. Chen, N. Dell, and F. Roesner. Computer security and privacy in the interactions between victim service providers and human trafficking survivors. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 89–104, 2019.

[15] C. Cobb, S. Bhagavatula, K. A. Garrett, A. Hoffman, V. Rao, and L. Bauer. "i would have to evaluate their objections": Privacy tensions between smart home device owners and incidental users. *Proceedings on Privacy Enhancing Technologies*, 2021.

[16] S. Consolvo, P. G. Kelley, T. Matthews, K. Thomas, L. Dunn, and E. Bursztein. "why wouldn't someone think of democracy as a target?": Security practices & challenges of people involved with {US}. political campaigns. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 1181–1198, 2021.

[17] M. Corporation. The STRIDE threat model. *Tech. rep., Microsoft Corporation*.

[18] A. Daffalla, L. Simko, T. Kohno, and A. G. Bardas. Defensive technology use by political activists during the sudanese revolution. In *2021 IEEE symposium on security and privacy (SP)*, pages 372–390. IEEE, 2021.

[19] E. Deldari, D. Freed, J. Poveda, and Y. Yao. An investigation of teenager experiences in social virtual reality from teenagers', parents', and bystanders' perspectives. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*, pages 1–17, 2023.

[20] M. Deng. Privacy preserving content protection. *Ph. D. Dissertation*, 2010.

[21] P. Emami-Naeini, J. Breda, W. Dai, T. Kohno, K. Laine, S. Patel, and F. Roesner. Understanding people's concerns and attitudes toward smart cities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–24, 2023.

[22] D. Freed, N. N. Bazarova, S. Consolvo, E. J. Han, P. G. Kelley, K. Thomas, and D. Cosley. Understanding digital-safety experiences of youth in the us. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2023.

[23] D. Freed, S. Havron, E. Tseng, A. Gallardo, R. Chatterjee, T. Ristenpart, and N. Dell. "is my phone hacked?" analyzing clinical computer security interventions with survivors of intimate partner violence. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–24, 2019.

[24] J. Freund and J. Jones. *Measuring and managing information risk: a FAIR approach*. Butterworth-Heinemann, 2014.

[25] A. Frik, L. Nurgalieva, J. Bernd, J. Lee, F. Schaub, and S. Egelman. Privacy and security threat models and mitigation strategies of older adults. In *Fifteenth symposium on usable privacy and security (SOUPS 2019)*, pages 21–40, 2019.

[26] A. Gallardo, C. Choy, J. Juneja, E. Bozkir, C. Cobb, L. Bauer, and L. Cranor. Speculative privacy concerns about ar glasses data collection. *Proceedings on Privacy Enhancing Technologies*, 2023.

[27] N. Goyal, L. Park, and L. Vasserman. "you have to prove the threat is real": Understanding the needs of female journalists and activists to document and report online harassment. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2022.

[28] T. Guberek, A. McDonald, S. Simioni, A. H. Mhaidli, K. Toyama, and F. Schaub. Keeping a low profile? technology, risk and privacy among undocumented immigrants. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–15, 2018.

[29] V. Hamilton, H. Barakat, and E. M. Redmiles. Risk, resilience and reward: Impacts of shifting to digital sex work. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–37, 2022.

[30] A. A. Hasegawa, D. Inoue, and M. Akiyama. How weird is usable privacy and security research?

[31] S. Havron, D. Freed, R. Chatterjee, D. McCoy, N. Dell, and T. Ristenpart. Clinical computer security for victims of intimate partner violence. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 105–122, 2019.

[32] J. Hayes, S. Kaushik, C. E. Price, and Y. Wang. Cooperative privacy and security: Learning from people with visual impairments and their allies. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, pages 1–20, 2019.

[33] Y. Huang, B. Obada-Obieh, and K. Beznosov. 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*, pages 1–13, 2020.

[34] J. Huh-Yoo and E. Rader. It's the wild, wild west: Lessons learned from irb members' risk perceptions toward digital research data. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1):1–22, 2020.

[35] J. Im, S. Schoenebeck, M. Iriarte, G. Grill, D. Wilkinson, A. Batool, R. Alharbi, A. Funwie, T. Gankhuu, E. Gilbert, et al. Women's perspectives on harm and justice after online harassment. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–23, 2022.

[36] N. JTFTI. Nist sp 800-39 managing information security risk: Organization, mission, and information system view. Technical report, Accessed 1/16/2020 from https://csrc. nist. gov/publications/detail/sp/800 . . . , 2011.

[37] M. Kane and W. M. Trochim. *Concept mapping for planning and evaluation.* Sage Publications, Inc, 2007.

[38] S. Kapoor, M. Sun, M. Wang, K. Jazwinska, and E. A. Watkins. Weaving privacy and power: On the privacy practices of labor organizers in the us technology industry. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–33, 2022.

[39] A. Kariryaa, G.-L. Savino, C. Stellmacher, and J. Schöning. Understanding users' knowledge about the privacy and security of browser extensions. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, pages 99–118. USENIX Association, Aug. 2021.

[40] P. G. Kelley, C. Cornejo, L. Hayes, E. S. Jin, A. Sedley, K. Thomas, Y. Yang, and A. Woodruff. "there will be less privacy, of course": How and why people in 10 countries expect {AI} will affect privacy in the future. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*, pages 579–603, 2023.

[41] T. Kohno, Y. Acar, and W. Loh. Ethical frameworks and computer security trolley problems: Foundations for conversations. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 5145–5162, Anaheim, CA, Aug. 2023. USENIX Association.

[42] P. Kumar, S. M. Naik, U. R. Devkar, M. Chetty, T. L. Clegg, and J. Vitak. 'no telling passcodes out because they're private' understanding children's mental models of privacy and security online. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):1–21, 2017.

[43] J. Lau, B. Zimmerman, and F. Schaub. Alexa, are you listening? privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on human-computer interaction*, 2(CSCW):1–31, 2018.

[44] K. Lebeck, K. Ruth, T. Kohno, and F. Roesner. Towards security and privacy for multi-user augmented reality: Foundations with end users. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 392–408. IEEE, 2018.

[45] A. Lerner, H. Y. He, A. Kawakami, S. C. Zeamer, and R. Hoyle. Privacy and activism in the transgender community. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[46] Y. Li, Y. Kou, J. S. Lee, and A. Kobsa. Tell me before you stream me: Managing information disclosure in video game live streaming. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–18, 2018.

[47] S. Linxen, C. Sturm, F. Brühlmann, V. Cassau, K. Opwis, and K. Reinecke. How weird is chi? In *Proceedings of the 2021 chi conference on human factors in computing systems*, pages 1–14, 2021.

[48] A. F. Luo, N. Warford, S. Dooley, R. Greenstadt, M. L. Mazurek, and N. McDonald. { How} library {IT} staff navigate privacy and security challenges and responsibilities. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 5647–5664, 2023.

[49] N. Malkin. Contextual integrity, explained: A more usable privacy definition. *IEEE Security & Privacy*, 21(1):58–65, 2022.

[50] S. Maqsood and S. Chiasson. "they think it's totally fine to talk to somebody on the internet they don't know": Teachers' perceptions and mitigation strategies of tweens' online risks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2021.

[51] A. McDonald, C. Barwulor, M. L. Mazurek, F. Schaub, and E. M. Redmiles. "it's stressful having all these phones": Investigating sex workers' safety goals, risks, and practices online. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 375–392, 2021.

[52] A. McDonald, C. Sugatan, T. Guberek, and F. Schaub. The annoying, the disturbing, and the weird: Challenges with phone numbers as identifiers and phone number recycling. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[53] N. McDonald, A. Larsen, A. Battisti, G. Madjaroff, A. Massey, and H. Mentis. Realizing choice: Online safeguards for couples adapting to cognitive challenges. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, pages 99–110, 2020.

[54] N. McDonald, S. Schoenebeck, and A. Forte. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for cscw and hci practice. *Proceedings of the ACM on human-computer interaction*, 3(CSCW):1–23, 2019.

[55] R. McNaney, C. Morgan, P. Kulkarni, J. Vega, F. Heidarivincheh, R. McConville, A. Whone, M. Kim, R. Kirkham, and I. Craddock. Exploring perceptions of cross-sectoral data sharing with people with parkinson's. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2022.

[56] N. Meng, D. Keküllüoğlu, and K. Vaniea. Owning and sharing: Privacy perceptions of smart speaker users. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–29, 2021.

[57] C. W. Munyendo, Y. Acar, and A. J. Aviv. "desperate times call for desperate measures": User concerns with mobile loan apps in kenya. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 2304–2319. IEEE, 2022.

[58] D. Napoli, K. Baig, S. Maqsood, and S. Chiasson. "i'm literally just hoping this will {Work:'}'obstacles blocking the online security and privacy of users with visual disabilities. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, pages 263–280, 2021.

[59] H. Nissenbaum. Privacy as contextual integrity. *Wash. L. Rev.*, 79:119, 2004.

[60] B. Obada-Obieh, Y. Huang, and K. Beznosov. Challenges and threats of mass telecommuting: a qualitative study of workers. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, pages 675–694, 2021.

[61] B. Obada-Obieh, L. Spagnolo, and K. Beznosov. Towards understanding privacy and trust in online reporting of sexual assault. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, pages 145–164, 2020.

[62] A. Ponticello, M. Fassl, and K. Krombholz. Exploring authentication for {Security-Sensitive} tasks on smart home voice assistants. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, pages 475–492, 2021.

[63] Y. Rashidi, T. Ahmed, F. Patel, E. Fath, A. Kapadia, C. Nippert-Eng, and N. M. Su. "you don't want to be the next meme": College students' workarounds to manage privacy in the era of pervasive photography. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 143–157, 2018.

[64] H. Ray, F. Wolf, R. Kuber, and A. J. Aviv. "warn them" or "just block them"?: Investigating privacy concerns among older and working age adults. *Proceedings on Privacy Enhancing Technologies*, 2021.

[65] E. M. Redmiles. "should I worry?" a cross-cultural examination of account security incident response. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 920–934. IEEE, 2019.

[66] R. W. Rogers. A protection motivation theory of fear appeals and attitude change. *The journal of psychology*, 91(1):93–114, 1975.

[67] M. Saeidi, M. Calvert, A. W. Au, A. Sarma, and R. B. Bobba. If this context then that concern: Exploring users' concerns with ifttt applets. *arXiv preprint arXiv:2012.12518*, 2020.

[68] V. Saini, Q. Duan, and V. Paruchuri. Threat modeling using attack trees. *Journal of Computing Sciences in Colleges*, 23, 04 2008.

[69] C. Salter, O. S. Saydjari, B. Schneier, and J. Wallner. Toward a secure system engineering methodolgy. In *Proceedings of the 1998 workshop on New security paradigms*, pages 2–10, 1998.

[70] N. Sambasivan, A. Batool, N. Ahmed, T. Matthews, K. Thomas, L. S. Gaytán-Lugo, D. Nemer, E. Bursztein, E. Churchill, and S. Consolvo. "they don't leave us alone anywhere we go" gender and digital abuse in south asia. In *proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

[71] N. Sambasivan, G. Checkley, A. Batool, N. Ahmed, D. Nemer, L. S. Gaytán-Lugo, T. Matthews, S. Consolvo, and E. Churchill. "privacy is not for me, it's for those rich women": Performative privacy practices on mobile phones by women in south asia. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 127–142, Baltimore, MD, Aug. 2018. USENIX Association.

[72] P. Samermit, A. Turner, P. G. Kelley, T. Matthews, V. Wu, S. Consolvo, and K. Thomas. Millions of people are watching you: Understanding the {Digital-Safety} needs and practices of creators. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 5629–5645, 2023.

[73] M. K. Scheuerman, J. A. Jiang, C. Fiesler, and J. R. Brubaker. A framework of severity for harmful content online. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–33, 2021.

[74] S. Shapiro, C. Bloom, B. Ballard, S. Slotter, M. Paes, J. McEwen, R. Xu, and S. Katcher. The PANOPTIC privacy threat model. Technical Report Version 1.1, MITRE, Apr. 2024.

[75] A. Shostack. *Threat modeling: Designing for security*. John Wiley & Sons, 2014.

[76] M. Sleeper, T. Matthews, K. O'Leary, A. Turner, J. P. Woelfer, M. Shelton, A. Oplinger, A. Schou, and S. Consolvo. Tough times at transitional homeless shelters: Considering the impact of financial insecurity on digital security and privacy. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

[77] J. Slupska, S. Cho, M. Begonia, R. Abu-Salma, N. Prakash, and M. Balakrishnan. "they look at vulnerability and use that to abuse you": Participatory threat modelling with migrant domestic workers. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 323–340, 2022.

[78] J. Slupska, S. D. Dawson Duckworth, L. Ma, and G. Neff. Participatory threat modelling: Exploring paths to reconfigure cybersecurity. In *extended abstracts of the 2021 CHI conference on human factors in computing systems*, pages 1–6, 2021.

[79] D. J. Solove. *Understanding privacy*. Harvard university press, 2010.

[80] E. Steinbrink, L. Reichert, M. Mende, and C. Reuter. Digital privacy perceptions of asylum seekers in germany: An empirical study about smartphone usage during the flight. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–24, 2021.

[81] S. Stephenson, M. Almansoori, P. Emami-Naeini, and R. Chatterjee. "it's the equivalent of feeling like you're in {Jail"}: Lessons from firsthand and secondhand accounts of {IoT-Enabled} intimate partner abuse. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 105–122, 2023.

[82] A. Strohmayer, R. Bellini, and J. Slupska. Safety as a grand challenge in pervasive computing: Using feminist epistemologies to shift the paradigm from security to safety. *IEEE Pervasive Computing*, 21(3):61–69, 2022.

[83] K. Sun, C. Sugatan, T. Afnan, H. Simon, S. A. Gelman, J. Radesky, and F. Schaub. "they see you're a girl if you pick a pink robot with a skirt": A qualitative study of how children conceptualize data processing and digital privacy risks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–34, 2021.

[84] K. Sun, Y. Zou, J. Radesky, C. Brooks, and F. Schaub. Child safety in the smart home: parents' perceptions, needs, and mitigation strategies. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–41, 2021.

[85] M. Tabassum, T. Kosinski, and H. R. Lipford. "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)*, pages 435–450, 2019.

[86] K. Take, K. Gallagher, A. Forte, D. McCoy, and R. Greenstadt. "it feels like whack-a-mole": User experiences of data removal from people search websites. *Proceedings on Privacy Enhancing Technologies*, 2022(3), 2022.

[87] K. Thomas, D. Akhawe, M. Bailey, D. Boneh, E. Bursztein, S. Consolvo, N. Dell, Z. Durumeric, P. G. Kelley, D. Kumar, et al. Sok: Hate, harassment, and the changing landscape of online abuse. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 247–267. IEEE, 2021.

[88] K. Thomas, P. G. Kelley, S. Consolvo, P. Samermit, and E. Bursztein. "it's common and a part of being a content creator": Understanding how creators experience and cope with hate and harassment online. In *Proceedings of the 2022 CHI conference on human factors in computing systems*, pages 1–15, 2022.

[89] E. Tseng, M. Sabet, R. Bellini, H. K. Sodhi, T. Ristenpart, and N. Dell. Care infrastructures for digital security in intimate partner violence. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–20, 2022.

[90] W. Usman, J. Hu, M. Wilson, and D. Zappala. Distrust of big tech and a desire for privacy: Understanding the motivations of people who have voluntarily adopted secure email. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*, pages 473–490, 2023.

[91] N. Warford, T. Matthews, K. Yang, O. Akgul, S. Consolvo, P. G. Kelley, N. Malkin, M. L. Mazurek, M. Sleeper, and K. Thomas. Sok: A framework for unifying at-risk user research. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 2344–2360. IEEE, 2022.

[92] H. Watson, E. Moju-Igbene, A. Kumari, and S. Das. "we hold each other accountable": Unpacking how social groups approach cybersecurity and privacy together. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[93] M. Wei, S. Consolvo, P. G. Kelley, T. Kohno, F. Roesner, and K. Thomas. "there's so much responsibility on users right now:" expert advice for staying safer from hate and harassment. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2023.

[94] D. Wilkinson and B. Knijnenburg. Many islands, many problems: An empirical examination of online safety behaviors in the caribbean. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–25, 2022.

[95] P. J. Wisniewski and X. Page. Privacy theories and frameworks. In *Modern socio-technical perspectives on privacy*, pages 15–41. Springer International Publishing Cham, 2022.

[96] H. Xia, Y. Wang, Y. Huang, and A. Shah. "our privacy needs to be protected at all costs" crowd workers' privacy experiences on amazon mechanical turk. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):1–22, 2017.

[97] S. Zhang, Y. Feng, and N. Sadeh. Facial recognition: Understanding privacy concerns and attitudes across increasingly diverse deployment scenarios. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, pages 243–262, 2021.

[98] D. Zhao, M. Inaba, and A. Monroy-Hernández. Understanding teenage perceptions and configurations of privacy on instagram. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–28, 2022.

[99] J. Zhao, G. Wang, C. Dally, P. Slovak, J. Edbrooke-Childs, M. Van Kleek, and N. Shadbolt. I make up a silly name' understanding children's perception of privacy risks online. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019.

[100] S. Zheng, N. Apthorpe, M. Chetty, and N. Feamster. User perceptions of smart home iot privacy. *Proceedings of the ACM on human-computer interaction*, 2(CSCW):1–20, 2018.

Appendix A.
Threat Modeling Definitions

Table 5 presents the definitions of threat modeling contained in papers from our corpus.

TABLE 5. DEFINITIONS OF "THREAT MODEL" OR "THREAT MODELING" IN OUR CORPUS

Title	Citation	Definition
"Should I worry?" A cross-cultural examination of account security	[65]	Who they would be most worried about gaining access to their account (e.g., friend, stranger) and why (e.g., what they would be concerned about this person doing or accessing).
Privacy and activism in the transgender community	[45]	Adversaries, threats, risks, defenses, and other factors they use to make decisions surrounding their goals
"There's so much responsibility on users right now:" Expert advice for staying safer from hate and harassment	[93]	Which online hate and harassment threats do experts believe most internet users should prioritize taking action to prevent or mitigate, and why?
Weaving privacy and power: on the privacy practices of labor organizers in the US technology industry	[38]	The process of taxonomizing the goals and risks of a particular situation or context and creating a plan to address them information security
Defensive Technology Use by Political Activists During the Sudanese Revolution	[18]	What are/were the dangers participants are/were facing as an activist? Who is an adversary to them? If they mention the government as an adversary: what arm(s) of the government might be harmful? For each: what are their capabilities? What do you use to defend against them? Is that enough to protect them?
"They Look at Vulnerability and Use That to Abuse You": Participatory Threat Modelling with Migrant Domestic Workers	[77]	Participatory threat modeling applies invites participants to identify and prioritise threats to their privacy and security. Therefore, it is an open-ended process which does not focus on a specific type of device or context but rather centres participants' perspectives.
Privacy and security threat models and mitigation strategies of older adults	[25]	The activities that can lead to security and privacy risks, along with the consequences of privacy and security violations.
"Millions of people are watching you": Understanding the digital-safety needs and practices of creators	[72]	Perceived or experienced digital-safety threats, attackers, and potential harms are they concerned about.

Appendix B.
Meta-Review

The following meta-review was prepared by the program committee for the 2025 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

B.1. Summary

This paper presents a systematization of knowledge on human-centered threat modeling in the security and privacy field by analyzing 78 papers. It asks asks (a) how researchers define human-centered threat modeling, (b) how researchers elicit a human-centered threat model from their subjects, and (c) what key components there are in human-centered threat modeling. The authors define a framework for future researchers to use when human-centered threat modeling.

B.2. Scientific Contribution

- Creates a New Tool to Enable Future Science
- Provides a Valuable Step Forward in an Established Field
- Establishes a New Research Direction

B.3. Reasons for Acceptance

1) This paper creates a new tool to enable future science. The presented framework for human-centered threat modeling has potential to better support and guide future research into human-centered threat models.
2) This paper both provides a valuable step forward in the established field and establishes a new research direction. Threat modeling analyzes systems for vulnerabilities and potential mitigations. This paper systematizes human-centered threat modeling, which focuses on threats against people, emphasizing users' goals and experiences.
3) This paper outlines the selection and survey of research works for the systematization in distinct and thorough steps that are well-written and easy to follow.