SoK: A Framework for Unifying At-Risk User Research

Noel Warford*, Tara Matthews†, Kaitlyn Yang*, Omer Akgul*, Sunny Consolvo†,
Patrick Gage Kelley†, Nathan Malkin*, Michelle L. Mazurek*, Manya Sleeper†, and Kurt Thomas†
*University of Maryland, †Google

2022 IEEE Symposium on Security and Privacy (SP) | 978-1-6654-1316-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/SP46214.2022.9833643

***Abstract*—At-risk users are people who experience risk factors that augment or amplify their chances of being digitally attacked and/or suffering disproportionate harms. In this systematization work, we present a framework for reasoning about at-risk users based on a wide-ranging meta-analysis of 95 papers. Across the varied populations that we examined (e.g., children, activists, people with disabilities), we identified 10 unifying *contextual risk factors*—such as *marginalization* and *access to a sensitive resource*—that augment or amplify digital-safety risks and their resulting harms. We also identified technical and non-technical *practices* that at-risk users adopt to attempt to protect themselves from digital-safety risks. We use this framework to discuss *barriers* that limit at-risk users' ability or willingness to take protective actions. We believe that researchers and technology creators can use our framework to identify and shape research investments to benefit at-risk users, and to guide technology design to better support at-risk users.**

I. Introduction

Anyone can experience attacks related to their security, privacy, or safety online (i.e., *digital safety*), but *at-risk users* have risk factors that augment or amplify their chances of being digitally attacked and/or suffering disproportionate harms. For example, some activists are surveilled by government actors due to their work [59, 90]; people who are LGBTQ+ face elevated risk of harassment by anonymous attackers on social media [15, 16]; and women in repressive regions experience pervasive sexual harassment online and sometimes severe consequences from their community as a result [88, 109].

A growing body of research has explored how the digital-safety needs of at-risk users may be unmet by existing security, privacy, and safety threat models that tend to focus on a mythical "average user." A common recommendation from researchers in this space is to consider at-risk users during the technology creation process (e.g., [13, 25, 63, 97, 108, 124]). However, for technology creators, it can be bewildering to consider dozens of different at-risk populations, each with disjoint and sometimes contradictory digital-safety needs. Accordingly, we argue that there is a need for synthesis: to organize what is known into a framework that can be used to reason about at-risk users' risks and needs, and to identify gaps in knowledge for future work.

We systematically identified and reviewed 95 papers focused on the digital-safety experiences of at-risk populations and developed a framework that can be used to reason about four research questions regarding at-risk users:

- **RQ1:** *Contextual risk factors*. What factors—such as a person's situation in society, relationships, or personal circumstances—contribute to digital-safety risks for at-risk users?
- **RQ2:** *Interactions*. How do these contextual risk factors interact to elevate the risk or severity of digital-safety attacks for at-risk users?
- **RQ3:** *Protective practices*. What protective practices are common across at-risk users when attempting to address their digital-safety risks?
- **RQ4:** *Barriers.* What barriers do at-risk users encounter in protecting themselves from digital-safety risks?

Based on an analysis across 31 distinct population categories (e.g., journalists, refugees, older adults), we identified 10 contextual risk factors that cross-cut at-risk populations, yielding a set of circumstances that technology creators and researchers can consider in research, design, and development. We also found that at-risk users currently rely on varied, often ad-hoc protective practices, ranging from leaning on social connections to relying on a patchwork of technical strategies to try to minimize risks and harms. We provide an *at-risk framework* comprised of these *contextual risk factors* and *protective practices*, which we use to discuss *barriers* that limit or prevent at-risk users from enacting digital protections, and to show how competing priorities, a lack of digital safety awareness, and broken technology assumptions compound the challenges at-risk users face.

We advocate for technology creators and researchers to consider at-risk users' needs in risk modeling and design. Our framework provides a blueprint for addressing these issues through research, education support, and technology creation, to better ensure that at-risk users can engage safely online, and in the process, to improve digital safety for everyone.

II. Who are at-risk users?

In this paper, we use *at-risk user* as an umbrella term for anyone with risk factors that augment or amplify their chances of being attacked digitally and/or suffering disproportionate harms from an attack. We refer to groups of at-risk users as *at-risk populations*. Due to a lack of consensus in the literature on how to refer to such users or populations, we chose these terms with the goal of drawing focus to external risks these users face.

A. Previous taxonomies of attacks, threats, and harms

Previous systematizations developed frameworks to broadly categorize attacks, threats, and harms, although none capture how these elements overlap or differ across distinct at-risk populations. In particular, Scheuerman et al. [92] and Thomas

DOI 10.1109/SP46214.2022.00116

et al. [103] developed frameworks for understanding classes of harms that may result from digital-safety attacks, such as reputational harm, financial harm, reduced sexual safety, reduced physical safety, and coercion. Scheuerman et al. [92] also provided a framework for assessing the severity of threats based on such harms. Thomas et al. [103], Sambasivan et al. [88], and Levy and Schneier [57] detailed how attacks vary based on the capabilities of attackers, such as having intimate access to a target, or privileged access to a target's devices or data. Our at-risk framework differs in that we isolate the contextual risk factors that can make at-risk users particularly vulnerable to such attacks, threats, or harms. We also document common protective practices at-risk users adopt and discuss barriers they face to staying safe.

B. Value of focusing on at-risk users

The challenges experienced by at-risk users can be inordinately complex, reflecting broader, societal "structural inequalities and social norms" [35, 65]. These inequalities, which vary globally, mean that particular care is required to integrate at-risk users' experiences and identities into the technology creation process [48, 65, 111].

We advocate for increased focus on at-risk users' needs by technology creators and researchers during threat modeling, research, design, and development. Accounting for at-risk users can also elevate the digital safety of all users by making "more pronounced the need[s] that many of us have" [29]. Providing better digital-safety tools and guarantees can have far-reaching impact both to at-risk users and general users. Additionally, providing choices and controls for at-risk users who know intimately the digital-safety threats they face can also benefit general users who may desire similar protections.

III. Methods

We synthesize 95 research papers from a cross-section of computer science conferences. Here, we discuss how we identified and analyzed these papers.[1]

A. Paper selection

Our dataset for this analysis was 95 papers describing digital-safety-related issues for various at-risk populations. We collected papers from five years (2016–2020) of conferences spanning the security, privacy, and human-computer interaction (HCI) communities: CCS, CHI, CSCW, IEEE S&P, NDSS, PETS, SOUPS, and USENIX Security. We first gathered links to every paper from these conferences on DBLP.[2] From those links, we collected paper titles, abstracts, and publication dates, resulting in 6,534 papers.

To refine this list, three researchers independently read titles and abstracts for each paper and marked them as 'relevant' to our research questions or not. At this stage, we interpreted relevance broadly, selecting any paper even slightly within scope. Papers that no researcher marked as relevant were removed. Papers marked as relevant by only one researcher were reviewed by a fourth researcher and discussed. This process identified 127 potentially relevant papers.

Authors with extensive experience working with at-risk populations added 12 papers from other sources and/or from outside the target date range, in order to cover a broader range of populations, for a total of 139 potentially relevant papers.

B. Codebook development

Our goal was to identify contextual risk factors, protective practices, and other patterns discussed by the papers in our dataset. As a first step, we inductively built a codebook by analyzing, in detail, a subset of papers well-aligned with our research questions. Most of the core concepts in our framework were identified at this stage, although inductive refinement continued throughout our analysis.

To select this initial subset, we extracted from the dataset an initial list of populations (e.g., survivors of intimate partner abuse [63], refugees [96], activists [25], children [124], etc.). We also synthesized an initial list of risk factors, for example, attributes of the population or the threats they faced that contributed to their digital-safety-related risks. We then selected our subset to ensure each population and risk factor on our list was represented, making sure to include some papers that combined multiple risk factors (e.g., low-income African American New York City residents [31] and foster teens [12]). This process yielded 27 papers.

We next analyzed these 27 papers and inductively built our codebook. We used the specific population discussed in each paper[3] as a unit of analysis [102]. One of four researchers read and summarized each paper. The full team used these summaries to iteratively build and refine our codebook [102], which included categories for risk factors, protective practices, and barriers to protection. We met throughout the process to develop, discuss, and refine codes. These detailed summaries also enabled us to examine relationships between codes and memo early ideas on themes [17].

C. Full analysis

Next, we used the codebook to analyze the remaining papers Because our initial paper selection steps were deliberately inclusive, we continued to refine the dataset during this phase. Any researcher could flag a paper for possible exclusion if it did not address any of our research questions; determinations were made after discussion. Our final analyzed dataset included 95 papers.

We randomly selected 20% of the remaining papers, which two researchers independently coded, making minor updates to the codebook as needed. The researchers assigned codes to each paper and memoed additional relevant details and context [17]. After completing this 20% sample, the researchers calculated agreement using Krippendorff's α before discussing and resolving all disagreements. The researchers then repeated

[1] Additional method details can be found at https://arxiv.org/abs/2112.07047

[2] See https://dblp.uni-trier.de/search

[3] For example, Simko et al. [96] reported on interviews with refugees *and* associated caseworkers; we treated this as two populations.

this procedure on another random 20%. Once sufficient reliability was reached, the researchers split the remaining papers between them to complete coding.

For the contextual risk factors and protective practices, we obtained $\alpha = [0.88, 1.00]$ after the second round of coding, and for the barriers, $\alpha = 0.83$ after the third round. As these are above the standard threshold of 0.8 for reliability [51], we then proceeded to individual coding. After coding all papers, the entire research team met to review the results, identify second-level codes and themes [17], and finalize our framework.

We conducted credibility checks of our framework and findings to verify they were accurate and clear. Seven experts who have worked with at-risk populations reviewed a version of this paper and met with us to discuss it. All experts found the framework sound and useful.

D. Limitations

Our 95-paper dataset is not exhaustive of all relevant papers published in the security, privacy, or HCI communities. However, given our systematic method of compiling papers, it should reasonably represent these communities' published understanding of at-risk users. Also, reflecting the current state of literature from these communities, the dataset papers skewed heavily toward Western, and specifically U.S., populations. In addition, literature from other fields, especially the social sciences, could offer relevant perspectives on the digital safety of at-risk populations. (See the appendix for details on geographic representation and researcher reflexivity.)

As research methods and best practices for understanding at-risk users are still being developed—often differing from one community to the next—the papers in our dataset also often did not focus on the same issues or investigate to the same depth. As a result, our synthesis of the contextual risk factors, protective practices, and barriers covered in this paper may not reflect all the challenges the population in question experiences. Our coding is, instead, a reflection of the current understanding in the sampled literature on each population.

Despite these limitations, we believe this work serves as a critical first step towards recognizing contextual risk factors and protective practices that span at-risk users. We advocate for future work that builds on this framework by including broader literature and cultural perspectives.

IV. Contextual risk factors

In our meta-analysis, we identified 10 contextual risk factors that augmented or amplified digital-safety risks. These risk factors, which form the first component of our at-risk framework, include: three *societal factors*, influenced by an at-risk user's role in their society and culture; three *relationship factors* stemming from who an at-risk user knows or interacts with; and four *personal circumstances* dependent on who an at-risk user is or their personal or professional activities. We note some common attacks for each risk factor, but do not consider being the target of a digital attack alone a risk factor.

We capture the presence of these risk factors within the papers from our dataset in Table I. A black circle (●) indicates that we identified at least one prior study of an at-risk population that reported risks related to that factor. We caution that the absence of a black circle for an at-risk population in Table I does ***not*** imply that the risk factor is irrelevant to the population, only that it was not reported in our dataset.

Next, we describe each contextual factor, focusing on the nature of the risk and types of associated attackers[4] and harms, as applicable. We also explore how risk factors may intersect to create more severe risks.

A. Societal factors

The first set of contextual risk factors involve *societal factors* amorphously driven by cultures and institutions. Attacks related to societal risk factors tended to be *diffusely targeted*, i.e., directed toward anyone in a population or an entire at-risk population simultaneously, rather than at a specific person.

Legal or political. The government, political affairs, or laws of a country can contribute to at-risk populations experiencing heightened digital-safety risks, including potentially sophisticated attacks from government actors. A key theme associated with this factor was the power differential between government or quasi-governmental actors and the targeted populations.

Government or quasi-government actors may be able to intercept communications from at-risk populations in various ways, such as physically seizing devices or data [55, 67], impersonating trusted entities [59], coercing platform or telephony providers to bypass security measures [90], or preventing internet access entirely [25]. For example, in 2019, the Sudanese government shut off the country's mobile data network to make organizing for activism as difficult as possible [25]. Similarly, the International Committee of the Red Cross—a non-governmental organization (NGO) that collects information that could be used by armed groups for non-humanitarian intelligence—reported being obligated to physically surrender devices to meet with those armed groups [55].

Governments may also be able to enact surveillance, leading to real or perceived threats of monitoring. For example, undocumented immigrants in the U.S. reported concern about posting their activities on social media, due to perceived government monitoring [40]. More generally, residents of several countries with government-controlled internet surveillance have reported modifying their behavior [46, 89, 109].

Because of the power differential, threats associated with this factor may also escalate into offline harms, such as detainment, incarceration, or deportation [25, 40]. These harms may also have wide-ranging societal impacts, including restricted or self-censored speech [49, 101] or lasting damage to trust in public institutions or figures [21, 68]. For example, attacks on people involved with U.S. political campaigns were described as intending to undermine the institution of U.S. elections [21].

Marginalization. Pervasive negative treatment or exclusion at a societal level, due to an individual's identity attributes or life experiences, may also elevate digital-safety risk.

[4]We use *attacker* broadly to refer to anyone who introduces digital-safety issues for an at-risk user, regardless of the severity or intention.

		Societal factors			Relationships			Personal circumstances			
Population Category	**Citations**	Legal or political	Marginalization	Social norms	Relationship with the attacker	Reliance on a third party	Access to other at-risk users	Prominence	Resource or time constrained	Underserved accessibility needs	Access to a sensitive resource
Children	[37, 39, 47, 52–54, 72, 74, 77, 124]			●		●				●	
Teens	[37–39, 71, 118–120]			●	●	●				●	
Foster teens	[12]				●	●				●	
Older adults	[36, 45, 73, 78]					●			●	●	
Activists	[9, 25, 49, 59, 101]	●						●	●		●
Activists × Transgender people	[56]	●	●					●			●
People involved with political campaigns × US	[21]	●					●	●	●		●
Teachers	[53]						●		●		
Journalists	[30, 67–70]	●					●	●	●		●
Sex workers	[13, 100]	●	●								
ER staff	[99]								●		●
NGO staff	[19, 55]	●					●	●			●
Crowd workers	[122]				●				●		
LGBTQ+ people	[15, 16, 20, 41, 56, 91]		●								
LGBTQ+ people × With HIV	[113, 114]		●								
Marginalized racial group × US	[104]		●								
People with an illness	[81, 87, 94]		●								
Older adults × With cognitive impairments	[14, 22, 36, 66, 76]					●			●	●	
People with visual impairments	[5–7, 32, 43, 112]			●		●				●	
People with other or multiple disabilities	[61, 82]		●							●	
Non-Western culture × Women	[8, 28, 88, 89, 109]		●	●	●	●			●		
Developing regions × Older adults	[46]	●		●						●	
Developing regions × Low SES	[2–4, 75, 86, 108]	●	●	●	●				●		
Developed regions × Low SES	[50, 83, 84, 97, 110, 115, 117]		●	●	●				●		
Developed regions × Low SES × Marginalized racial group	[31]		●	●					●		
Undocumented immigrants	[40]	●	●							●	
Refugees	[96]			●		●			●	●	
People involved with armed conflict	[95]	●									
Survivors of sexual assault	[10, 79]		●		●	●					
Survivors of intimate partner abuse	[18, 33–35, 42, 63]		●		●	●			●		
Survivors of trafficking	[19]	●			●	●			●		

TABLE I: Meta-analysis showing how the contextual risk factors from our at-risk framework apply to categories of populations in our dataset. A black circle means one or more citations indicated the contextual risk factor was relevant. Note that the absence of a black circle does not mean the population does not have that digital-safety risk. Rather, it means the set of papers cited did not discuss or explore that risk factor as defined in this framework.

Attackers may target at-risk populations with online hate and harassment due to their beliefs, identity, or social status, such as people who are LGBTQ+ [15, 16, 41, 56, 91, 113, 114], undocumented immigrants [40], marginalized racial and ethnic groups [31, 104], or people with low socioeconomic status (SES) [31, 50]. These threats were typified by a broad set of potential attackers, and contending with hate and harassment led to emotional distress [16, 104, 109]. For this factor, past research has described the perception that anyone may be a potential attacker [91], exacerbated by the fact that attackers may hide behind anonymous identities [88, 104].

Harassment experienced by marginalized populations can be subtle or even unintentional. For example, in To et al. [104], people in marginalized racial and ethnic groups have reported experiencing widespread microaggressions pertaining to their race, via their interpersonal interactions with others on- and offline. Given the prevalence of this threat, some populations may be reluctant to fully participate online due to the risk of revealing a marginalized characteristic [10, 13, 35, 35, 79, 94, 113, 114].

Given that this factor can be associated with stigmatized information, attackers may also coerce at-risk users by threatening to leak information that could be harmful. This can lead to reputation damage [109, 113] or sexual violence [40, 88]. For example, attackers may threaten to "out" transgender individuals [56] or in conservative regions, may threaten to

leak chat logs between a man and a woman to damage the woman's reputation [88, 109]. Algorithmic bias may also create or exacerbate digital-safety risks in this space. For example, automated gender recognition systems may misgender transgender people [41].

Social norms. Social norms are informal rules that govern behavior in society. Some norms can particularly restrict options for members of an at-risk population, leading to increased digital-safety risks when there is a mismatch between the expectations of technology creators and the lived experiences of that population.

For example, technology creators may assume that devices and accounts are personal and private. However, device-sharing norms mean that this assumption may not hold for a variety of at-risk populations, limiting the privacy protections afforded by private personal devices and accounts. For example, Sambasivan et al. and Ahmed et al. found that women in South Asia were expected to share devices or accounts with family members [2, 89]. Similar device sharing norms were also found among rural women in Greenland [117].

Technology creators may also assume that users understand implicit norms around when it is appropriate (or dangerous) to share personal information. However, some at-risk populations face changing norms, which can lead to unexpected threats. For example, alongside numerous other challenges, refugees must adapt to the technology norms and associated digital-safety risks of their new country of residence. For refugees in the U.S. specifically, certain types of information, such as Social Security numbers, were more sensitive than expected, placing refugees at increased risk of financial harm until they learned when it was appropriate to disclose this type of information [96].

B. Relationships

The second set of risk factors is driven by the *relationships* of at-risk users, including direct relationships with an attacker and relationships with a third party. These risks tended to be associated with *focused targeting*, in which attackers pursue specific at-risk users (often with intense motivation).

Relationship with the attacker. A personal relationship with an attacker can lead to heightened risk of what Levy and Schneier [57] classify as "intimate threats." It can be particularly difficult for individuals targeted by intimate threats to prevent and detect harm, because the attackers are likely to have physical access to the target's devices and accounts, implicit or explicit authority over the target, and potentially detailed knowledge about the target they can leverage [57]. Intimate threats can involve a wide range of digital-safety attacks including surveillance, device or account compromise, destruction of data or devices, harassment, and more [57].

For example, survivors of intimate partner abuse [33, 35, 42, 63] were reported as often facing relentless attacks from abusers aiming to limit the survivor's autonomy. These abusers may have physical or digital access to the survivors' devices and accounts, including surveillance via coercive physical access [63] or spyware [18, 35]. Similarly, survivors of trafficking reported worrying about being recaptured because, prior to escape, the trafficker had full access to the survivor's digital life [19]. Stalking attacks are also often characterized by a *relationship with the attacker*, as "the majority of stalkers are . . . obsessionally focused on a specific person with whom they have had some previous relationship," often an intimate one [26]. Workers on online crowdsourcing platforms also experienced a severe power imbalance with requesters on those platforms, who could decide whether or not to pay workers based on their responses, and could gather detailed information on them via completed tasks [122].

Reliance on a third party. Individuals may be at risk due to their *reliance on a third party* for safety-focused care or help with essential tasks. Unlike intimate threats associated with a *relationship with the attacker*, the third parties in this context typically have helpful or safety-focused reasons for "privacy invasions" [57]. However, even with generally supportive intentions, these invasions can increase the individual's digital-safety risks due to reduced privacy (from the third party), leading at-risk users to feel uncomfortable or that their autonomy is limited [76]. These types of risks can also increase the attack surface (e.g., through the third-party), or leave at-risk users vulnerable to attackers that impersonate the third party.

For example, children and teens often shared their devices and personal data with parents or caregivers [52, 72], or had applications or privacy settings enabled that allowed parental monitoring [38, 39, 118]. People with visual impairments sent photos to crowdsourced services for object identification, but also worried that there might be sensitive content in the photos (e.g., credit card numbers) [7]. Older adults with cognitive impairments depended on help from caregivers for digital tasks [36, 76]. Refugees relied on case workers for processing personal data, applying to jobs, or applying for support [96].

Access to other at-risk users. Having access to another at-risk person or population (the "primary target") can also put someone at greater risk of focused, stepping-stone attacks that ultimately aim to harm the primary target. For example, children of survivors of intimate partner abuse may be targeted by abusers in order to regain access to the survivor [63], journalists may be targeted to try to gain access to their sources [30, 67–70], and elementary school teachers may be targeted to gain access to their students [53].

C. Personal circumstances

Our third set of contextual risk factors cover *personal circumstances* that can increase risk, such as public prominence or socioeconomic constraints. All of these factors tended to involve *diffuse targeting* of groups, except *prominence*, which involved *focused targeting* of specific individuals.

Prominence. At-risk users who stand out in a population, because they are well-known publicly or have noticeable attributes (e.g., accomplishments, outspokenness, attractiveness,

etc.), may face heightened risk associated with their prominence. *Prominence* can expose individuals to new attacks, like *focused targeting*. Celebrities, for example, may experience parasocial relationships with audiences, increasing the risk of stalking or personal information leaking [121].

Prominence may also exacerbate risks associated with *legal or political* factors or *marginalization*, a dynamic identified for prominent journalists [30, 67–69], politicians [21], NGO staff [55], and activists [25, 59]. In these cases, the level of prominence influenced the severity of risks. For example, people involved in a national U.S. political campaign were more at risk of focused attacks than campaigns receiving only state or local attention [21].

Resource or time constrained. Populations can experience elevated digital-safety risks and decreased ability to respond to digital-safety threats if they have limited access to technology or other resources (e.g., money, devices, connectivity), or limited ability to set aside time to cope with risks. At-risk users experiencing this risk factor face constraints that go beyond what is typical. Individuals with low-SES, for example, may not be able to afford private personal devices, or may need to rely on old devices that can no longer receive security updates [31, 97, 108, 110].

At-risk users can also face extraordinary time constraints. Political workers, for example, operated within extremely limited election timelines, which made it challenging to set up the security infrastructure needed to counter nation-state attackers [21]. Hospital emergency departments similarly "have strong availability demands ... and must provide services as quickly as possible" [99], leading to password reuse and use of unsecured personal devices in a sensitive work context.

Underserved accessibility needs. Some at-risk users have accessibility needs that are underserved by current technology, contributing to their digital-safety risks. In our dataset, this included accessibility needs due to a disability, neurodiversity, a language barrier, or developmental maturity. Inaccessible technology can cause anxiety about, and susceptibility to, potential attacks.

Members of populations experiencing this risk factor have described general anxiety about falling prey to "hackers" or vague bad actors, as well as worries about their ability to effectively protect themselves with existing technology. For example, some older adults reported asking trusted sources for digital-safety help because they did not feel confident protecting themselves [36]. This can stem from negative past experiences or from inaccessible language in online resources about digital safety [78]. Assumptions built into systems were unrealistic for older adults with mild cognitive impairments, who were sometimes unable to remember passwords and other crucial information [76]. In some cases, they had trouble remembering whether they made a particular purchase [66], making it difficult to differentiate an attack from a memory lapse.

Similarly, people with disabilities may struggle with a lack of accessible technology for some tasks [5–7, 32, 43, 61, 112]. For example, people with visual impairments reported to Wang et al. the conflict between the need to use screen readers in public and concerns about eavesdropping and safety [112].

Additionally, refugees with developing English skills reported struggling with language accessibility, finding it difficult to distinguish legitimate callers from scammers when the call was in English [96]. Zhao et al. found that children do not completely understand certain online privacy risks due to their age and development [124].

Access to a sensitive resource. Access to a sensitive resource (e.g., sensitive data, credentials, money) can increase the risk of attacks aimed at co-opting this access. In most cases, the at-risk users' professional activities provided them privileged access to these resources.

For example, emergency department staff may be targeted for their access to patient medical data [99], journalists for their access to original source material, such as legal documents or financial records [30, 67, 68], and executive staff for their ability to approve wire transfers [27].

D. How do contextual risk factors interact?

Most at-risk populations in Table I experienced more than one contextual risk factor. Each factor contributed to risk on its own, but risk factors also combined to yield new digital-safety risks or amplify existing risks. Thus, we argue that technology creators and researchers should consider all of an at-risk population's risk factors together when possible. This is related to prior work on intersectionality [24, 93], which considers how multiple marginalized identities or circumstances can combine to create unique modes of discrimination. We provide an illustrative list of examples from our thematic analysis below, chosen because they represented experiences reported in multiple papers across populations and/or provided understandable illustrations of how risk factors could combine to yield new risks or amplify each other.

***Prominence* added *focused targeting* to other risk factors.** When added to other risk factors, *prominence* appeared to make it more likely that an at-risk user would experience *focused targeting* that amplified their other risk factors. For example, political campaign workers had *access to a sensitive resource*, but *prominent* politicians were more likely to be subjected to *focused targeting* to access those sensitive resources [21]. Similarly, while many transgender people reported experiencing *marginalization*, the *prominence* of transgender activists resulted in highly targeted hate and harassment attacks [56].

***Resource or time constraints* made it harder to cope with other risk factors.** Populations who are *resource or time constrained* were more likely to experience worse outcomes pertaining to their other risk factors, because they did not have the time, money, or other resources to effectively protect themselves or recover from attacks. For example, the primary risk factor for survivors of intimate partner abuse was typically their *relationship with the attacker*, but prior research has em-

phasized that survivors with low SES had particular difficulty protecting their (and their children's) digital lives [63].

***Underserved accessibility needs* and *reliance on a third party* often combined.** In our dataset, at-risk users with *underserved accessibility needs* also tended to *rely on a third party* for care or help with technology, at least sometimes. This included children, teens, older adults, people with visual impairments, and refugees, as shown in Table I. For example, older adults experiencing mild cognitive impairment often forgot important passwords and information (related to *underserved accessibility needs*), which reinforced the need to share these passwords and information with caregivers (subjecting them to risks associated with *reliance on a third party*) [76].

V. Protective Practices

At-risk users employed practices they perceived would help them prevent, mitigate, or respond to digital-safety risks. These *protective practices* form the second part of our at-risk framework. The risk factors presented above helped drive user decisions about which protective practices to use, but these practices were not always ideal or even effective. Protective practices involved tradeoffs and highlight barriers to technology use (which we explore in Section VI). Different users weighed the pros and cons differently, sometimes leading to seemingly contradictory choices. We catalog these imperfect practices to show what at-risk users currently do given their risks, and to set the stage for a discussion of barriers to protections in Section VI, both of which provide context for how to design technologies intended to support at-risk users.

Our meta-analysis identified three categories of protective practices: *social strategies* where at-risk users relied on their social connections to respond to threats; *distancing behaviors* where at-risk users distanced themselves from, or entirely abandoned, certain accounts and technologies; and *technical solutions* that involved leveraging technical tools and mechanisms to prevent or respond to threats. We found these strategies were not mutually exclusive, with at-risk users commonly relying on multiple strategies simultaneously.

A. Social strategies

At-risk users frequently relied on social connections (in-person or online) to overcome digital-safety threats. This included relying on family or peers for trusted advice and support, vetting the identities of people they interacted with online, and controlling social interactions to minimize harms.

Informal help from trusted family and peers. A popular protective practice among at-risk users experiencing the *marginalization*, *social norms*, or *underserved accessibility needs* factors was to informally seek help from trusted family and peers. The at-risk user may be seeking direct help from someone perceived to be more knowledgeable or resourced, or may simply be seeking emotional support. For example, children and teenagers reportedly sought help from parents to understand security warnings [47] or deal with strange, scary, or confusing internet experiences [119, 120, 124]. Older adults who self-identified as having low technical understanding sought assistance from family members when addressing digital-safety concerns [36, 78]. At times, women in South Asia relied on family members for emotional support when harassed online [88, 89]. Informal support sometimes also came from anonymous peers, such as transgender individuals relying on social media to connect with other LGBTQ+ peers for emotional support when they experienced hate and harassment [56, 91], or people in armed conflict zones connecting to share critical information [95].

Formal help from trusted organizations. Public recognition of digital-safety threats facing at-risk users—particularly with respect to the *legal or political*, *marginalization*, *relationship with the attacker*, or *resource or time constrained* risk factors—has prompted trusted organizations to offer support. For example, organizations that assist survivors of trafficking and intimate partner abuse, including NGOs [19] and government agencies [33, 35, 42], have provided assistance setting up digital-safety protections that are interwoven with continuous care [42]. Similarly, NGOs have assisted refugees with technology-required tasks, such as applying for work or submitting legal documents [96]. Public institutions, like libraries, have provided assistance and access to computers and connectivity to people with low SES [110]. While at-risk users may rely on these institutions for a broad variety of technical assistance [96, 110], digital-safety support was often explicitly reported as a key benefit. This assistance could incur additional digital-safety risks due to *reliance on a third party*, but the information and aid provided was often critical, both for digital safety and other basic needs.

Formal support may also stem from dedicated professional resources, particularly for those who have *access to a sensitive resource* or *access to other at-risk users*. For example, journalists have benefited from institutional support and accumulated best practices, as exemplified by the International Consortium of Investigative Journalists' (ICIJ) tools for digital safety in the Panama Papers investigation [69]. Elementary school teachers using technology in classrooms—who have a responsibility for ensuring the digital safety of their students—have sought assistance from school media specialists or librarians [53]. Political campaign workers in the U.S. have received training from organizations such as Defending Digital Campaigns and the Center for Democracy & Technology [21].

Vetting identities to avoid potential attackers. At-risk users who experience *marginalization* or *legal or political* risk factors have a documented need to verify that new people they encounter online are safe to associate with or to admit into a private space. For example, social media community moderators for transgender communities of color asked potential new members questions about topics relating to race and LGBTQ+ issues to help protect their spaces from potential attackers [56]. Women from non-Western cultures scrutinized online profiles to prevent men from infiltrating their private discussion spaces and subjecting them to sexual harassment [88]. Sex workers informally shared lists of abusive clients and aggressors to

be avoided [13]. Undocumented immigrants restricted communication with untrusted parties until sufficiently vetted, due to concerns of impersonation by the police or immigration enforcement [40]. Activists used physical in-person meetings or mutually trusted peers to vet new individuals before adding them to sensitive group chats [25]. Unifying these practices is a lack of dependable digital signals of authenticity or identity, resulting in at-risk users relying on ad-hoc techniques to establish trust.

Preemptive disclosure for control. At-risk users experiencing the *marginalization* or *legal or political* risk factors sometimes proactively revealed sensitive personal information in a controlled manner to disempower attackers and preempt future emotional harm. For example, some men seeking men on dating apps opted to publicly disclose their HIV-positive status, both to simplify navigating sexual negotiations and as a path toward destigmatization [113, 114]. Some transgender people chose to "out" themselves to avoid the emotional stress of maintaining a secret identity and the risk of coercion or extortion [56]. Some people with disabilities—especially those with "visible" disabilities—chose to disclose these disabilities on dating apps, to reduce the chance of undesirable connections [82]. Some activists engaged in an "anti-surveillance tactic of openness" by collapsing public and private personas, rendering information they shared—like political beliefs—unusable against them in a separate context if leaked [90]. This practice is not applicable to all at-risk users, as it requires a willingness to be hyper-public, which carries potential risks akin to those that *prominent* at-risk users face.

Social pleas. Some at-risk users who experience the *marginalization* or *social norms* risk factors reported reaching out directly to attackers with a plea to cease activities, such as content leaks or harassment. For example, some women in South Asia either independently, or with the assistance of family, asked attackers within their community to cease online sexual harassment [88, 109]. Similarly, some teens engaged with peers [120] or parents [77] to have embarrassing content removed from social media. This practice leverages the attacker's empathy to help resolve digital-safety threats.

B. Distancing behaviors

The second category of protective practices in our framework involves *distancing*, or limiting use of technology. Distancing behaviors were prevalent in our dataset, used by most populations. The fact that limited or non-participation often felt like the safer choice may perhaps be a troubling signal to technology creators. Here we highlight two common themes.

Censoring online sharing. Some at-risk users carefully self-censored personal content they shared online. This was especially common for users experiencing the *marginalization* risk factor, who often did not feel safe sharing personal information or did not want to reveal stigmatized aspects of their lives to broad online audiences. For example, some LGBTQ+ parents refrained from sharing family or personal photos to try to avoid "outing" themselves beyond specific social circles [16]. In another case, some HIV-positive men seeking men reported not sharing their HIV status on dating apps to avoid unwanted stigmatization [113]. (As noted in the previous section, some members of this population did preemptively disclose their HIV status, demonstrating that protective practices within a population can vary.)

Reducing one's digital footprint. Some at-risk users dramatically reduced technology use to avoid attackers with extensive access, knowledge, and power, as was often the case with the *relationship with the attacker* and *legal or political* risk factors. For example, Chen et al. [19] found that survivors of human trafficking took relatively extreme measures to try to protect their new location, including abandoning devices, deleting social media accounts, and severely restricting online sharing. Survivors of intimate partner abuse employed similar measures when they suspected devices or accounts may have been compromised by their abuser [34]. As another example, activists in the Sudanese Revolution, concerned about the government confiscating their devices, used coded communications and regularly deleted sensitive data [25].

C. Technical solutions

At-risk users also employed technical solutions to protect their digital safety, such as specialized software or settings in common apps and services. Across these strategies, at-risk users did their best to protect themselves based on their knowledge and experience; however, the technical solutions they chose did not always provide the desired protections.

Secure communication and encryption. At-risk users experiencing the *legal or political* risk factor, such as journalists, activists, or politicians, commonly reported using encrypted messaging platforms. These at-risk users frequently wished to secure their communications against monitoring by a nation-state (or similarly resourced) attacker. For example, journalists working together to report on the Panama Papers used PGP as part of organizational security policies established by the ICIJ [69]. Similarly, journalists and activists in varied international contexts reported using encrypted chat apps to communicate over networks directly controlled by their nation-state adversaries [25, 90]. Transgender activists in the U.S., particularly those engaged in more *prominent* activities, similarly used encrypted chat to protect their communications [56]. NGOs, who had *access to other at-risk users*, sought to protect those at-risk users by using encrypted email for internal communications [19].

Beyond encrypted messaging, some at-risk users experiencing the *legal or political* risk factor reported encrypting files or entire devices to protect content from attackers—including nation-states and other focused attackers—who might gain physical access. At-risk users employing this technique included journalists, human rights organizations, and activists [55, 59, 67].

Strong(er) authentication. Strong authentication is commonly advised for internet users in general [80, 85], but can be especially important for users at higher risk of device

or account compromises. While truly strong authentication practices were rare in our dataset, we identified a few authentication trends in subsets of select populations.

While frequently changing passwords is no longer considered a best practice generally [80], some at-risk users did this to address specific practical concerns. For example, at-risk users who have a *relationship with the attacker*, such as survivors of intimate partner abuse or trafficking, coped with attackers who may have ready access to their passwords (e.g., through coercion, co-presence leaks, or remote surveillance [33, 63]) by regularly changing passwords on accounts known to the attacker [19, 33]. Some users with visual impairments also chose to change passwords frequently in case others might have observed them inputting their passwords [5].

While not mentioned frequently in our dataset, our analysis revealed some cases of successful adoption of two-factor authentication (2FA). For example, in Matthews et al. [63], only a small number of survivors of intimate partner abuse in their study reported using 2FA to protect their accounts, despite highly motivated attackers. Only a minority of participants involved with political campaigns reported adopting the strongest form of 2FA, despite commonly experiencing phishing attacks [21]. Journalists collaborating on the Panama Papers (*access to a sensitive resource*) were required by the ICIJ to use 2FA [69], though journalists are more broadly described as having limited awareness of 2FA's benefits [68].[5]

Privacy settings and access control. We identified three categories of privacy settings (within widely used apps and services) that were discussed by multiple papers across a variety of at-risk populations and risk factors: location privacy [19, 28, 36, 63], social media visibility settings [19, 28, 36, 46, 63, 124], and blocking undesired contacts [28, 36, 63, 86, 88]. For users with a *relationship with the attacker* or *access to other at-risk users*, these privacy settings were commonly discussed as protecting highly sensitive information that, if obtained by an attacker, could compromise safety [19, 35, 63]. For example, some women in South Asia used platform controls to block unwanted contact from abusers on social media [88].

For at-risk users in professional settings with *access to a sensitive resource*—such as NGO staff, political campaign workers, and journalists—access control settings were also important for limiting who could access sensitive digital resources [21, 55, 69].

Online identity management. Some at-risk users reported attempting to hide their identity online via careful account management, including multiple and pseudonymous accounts.

Some at-risk users used multiple accounts or devices to maintain boundaries between different facets of their identities. *Marginalization* was one driver for this behavior; for example, sex workers (whose profession was stigmatized) reported keeping separate work and personal accounts [13]. Alternatively, survivors of sexual assault used throwaway accounts to seek support online without revealing their identities [10].

[5]We note that 2FA has become more commonplace since some of these papers were published.

At-risk users also used multiple accounts or devices in response to potential account hijacking or surveillance. Survivors of intimate partner abuse, whose *relationship with the attacker* tended to give the attacker opportunities to access their accounts or devices, reported creating new accounts or purchasing new devices to avoid revictimization after escape [63]. Users experiencing *legal or political* risks, such as activists, reported using multiple accounts and devices to protect themselves from potential nation-state attackers. For example, activists in the Sudanese Revolution reported using SIM cards from other countries, creating fake U.S. phone numbers online, and asking relatives and friends overseas to verify social media accounts in an effort to thwart government surveillance [25]. Marczak et al. found similar evidence of an activist using multiple SIM cards to try to prevent governments from linking calls made in different countries and contexts [60].

In some cases—particularly cases of *marginalization*—rather than use entirely new accounts, at-risk users reported ad-hoc pseudonymity strategies for existing accounts. For example, in order to prevent personal photos being used for digital abuse, some women in South Asia chose neutral images, like flowers, as profile photos [88]. This strategy was shared by some transgender people, particularly activists (who also experienced elevated risk due to *prominence*) [56]. Low-income Black Americans similarly reported sometimes using emojis rather than contact names to protect their contacts' identities on devices they thought might be compromised [31].

Network security. Only two at-risk populations in our review were reported to have used network security tools—like Tor or virtual private networks—to disguise their web traffic. Certain activists [25, 59] and journalists [67] were aware of and used these technologies, but these tools did not appear elsewhere in our dataset—and not every user in those populations found these tools equally useful or usable [59, 68]. This could mean that other at-risk populations are largely unaware of these tools, that existing tools are not perceived as useful for the particular digital-safety concerns of other at-risk populations, or simply that researchers have not fully investigated this question with other at-risk populations.

Tracking and monitoring applications. Caregivers (of at-risk users who *rely on a third party*) sometimes used tracking or monitoring applications with the goal of protecting the at-risk user's digital safety. Parents of foster teens reported using router settings to limit internet access at certain times of day and parental control apps to watch for potentially dangerous behavior [12]. These applications can create conflict between autonomy and privacy, but some children and teenagers indicated that they can be helpful, particularly when used as part of an ongoing dialogue with parents [37, 39].

VI. Barriers to Protective Practices

In the previous section, we documented a range of protective practices at-risk users employed, motivated by the risk factors they experienced. In our thematic analysis, we also identified a

variety of barriers that limited or prevented at-risk users from effectively adopting these practices. We apply the contextual risk factors and protective practices of our at-risk framework to discuss three categories of barriers: *competing priorities*, *lack of knowledge or experience*, and *broken technology assumptions*.

A. Competing priorities

Digital safety is not and cannot always be the top priority for all users; this is perhaps even more true for at-risk users with competing, often critical, needs. Some of the many competing priorities that appeared in our dataset included basic needs like food, income, or physical health and safety; social participation and compliance with social norms; and caring for others. We found that specific competing priorities were often associated with particular risk factors.

It is well understood that in general, users prioritize convenience, simplicity, and their tasks and goals over digital safety [1, 44]. Our analysis revealed a tendency for at-risk users to have layered or more severe conflicts between digital safety and other needs, making the leap to safer behaviors especially difficult.

Basic needs. People who are *resource or time constrained* in our dataset often prioritized other critical needs over digital safety, despite the potential for increased digital-safety risk. For example, Elliott and Brody [31] reported that low-income Black New Yorkers used apps to find cheaper food, despite suspecting the apps were insecure. Similarly, people experiencing homelessness described discomfort using public Wi-Fi, but used it anyway for critical needs such as applying for government assistance, housing, and jobs [97].

Relatedly, people who *rely on a third party* often gave up control of digital safety to accomplish basic tasks. Refugees, for example, shared account information, including passwords, with caseworkers to obtain social services or apply for jobs [96]. Some people who are visually impaired used crowdsourced assistive technology for tasks like identifying medicines correctly, despite the risk of exposing sensitive content in the background of the photos crowdworkers would evaluate [7]. Frik et al. [36] found that some older adults were willing to accept in-home surveillance, giving up privacy to maintain some autonomy and independent living.

Participation and connection. As noted in SectionV-B above, people who experience *marginalization* sometimes chose to distance or fully disconnect in order to keep themselves safe. Members of these populations who opted to engage were often reported as knowing that it increased their risk. For example, Blackwell et al. reported on LGBT+ parents implicitly "outing" themselves via "everyday" social media posts about their children and families [16].

We noted similar behavior related to *legal or political* risk, where the need to communicate with other activists, both to organize events and simply to be part of a community, motivated potentially risky modes of communication [90].

Similarly, maintaining good social standing in a family or community sometimes required people experiencing the *social norms* factor to deprioritize digital safety. Our dataset included, for example, women in South Asia and Saudi Arabia, as well as people in South Africa, sharing accounts, devices, and credentials with family members to meet cultural or community expectations [8, 86, 88, 89]. Some users in South Africa accepted Facebook friend requests from strangers, even though they were uncomfortable with the resultant potential for sharing, to avoid being rude [86]. Teenagers noted that they did not discuss or share risky digital-safety experiences with their parents to avoid awkwardness or a possible negative reaction [119].

Caring for others. Another example of competing priorities involved at-risk users taking on additional risk to help, care for, or support others. We noted this occurring frequently in connection with the *marginalization* factor, in part to reduce overall stigma via normalization. For example, Warner et al. found that some gay men disclosed their HIV+ status on a dating app in part to increase visibility [113]; similarly, separate studies found that LGBTQ+ parents and transgender individuals posted on social media, despite possible risks, in part to increase visibility and reassure others that they were not alone [16, 56].

Our analysis revealed cases in which at-risk users prioritized others' autonomy over their own digital safety, particularly in the case of *access to other at-risk users*. McGregor et al. [68] quoted a journalist who valued avoiding "impos[ing] any kind of burden on a source" over the journalist's own digital safety. Similarly, NGO staff with *access to other at-risk users* reported the need to balance security with the autonomy of the survivors they worked with: Chen et al. [19] quoted a staff member working with trafficking survivors who tended to suggest that they turn off phone location, "But I do kind of leave it up to them. It's not mandatory.... 'Cause we come from an empowering place. We don't want to be telling them what to do."

Caring for others can also lead at-risk users to prioritize efficiency in achieving their primary goals over digital safety, something we noted particularly in the cases of *legal or political*, *prominence*, or *access to a sensitive resource*, which frequently coincided with a profession or activity that benefited others. A human rights activist, for example, wondered, "Should I spend half a day figuring out digital security, or do work?" [59]. In several studies, journalists, political campaign workers, and emergency department personnel reported mixing personal and work data and devices for efficiency in their work [21, 68, 90, 99].

B. Lack of knowledge or experience

In our dataset, nearly all at-risk populations had less digital-safety knowledge than they needed to mitigate the risks they faced. While prior work has found that general users tend to have limited digital-safety knowledge [116], this was exacerbated for at-risk users who, by definition, faced serious digital-safety risks and thus usually needed to understand and

deploy more robust protections than a typical user. The at-risk framework helps us understand current patterns in how at-risk users experience digital-safety knowledge limitations.

***Legal or political* and *prominence* tended to attract sophisticated attackers, leading to stark knowledge asymmetries.** It was difficult for at-risk populations—like activists, undocumented immigrants, journalists, or political campaign workers—to surmount the knowledge differential needed to counter the nation-state attackers that might or did target them [21, 25, 67, 68, 101]. For example, undocumented immigrants in the U.S. were often unaware that Immigration and Customs Enforcement might target them for surveillance via government requests for information to the social media platforms they used [40].

***Relationship with the attacker* involved intimate threats that required expertise to counter.** Even in the absence of sophisticated attacks, these threats can require substantial, robust protective measures. For example, survivors of intimate partner abuse contended with a motivated attacker who may launch repeated attacks leveraging intimate knowledge about them, physical access to their devices and accounts, and relational power [21, 34, 57].

Some *resource or time constrained* users had limited technology experience. Examples included some people with low SES, people living in developing regions, and survivors of intimate partner abuse or trafficking, who did not have regular access to new or trusted devices or the internet, which limited their ability to gain technology experience and skills [19, 46, 63, 97]. Other populations, like journalists and people involved with political campaigns, commonly did not have the time to develop the technical skills to counter the digital-safety risks they faced [21, 67, 68]. Several studies reported that at-risk users who are *resource or time constrained* (emergency department workers [99], older adults [36, 45], refugees [96], and others) did not understand digital-safety settings that were, in theory, available to them.

C. Broken technology assumptions

The atypical threat models at-risk users face sometimes break assumptions that are built into secure system designs. Because of these assumptions, digital-safety best practices and technologies may be inaccessible, non-functional, or only minimally useful to at-risk users, often with magnified consequences.

One person per device or account. A common assumption of technology creators is that for every given device and account, there will be exactly one user who has access, despite prior work showing that convenience-based device sharing between trusted parties is common [62].[6] This core assumption fails for several at-risk populations. At-risk users who *rely on a third party* may share account information with these trusted parties in order to accomplish important tasks [53, 96] or receive needed monitoring [76]. Different cultural privacy models have resulted in certain users sharing devices and accounts due to *social norms* [4, 89] or even the *legal or political* requirements of where they live [8]. Additionally, at-risk users who have a *relationship with the attacker* were frequently forced to give these attackers access to their devices under duress, directly breaking this assumption [18, 19, 33, 35, 42, 63].

Everyone has sufficient technology access. Some security techniques assume minimum levels of technology access that are out of reach for some at-risk users who are *resource constrained*. For example, 2FA that depends on a mobile phone (e.g., via SMS or a code-generating app) may not be available to at-risk users experiencing homelessness who are unable to consistently pay a phone bill or have an old device without space for new apps [97]. Some low-income Black Americans reported needing to stay on family mobile plans—which could have enabled surveillance from untrusted relations—because they could not afford separate service [31]. U.S. sanctions on Sudan entirely prevented Sudanese activists from enabling 2FA on a particular social media platform, since that platform's 2FA system was prohibited from recognizing Sudanese phone numbers [25].

Physical and cognitive capacities are universal. Even for general users, advice about authentication tends to assume impossible cognitive capabilities, such as memorizing a unique, strong password for every account [80, 107]. This mistaken assumption is compounded in cases of *underserved accessibility needs*. For example, several papers have identified password memorization as a critical challenge for older adults with cognitive impairments [36, 76]. Password memorization can be difficult for people with disabilities that relate to alphanumeric comprehension, like dyslexia or aphasia [61].

Concepts, values, and experiences are universal. Other digital-safety paradigms rest on ideas, definitions, values, and morals that may not apply universally. Shortcomings in these assumptions can create digital-safety risks, especially for at-risk users with different *social norms*. For example, translations from English to Khmer of concepts like *privacy* within social media settings were hard to understand in the strongly community-oriented culture in Cambodia [46]. Similarly, some refugees who immigrated to the U.S. came from cultures where birthdays were not recorded; when these users were assigned a default value of January 1 in the U.S., authentication systems that relied on knowledge or entropic distribution of birthdays were less effective [96]. Separately, Barwulor et al. [13] found that moral codes and laws enacted by the U.S. government and enforced by U.S. companies made it difficult for sex workers in other countries, where their work is legal, to access safe payment and advertisement platforms, placing their physical and digital safety at risk.

Limiting digital-safety options is good for everyone. Digital-safety options that are too numerous can be hard to use [98].

[6] We note that since many of the papers in our dataset were published, there has been progress in addressing this issue, particularly in supporting child and family accounts. Nonetheless, we believe more can still be done to disrupt this assumption.

But limiting the nuanced control enabled by digital-safety options to meet the usability needs of typical users may not always work well for at-risk users. Risk factors that increased the chances of focused targeting—which included *prominence*, *relationship with the attacker*, *reliance on a third party*, and *access to other at-risk users*—can lead at-risk users to have highly contextual digital-safety needs. For example, users protecting against a focused, intimate attacker, must account for nuances in the current relationship, the attacker's mood, whether or not they are physically copresent, and more [57, 63]. Because of this nuance, at-risk users may benefit from additional options or enhanced transparency that they can deploy in specific alignment with their goals. When these options are not easily available or understood, it may lead at-risk users to fall back on *distancing behaviors* in which they try to stay safe at the cost of fully engaging with technology [25, 34, 36, 40, 56, 67].

VII. Implications and Future Directions

The at-risk framework can be used in multiple ways by researchers and technology creators, including guiding research and developing technologies to be inclusive of at-risk users.

A. Research

The framework, as applied to our dataset in Table I, can be used to help identify where knowledge of at-risk users is underdeveloped, sparse, or missing, giving researchers a way to prioritize their efforts. The framework can also be used to guide development of study designs and research questions.

Identify the *who* and the *what*. The digital-safety community would benefit from research about all at-risk populations and factors—this includes expanding knowledge about those that have already been studied and creating new knowledge about those that have not. In this complex and nuanced space, even after several papers have reported on a population, researchers may still have much to learn about the populations' digital-safety needs.

Nonetheless, our meta-analysis makes clear that certain at-risk populations and risk factors have received particularly limited attention, at least recently and within the digital-safety community. We observed a general tendency to study participants from Western cultures, especially the U.S. Studying people from other regions and cultures could shed light on new risks related to *social norms*, un- or understudied interactions among risk factors, or even new risk factors. We advocate for more research involving geographically and culturally diverse at-risk participants, conducted by researchers who represent the world.

We also saw multiple studies exploring populations that experienced *marginalization*, focusing on a fairly narrow set of risk experiences (i.e., some populations' only black circle (●) in Table I was for the *marginalization* factor); future work could expand our understanding of digital-safety risks for these populations. Other groups often referenced as at-risk in popular media (e.g., real estate agents who have *access to sensitive resources* [106]; celebrities who face digital-safety risks due to *prominence* and parasocial relationships [121]) have not been studied in the research venues we analyzed. New or additional foundational research may be needed to inform the digital-safety community about these populations' experiences and needs.

Our framework also highlights how risk factors often combine and interact, but the literature in our dataset has not deeply explored this topic, precluding a thorough synthesis. We advocate *interactions* as a critical area of future research—one which our framework can support (e.g., enumerating the risk factors to consider for interactions).

Finally, in crisis situations, such as natural disasters or war, people may lose access to essential resources, their circumstances may change, or they may change their behaviors in exchange for critical services in ways that amplify their risk of attack or the severity of resulting harms. Recent papers on the impacts of the COVID-19 pandemic provide evidence that such changes can create new risks and amplify existing risks (e.g., [105, 123]). Future work on digital safety during various kinds of crisis situations could yield a new risk factor, or expand the definitions of existing factors.

Guide the *how*. The framework can also be used to help shape study designs and reporting, particularly interview or survey questions. For a population about which little is known, asking about all 10 contextual risk factors can ensure fairly comprehensive coverage of digital-safety concerns. For populations where some research exists, researchers can use the framework to explore how previously un- or understudied risk factors may (or may not) apply, or add depth on the impact of a specific, previously identified risk factor of interest. Researchers can also use the protective practices portion of the framework to explore more comprehensively how their participants currently protect themselves and why they choose their practices. We hope that using the framework to guide research and reporting can enable better comparisons among studies, helping to uncover when disparate populations have overlapping (or distinctive) digital-safety practices and needs. Beyond this, it is important for researchers to carefully plan ethical methods when working with at-risk users, guidance for which is an emerging area of research [23, 58].

B. Technology development

Technology creators can use the at-risk framework to better support a wide range of at-risk users in their products.

Consider at-risk users at scale. Our framework does not replace direct engagement with at-risk users, and we advocate for such engagement when appropriate. However, doing so selectively and ethically is important, and there are still open questions about ethical methods for at-risk user research (e.g., how do researchers not overtax already stressed and resource-constrained groups?). Meanwhile, many papers in our dataset recommended considering the impact of a technology design on the at-risk population they studied, which is incredibly important but difficult to scale across populations without a guiding framework. Our framework simplifies the challenging

but important process of thinking through potential risks and needs of multiple at-risk populations together. It does this by providing 10 contextual risk factors that organize patterns of risks and needs, and a set of protective practices at-risk users currently deploy to (sometimes ineffectively) cope. Using our framework can help researchers and technology creators prepare for user research, at interim points during multi-phased technology creation projects, or when user research across multiple at-risk populations is not an ethical option.

Each risk factor suggests specific, sometimes overlapping, technology needs. For *social norms*, *relationship with the attacker*, and *reliance on a third party*, users often could not keep their devices and accounts private. These users might benefit from the ability to keep select technology use and data secret on shared devices and accounts, or perhaps to enable digital traces that are ambiguous or imprecise to support plausible deniability [11]. Users with *access to a sensitive resource* would benefit from robust protections for the sensitive resource (e.g., encryption, strong authentication). Users facing focused and/or sophisticated attackers (e.g., *legal or political*, *prominence*, *relationship with the attacker*, *access to other at-risk users*) could benefit from easier-to-use versions of strong protections (such as hardware-based 2FA, strong passwords, and encryption) coupled with guided set-up flows and education. Users with *prominence* or who experience *marginalization* would benefit from support for managing bulk or pervasive attacks from potentially anyone.

Our systematization of protective practices (Section V) and barriers (Section VI) together highlight how at-risk users cope with their risks, sometimes in ways that are not completely effective or that introduce vulnerabilities. Notably, at-risk users commonly employed non-technical practices—such as a host of social strategies and distancing behaviors—and it is important for technology creators to understand and not disrupt the important role these practices play. The protective practices we identified also show that at-risk users are not commonly using some existing digital-safety solutions (at least as reported in the dataset), suggesting areas where technology creators could improve accessibility and/or usability. Further, we discuss broken assumptions that contribute to ineffective protections (Section VI-C), and encourage technology creators to consider these in new technologies.

Balance tensions. Those creating technology for at-risk users will have to contend with inherent tensions: "perfect" digital-safety protections usually do not exist. At-risk users already use a variety of practices, technical and otherwise, to address their pressing digital-safety concerns (Section V). These practices all have some protective value, but may also come with significant downsides, like reduced social participation, lack of agency, and loss of transparency. For example, *distancing* from technology (Section V-B) was a common protective practice across our dataset, but it also reduced access to *social support*, which was another essential protective practice broadly employed by at-risk users (Section V-A). Technology creators should understand that any technical intervention is likely to introduce benefits and drawbacks, which should be carefully studied to help ensure such tensions are understood, manageable, and net beneficial. Careful evaluation of potential designs using our framework can help technology teams reason about how to balance various risks and benefits.

Balance usability and options. In Section VI-C, we discussed at-risk users' need for more nuanced digital safety options than typical users. At the same time, adding options and transparency must be balanced with a very real need for usability. At-risk users experience heightened stress associated with higher potential for digital harm, as well as stress from their particular risk factors (e.g., *resource constraints*, *marginalization*, etc.).

To balance the need for options that addresses the unique needs of at-risk users with the high bar for usability, we encourage practitioners to make transparency and controls actionable and manageable. For example, transparency features that make users aware of a threat can be more actionable if they provide clear next steps to fix the issue and then guide the user through relevant protective measures. Similarly, layered or directed designs can help users find the options and controls that best meet their needs. Equally important are easy-to-understand defaults that minimize barriers to deploying protections and are carefully selected to support at-risk users with many competing priorities (Section VI-A).

VIII. Conclusion

Over the past several years, a growing body of research has focused on digital-safety risks for at-risk users; however, guidance drawn from varied populations can be difficult for researchers and technology creators to apply in practice. To make this more tractable, we systematically analyzed 95 papers focused on varied populations and created an *at-risk framework*: 10 *contextual risk factors* that can augment or amplify common, high-priority digital-safety risks and their resulting harms, and the *protective practices* at-risk users employ to mitigate these risks. We used our framework to discuss *barriers* at-risk users face enacting digital protections. Going forward, our framework can be used to identify opportunities for future research and to provide a structure for researchers and technology creators to scalably and more comprehensively ensure that everyone—including at-risk users—can engage safely online.

IX. Acknowledgements

We thank our colleagues and experts who provided feedback on this work, including Allison McDonald, Andrew Botros, Beng Lim, Emerson Murphy-Hill, Florian Schaub, Franziska Roesner, Jill Palzkill Woelfer, Josh Lovejoy, Nafis Zebarjadi, Patrawat Samermit, Reena Jana, Shaun Kane, Stephan Somogyi, and Tu Tsao. This material is based upon work supported by DARPA under grant HR00112010011. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Government or DARPA. Approved for public release; distribution is unlimited.

References

[1] Anne Adams and Martina Angela Sasse. Users are not the enemy. *Communications of the ACM*, 42(12):40–46, 1999.

[2] Syed Ishtiaque Ahmed, Md. Romael Haque, Jay Chen, and Nicola Dell. Digital privacy challenges with shared mobile phone use in Bangladesh. *PACM HCI*, 1(CSCW):1–20, 2017.

[3] Syed Ishtiaque Ahmed, Md. Romael Haque, Shion Guha, Md. Rashidujjaman Rifat, and Nicola Dell. Privacy, security, and surveillance in the global south: A study of biometric mobile SIM registration in Bangladesh. In *Proc. CHI*, 2017.

[4] Syed Ishtiaque Ahmed, Md. Romael Haque, Irtaza Haider, Jay Chen, and Nicola Dell. "Everyone has some personal stuff": Designing to support digital privacy with shared mobile phone use in Bangladesh. In *Proc. CHI*, 2019.

[5] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. Privacy concerns and behaviors of people with visual impairments. In *Proc. CHI*, 2015.

[6] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. Addressing physical safety, security, and privacy for people with visual impairments. In *Proc. SOUPS*, 2016.

[7] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. "I am uncomfortable sharing what I can't see": Privacy concerns of the visually impaired with camera based assistive applications. In *Proc. USENIX Security*, 2020.

[8] Deena Alghamdi, Ivan Flechais, and Marina Jirotka. Security practices for households bank customers in the Kingdom of Saudi Arabia. In *Proc. SOUPS*, 2015.

[9] Adriana Alvarado Garcia, Alyson L Young, and Lynn Dombrowski. On making data actionable: How activists use imperfect data to foster social change for human rights violations in Mexico. *PACM HCI*, 1(CSCW):1–19, 2017.

[10] Nazanin Andalibi, Oliver L. Haimson, Munmun De Choudhury, and Andrea Forte. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In *Proc. CHI*, 2016.

[11] Paul M. Aoki and Allison Woodruff. Making space for stories: ambiguity in the design of personal communication systems. In *Proc. CHI*, 2005.

[12] Karla Badillo-Urquiola, Xinru Page, and Pamela J. Wisniewski. Risk vs. restriction: The tension between providing a sense of normalcy and keeping foster teens safe online. In *Proc. CHI*, 2019.

[13] Catherine Barwulor, Allison McDonald, Eszter Hargittai, and Elissa M Redmiles. "Disadvantaged in the american-dominated internet": Sex, work, and technology. In *Proc. CHI*, 2021.

[14] Clara Berridge, Jodi Halpern, and Karen Levy. Cameras on beds: The ethics of surveillance in nursing home rooms. *AJOB Empirical Bioethics*, 10(1):55–62, 2019.

[15] Lindsay Blackwell, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe. Classification and its consequences for online harassment: Design insights from HeartMob. In *PACM HCI*, 2017.

[16] Lindsay Blackwell, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck. LGBT parents and social media: Advocacy, privacy, and disclosure during shifting social movements. In *Proc. CHI*, 2016.

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

[18] Rahul Chatterjee, Periwinkle Doerfler, Hadas Orgad, Sam Havron, Jackeline Palmer, Diana Freed, Karen Levy, Nicola Dell, Damon McCoy, and Thomas Ristenpart. The spyware used in intimate partner violence. In *Proc. IEEE S&P*, 2018.

[19] Christine Chen, Nicola Dell, and Franziska Roesner. Computer security and privacy in the interactions between victim service providers and human trafficking survivors. In *Proc. USENIX Security*, 2019.

[20] Taejoong Chung, Jinyoung Han, Daejin Choi, Ted Taekyoung Kwon, Jong-Youn Rha, and Hyunchul Kim. Privacy leakage in event-based social networks: A Meetup case study. *PACM HCI*, 1(CSCW):1–22, 2017.

[21] Sunny Consolvo, Patrick Gage Kelley, Tara Matthews, Kurt Thomas, Lee Dunn, and Elie Bursztein. "Why wouldn't someone think of democracy as a target?": Security practices & challenges of people involved with U.S. political campaigns. In *Proc. USENIX Security*, 2021.

[22] Raymundo Cornejo, Robin Brewer, Caroline Edasis, and Anne Marie Piper. Vulnerability, sharing, and privacy: Analyzing art therapy for older adults with dementia. In *Proc. CSCW, 2016*, 2016.

[23] Sasha Costanza-Chock. Design justice: Towards an intersectional feminist framework for design theory and practice. *Proceedings of the Design Research Society*, 2018.

[24] Kimberlé Crenshaw. Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, pages 139–168, 1989.

[25] Alaa Daffalla, Lucy Simko, Tadayoshi Kohno, and Alexandru G Bardas. Defensive technology use by political activists during the Sudanese revolution. In *Proc. IEEE S&P*, 2021.

[26] Keith E. Davis and Irene H. Frieze. Research on stalking: What do we know and where do we go? *Violence and Victims*, 15(4):473–487, 2000.

[27] Steve Dent. Security firm details how hackers stole $1.3 million in wire transfers. https://www.engadget.com/hackers-steal-1-3-million-wire-transfer-100039219.html, 2020.

[28] Jayati Dev, Pablo Moriano, and L Jean Camp. Lessons learnt from comparing WhatsApp privacy concerns across Saudi and Indian populations. In *Proc. SOUPS*, 2020.

[29] Paul Dunphy, John Vines, Lizzie Coles-Kemp, Rachel Clarke, Vasilis Vlachokyriakos, Peter Wright, John McCarthy, and Patrick Olivier. Understanding the experience-centeredness of privacy and security technologies. In *Proc. NSPW*, 2014.

[30] Kasra EdalatNejad, Wouter Lueks, Julien Pierre Martin, Soline Ledésert, Anne L'Hôte, Bruno Thomas, Laurent Girod, and Carmela Troncoso. DatashareNetwork: A decentralized privacy-preserving search engine for investigative journalists. In *Proc. USENIX Security*, 2020.

[31] Ame Elliott and Sara Brody. Straight talk: New Yorkers on mobile messaging and implications for privacy. Technical report, Simply Secure, 2015.

[32] Valerie Fanelle, Sepideh Karimi, Aditi Shah, Bharath Subramanian, and Sauvik Das. Blind and human: Exploring more usable audio CAPTCHA designs. In *Proc. SOUPS*, 2020.

[33] Diana Freed, Sam Havron, Emily Tseng, Andrea Gallardo, Rahul Chatterjee, Thomas Ristenpart, and Nicola Dell. "Is my phone hacked?" analyzing clinical computer security interventions with survivors of intimate partner violence. *PACM HCI*, 3(CSCW):1–24, 2019.

[34] Diana Freed, Jackeline Palmer, Diana Minchala, Karen Levy, Thomas Ristenpart, and Nicola Dell. Digital technologies and intimate partner violence: A qualitative analysis with multiple stakeholders. *PACM HCI*, 1(CSCW):1–22, 2017.

[35] Diana Freed, Jackeline Palmer, Diana Minchala, Karen Levy, Thomas Ristenpart, and Nicola Dell. "A Stalker's Paradise": How intimate partner abusers exploit technology. In *Proc. CHI*, 2018.

[36] Alisa Frik, Leysan Nurgalieva, Julia Bernd, Joyce Lee, Florian Schaub, and Serge Egelman. Privacy and security threat models and mitigation strategies of older adults. In *Proc. SOUPS*, 2019.

[37] Arup Kumar Ghosh, Karla Badillo-Urquiola, Shion Guha, Joseph J. LaViola Jr, and Pamela J. Wisniewski. Safety vs. surveillance: What children have to say about mobile apps for parental control. In *Proc. CHI*, 2018.

[38] Arup Kumar Ghosh, Karla Badillo-Urquiola, Mary Beth Rosson, Heng Xu, John M Carroll, and Pamela J. Wisniewski. Matter of control or safety? Examining parental use of technical monitoring apps on teens' mobile devices. In *Proc. CHI*, 2018.

[39] Arup Kumar Ghosh, Charles E. Hughes, and Pamela J. Wisniewski. Circle of trust: A new approach to mobile online safety for families. In *Proc. CHI*, 2020.

[40] Tamy Guberek, Allison McDonald, Sylvia Simioni, Abraham H. Mhaidli, Kentaro Toyama, and Florian Schaub. Keeping a low profile?: Technology, risk and privacy among undocumented immigrants. In *Proc. CHI*, 2018.

[41] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. Gender recognition or gender reductionism? The social implications of embedded gender recognition systems. In *Proc. CHI*, 2018.

[42] Sam Havron, Diana Freed, Rahul Chatterjee, Damon McCoy, Nicola Dell, and Thomas Ristenpart. Clinical computer security for victims of intimate partner violence. In *Proc. USENIX Security*, 2019.

[43] Jordan Hayes, Smirity Kaushik, Charlotte Emily Price, and Yang Wang. Cooperative privacy and security: Learning from people with visual impairments and their allies. In *Proc. SOUPS*, 2019.

[44] Cormac Herley. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proc. NSPW*, 2009.

[45] Dominik Hornung, Claudia Müller, Irina Shklovski, Timo Jakobi, and Volker Wulf. Navigating relationships and boundaries: Concerns around ICT-uptake for elderly people. In *Proc. CHI*, 2017.

[46] Margaret C Jack, Pang Sovannaroth, and Nicola Dell. "Privacy is not a concept, but a way of dealing with life": Localization of transnational technology platforms and liminal privacy practices in Cambodia. *PACM HCI*, 3(CSCW):1–19, 2019.

[47] Rebecca Jeong and Sonia Chiasson. 'Lime', 'open lock', and 'blocked': Children's perception of colors, symbols, and words in cybersecurity warnings. In *Proc. CHI*, 2020.

[48] Os Keyes, Josephine Hoy, and Margaret Drouhard. Human-computer insurrection: Notes on an anarchist HCI. In *Proc. CHI*, 2019.

[49] Yong Ming Kow, Yubo Kou, Bryan Semaan, and Waikuen Cheng. Mediating the undercurrents: Using social media to sustain a social movement. In *Proc. CHI*, 2016.

[50] Sandjar Kozubaev, Fernando Rochaix, Carl DiSalvo, and Christopher A. Le Dantec. Spaces and traces: Implications of smart technology in public housing. In *Proc. CHI*, 2019.

[51] Klaus Krippendorff. Testing the reliability of content analysis data: What is involved and why. In Klaus Krippendorff and Mary Angela Bock, editors, *The Content Analysis Reader*, chapter 6.2, pages 350–357. SAGE, 2009.

[52] Priya Kumar, Shalmali Milind Naik, Utkarsha Ramesh Devkar, Marshini Chetty, Tamara L. Clegg, and Jessica Vitak. 'No telling passcodes out because they're private': Understanding children's mental models of privacy and security online. *PACM HCI*, 1(CSCW):1–21, 2017.

[53] Priya C. Kumar, Marshini Chetty, Tamara L. Clegg, and Jessica Vitak. Privacy and security considerations for digital technology use in elementary schools. In *Proc. CHI*, 2019.

[54] Elmer Lastdrager, Inés Carvajal Gallardo, Pieter Hartel, and Marianne Junger. How effective is anti-phishing training for children? In *Proc. SOUPS*, 2017.

[55] Stevens Le Blond, Alejandro Cuevas, Juan Ramón Troncoso-Pastoriza, Philipp Jovanovic, Bryan Ford, and Jean-Pierre Hubaux. On enforcing the digital immunity of a large humanitarian organization. In *Proc. IEEE S&P*, 2018.

[56] Ada Lerner, Helen Yuxun He, Anna Kawakami, Silvia Catherine Zeamer, and Roberto Hoyle. Privacy and activism in the transgender community. In *Proc. CHI*, 2020.

[57] Karen Levy and Bruce Schneier. Privacy threats in intimate relationships. *Journal of Cybersecurity*, 6(1):tyaa006, 2020.

[58] Calvin A. Liang, Sean A. Munson, and Julie A. Kientz. Embracing four tensions in human-computer interaction research with marginalized people. *ACM TOCHI*, 28(2):1–47, 2021.

[59] William R. Marczak and Vern Paxson. Social engineering attacks on government opponents: Target perspectives. In *Proc. PETS*, 2017.

[60] William R. Marczak, John Scott-Railton, Morgan Marquis-Boire, and Vern Paxson. When governments hack opponents: A look at actors and technology. In *Proc. USENIX Security*, 2014.

[61] Sonali Tukaram Marne, Mahdi Nasrullah Al-Ameen, and Matthew K. Wright. Learning system-assigned passwords: A preliminary study on the people with learning disabilities. In *Proc. SOUPS*, 2017.

[62] Tara Matthews, Kerwell Liao, Anna Turner, Marianne Berkovich, Robert Reeder, and Sunny Consolvo. "She'll just grab any device that's closer": A study of everyday device & account sharing in households. In *Proc. CHI*, 2016.

[63] Tara Matthews, Kathleen O'Leary, Anna Turner, Manya Sleeper, Jill Palzkill Woelfer, Martin Shelton, Cori Manthorne, Elizabeth F Churchill, and Sunny Consolvo. Stories from survivors: Privacy & security practices when coping with intimate partner abuse. In *Proc. CHI*, 2017.

[64] Leslie McCall. The complexity of intersectionality. *Signs: Journal of Women in Culture and Society*, 30(3):1771—-1800, 2005.

[65] Nora McDonald, Karla Badillo-Urquiola, Morgan G Ames, Nicola Dell, Elizabeth Keneski, Manya Sleeper, and Pamela J Wisniewski. Privacy and power: Acknowledging the importance of privacy research and design for vulnerable populations. In *Proc. CHI Extended Abstracts*, 2020.

[66] Nora McDonald, Alison Larsen, Allison Battisti, Galina Madjaroff, Aaron Massey, and Helena Mentis. Realizing choice: Online safeguards for couples adapting to cognitive challenges. In *Proc. SOUPS*, 2020.

[67] Susan E. McGregor, Polina Charters, Tobin Holliday, and Franziska Roesner. Investigating the computer security practices and needs of journalists. In *Proc. USENIX Security*, 2015.

[68] Susan E. McGregor, Franziska Roesner, and Kelly Caine. Individual versus organizational computer security and privacy concerns in journalism. In *Proc. SOUPS*, 2016.

[69] Susan E. McGregor, Elizabeth Anne Watkins, Mahdi Nasrullah Al-Ameen, Kelly Caine, and Franziska Roesner. When the weakest link is strong: Secure collaboration in the case of the Panama Papers. In *Proc. USENIX Security*, 2017.

[70] Susan E. McGregor, Elizabeth Anne Watkins, and Kelly Caine. Would you Slack that? The impact of security and privacy on cooperative newsroom work. *PACM HCI*, 1(CSCW):1–22, 2017.

[71] Bridget Christine McHugh, Pamela J Wisniewski, Mary Beth Rosson, Heng Xu, and John M Carroll. Most teens bounce back: Using diary methods to examine how quickly teens recover from episodic online risk exposure. *PACM HCI*, 1(CSCW):1–19, 2017.

[72] Brenna McNally, Priya Kumar, Chelsea Hordatt, Matthew Louis Mauriello, Shalmali Naik, Leyla Norooz, Alazandra Shorter, Evan Golub, and Allison Druin. Co-designing mobile online safety applications with children. In *Proc. CHI*, 2018.

[73] Andrew R. McNeill, Lynne Coventry, Jake Pywell, and Pam Briggs. Privacy considerations when designing social network systems to support successful ageing. In *Proc. CHI*, 2017.

[74] Emily McReynolds, Sarah Hubbard, Timothy Lau, Aditya Saraf, Maya Cakmak, and Franziska Roesner. Toys that listen: A study of parents, children, and internet-connected toys. In *Proc. CHI*, 2017.

[75] Hamid Mehmood, Tallal Ahmad, Lubna Razaq, Shrirang Mare, Maryem Zafar Usmani, Richard Anderson, and Agha Ali Raza. Towards digitization of collaborative savings among low-income groups. *PACM HCI*, 3(CSCW):1–30, 2019.

[76] Helena M Mentis, Galina Madjaroff, and Aaron K Massey. Upside and downside risk in online security for older adults with mild cognitive impairment. In *Proc. CHI*, 2019.

[77] Carol Moser, Tianying Chen, and Sarita Y. Schoenebeck. Parents' and children's preferences about parents sharing about children on social media. In *Proc. CHI*, 2017.

[78] James Nicholson, Lynne Coventry, and Pamela Briggs. "If it's important it will be a headline": Cybersecurity information seeking in older adults. In *Proc. CHI*, 2019.

[79] Borke Obada-Obieh, Lucrezia Spagnolo, and Konstantin Beznosov. Towards understanding privacy and trust in online reporting of sexual assault. In *Proc. SOUPS*, 2020.

[80] National Institute of Standards and Technology. Digital identity guidelines. Technical Report National Institute of Standards and Technology Special Publication 800-63-3, U.S. Department of Commerce, Washington, D.C., 2020.

[81] Justin Petelka, Lucy Van Kleunen, Liam Albright, Elizabeth Murnane, Stephen Voida, and Jaime Snyder. Being (in) visible: Privacy, transparency, and disclosure in the self-management of bipolar disorder. In *Proc. CHI*, 2020.

[82] John R. Porter, Kiley Sobel, Sarah E Fox, Cynthia L Bennett, and Julie A Kientz. Filtered out: Disability disclosure practices in online dating communities. *PACM HCI*, 1(CSCW):1–13, 2017.

[83] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. How I learned to be secure: a census-representative survey of security advice sources and behavior. In *Proc. ACM CCS*, 2016.

[84] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. Where is the digital divide? a survey of security, privacy, and socioeconomics. In *Proc. CHI*, 2017.

[85] Elissa M. Redmiles, Noel Warford, Amritha Jayanti, Aravind Koneru, Sean Kross, Miraida Morales, Rock Stevens, and Michelle L. Mazurek. A comprehensive quality evaluation of security and privacy advice on the web. In *Proc. USENIX Security*, 2020.

[86] Jake Reichel, Fleming Peck, Mikako Inaba, Bisrat Moges, Brahmnoor Singh Chawla, and Marshini Chetty. 'I have too much respect for my elders': Understanding South African mobile users' perceptions of privacy and current behaviors on Facebook and WhatsApp. In *Proc. USENIX Security*, 2020.

[87] Sabirat Rubya and Svetlana Yarosh. Interpretations of online anonymity in Alcoholics Anonymous and Narcotics Anonymous. *PACM HCI*, 1(CSCW):1–22, 2017.

[88] Nithya Sambasivan, Amna Batool, Nova Ahmed, Tara Matthews, Kurt Thomas, Laura Sanely Gaytán-Lugo, David Nemer, Elie Bursztein, Elizabeth Churchill, and Sunny Consolvo. "They don't leave us alone anywhere we go": Gender and digital abuse in South Asia. In *Proc. CHI*, 2019.

[89] Nithya Sambasivan, Garen Checkley, Amna Batool, Nova Ahmed, David Nemer, Laura Sanely Gaytán-Lugo, Tara Matthews, Sunny Consolvo, and Elizabeth Churchill. "Privacy is not for me, it's for those rich women": Performative privacy practices on mobile phones by women in South Asia. In *Proc. SOUPS*, 2018.

[90] Pedro Sanches, Vasiliki Tsaknaki, Asreen Rostami, and Barry Brown. Under surveillance: Technology practices of those monitored by the state. In *Proc. CHI*, 2020.

[91] Morgan Klaus Scheuerman, Stacy M. Branham, and Foad Hamidi. Safe spaces and safe places: Unpacking technology-mediated experiences of safety and harm with transgender people. *PACM HCI*, 2(CSCW):1–27, 2018.

[92] Morgan Klaus Scheuerman, Jialun Aaron Jiang, Casey Fiesler, and Jed R Brubaker. A framework of severity for harmful content online. *PACM HCI*, 5(CSCW2):1–33, 2021.

[93] Ari Schlesinger, W. Keith Edwards, and Rebecca E Grinter. Intersectional HCI: Engaging identity through gender, race, and class. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 5412–5427, 2017.

[94] Zachary Schmitt and Svetlana Yarosh. Participatory design of technologies to support recovery from substance use disorders. *PACM HCI*, 2(CSCW):1–27, 2018.

[95] Irina Shklovski and Volker Wulf. The use of private mobile phones at war: Accounts from the Donbas conflict. In *Proc. CHI*, 2018.

[96] Lucy Simko, Ada Lerner, Samia Ibtasam, Franziska Roesner, and Tadayoshi Kohno. Computer security and privacy for refugees in the United States. In *Proc. IEEE S&P*, 2018.

[97] Manya Sleeper, Tara Matthews, Kathleen O'Leary, Anna Turner, Jill Palzkill Woelfer, Martin Shelton, Andrew Oplinger, Andreas Schou, and Sunny Consolvo. Tough times at transitional homeless shelters: Considering the impact of financial insecurity on digital security and privacy. In *Proc. CHI*, 2019.

[98] Brian Stanton, Mary Theofanos, Sandra Spickard Prettyman, and Susanne Furman. Security fatigue. *IT Professional*, 18(5):26–32, 2016.

[99] Elizabeth Stobert, David Barrera, Valérie Homier, and Daniel Kollek. Understanding cybersecurity practices in emergency departments. In *Proc. CHI*, 2020.

[100] Angelika Strohmayer, Jenn Clamen, and Mary Laing. Technologies for social justice: Lessons from sex workers on the front lines. In *Proc. CHI*, 2019.

[101] Borislav Tadic, Markus Rohde, Volker Wulf, and David Randall. ICT use by prominent activists in Republika Srpska. In *Proc. CHI*, 2016.

[102] David R. Thomas. A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2):237–246, 2006.

[103] Kurt Thomas, Devdatta Akhawe, Michael Bailey, Dan Boneh, Elie Bursztein, Sunny Consolvo, Nicola Dell, Zakir Durumeric, Patrick Gage Kelley, Deepak Kumar, et al. SoK: Hate, harassment, and the changing landscape of online abuse. In *Proc. IEEE S&P*, 2021.

[104] Alexandra To, Wenxia Sweeney, Jessica Hammer, and Geoff Kaufman. "They just don't get it": Towards social technologies for coping with interpersonal racism. *PACM HCI*, 4(CSCW1):1–29, 2020.

[105] Emily Tseng, Diana Freed, Kristen Engel, Thomas Ristenpart, and Nicola Dell. A digital safety dilemma: Analysis of computer-mediated computer security interventions for intimate partner violence during COVID-19. In *Proc. CHI*, 2021.

[106] U.S. Federal Bureau of Investigation. Business e-mail compromise the 12 billion dollar scam. https://www.ic3.gov/Media/Y2018/PSA180712, 2018.

[107] U.S. Federal Bureau of Investigation. FBI Tech Tuesday: Strong passphrases and account protection. https://www.fbi.gov/contact-us/field-offices/phoenix/news/press-releases/fbi-tech-tuesday-strong-passphrases-and-account-protection, 2021.

[108] Aditya Vashistha, Richard Anderson, and Shrirang Mare. Examining security and privacy research in developing regions. In *Proc. COMPASS*, 2018.

[109] Aditya Vashistha, Abhinav Garg, Richard Anderson, and

Agha Ali Raza. Threats, abuses, flirting, and blackmail: Gender inequity in social media voice forums. In *Proc. CHI*, 2019.
[110] Jessica Vitak, Yuting Liao, Mega Subramaniam, and Priya Kumar. "I knew it was too good to be true": The challenges economically disadvantaged internet users face in assessing trustworthiness, avoiding scams, and developing self-efficacy online. *PACM HCI*, 2(CSCW):1–25, 2018.
[111] Ashley Marie Walker, Yaxing Yao, Christine Geeng, Roberto Hoyle, and Pamela Wisniewski. Moving beyond 'one size fits all' research considerations for working with vulnerable populations. *Interactions*, 26(6):34–39, 2019.
[112] Ruolin Wang, Chun Yu, Xing-Dong Yang, Weijie He, and Yuanchun Shi. EarTouch: Facilitating smartphone use for visually impaired people in mobile and public scenarios. In *Proc. CHI*, 2019.
[113] Mark Warner, Andreas Gutmann, M Angela Sasse, and Ann Blandford. Privacy unraveling around explicit HIV status disclosure fields in the online geosocial hookup app Grindr. *PACM HCI*, 2(CSCW):1–22, 2018.
[114] Mark Warner, Agnieszka Kitkowska, Jo Gibbs, Juan F Maestre, and Ann Blandford. Evaluating 'prefer not to say' around sensitive disclosures. In *Proc. CHI*, 2020.
[115] Jeffrey Warshaw, Nina Taft, and Allison Woodruff. Intuitions, analytics, and killing ants: Inference literacy of high school-educated adults in the US. In *Proc. SOUPS*, 2016.
[116] Rick Wash and Emilee Rader. Too much knowledge? Security beliefs and protective behaviors among United States internet users. In *Proc. SOUPS*, 2015.
[117] Nicola Wendt, Rikke Bjerg Jensen, and Lizzie Coles-Kemp. Civic empowerment through digitalisation: The case of Greenlandic women. In *Proc. CHI*, 2020.
[118] Pamela Wisniewski, Arup Kumar Ghosh, Heng Xu, Mary Beth Rosson, and John M Carroll. Parental control vs. teen self-regulation: Is there a middle ground for mobile online safety? In *Proc. CSCW, 2017*, 2017.
[119] Pamela Wisniewski, Heng Xu, Mary Beth Rosson, and John M Carroll. Parents just don't understand: Why teens don't talk to parents about their online risk experiences. In *Proc. CSCW, 2017*, 2017.
[120] Pamela Wisniewski, Heng Xu, Mary Beth Rosson, Daniel F Perkins, and John M Carroll. Dear diary: Teens reflect on their weekly online risk experiences. In *Proc. CHI*, 2016.
[121] Brittany Wong. Wait, what the heck is a 'parasocial relationship'? https://www.huffpost.com/entry/parasocial-relationships-with-celebrities_l_60a56a18e4b0d45b75248115, 2021.
[122] Huichuan Xia, Yang Wang, Yun Huang, and Anuj Shah. "Our privacy needs to be protected at all costs": Crowd workers' privacy experiences on Amazon Mechanical Turk. *PACM HCI*, 1(CSCW):1–22, 2017.
[123] Fujiko Robledo Yamamoto, Amy Voida, and Stephen Voida. From therapy to teletherapy: Relocating mental health services online. *PACM HCI*, 5(CSCW2), 2021.
[124] Jun Zhao, Ge Wang, Carys Dally, Petr Slovak, Julian Edbrooke-Childs, Max Van Kleek, and Nigel Shadbolt. 'I make up a silly name': Understanding children's perception of privacy risks online. In *Proc. CHI*, 2019.

APPENDIX

Population categories and intersections. Provisional categories [64] of at-risk populations represented by our dataset are listed in the first column of Table I. While these population categories flatten some of the demographic richness reported in individual dataset papers, where possible we unpack intersectional issues using our thematic analysis and examples that include more detailed participant descriptions throughout the paper. For example, the "activists" category in Table I describes activists across four continents, but examples note the specific geography when relevant. Further, these population categories are not exhaustive; instead they represent our dataset, demonstrating how risk factors can differ across categories and intersect for a single category.

Biases in geographic representation. At-risk populations are found around the world, with social and structural circumstances that vary globally. The papers identified in our review skewed heavily toward Western, and specifically U.S., populations, but do include some studies with participants from other regions. About half of the papers in our dataset (48 out of 95 papers) included participants from the U.S. only, and 68% of papers included participants from Western countries only. Populations about which the literature reported perspectives from around the world (including most or all continents), included activists, journalists, NGO staff, crowdworkers, LGBTQ+ people, and survivors of trafficking. Other non-Western perspectives were reported for women in South Asia, older adults in Cambodia, people in developing regions (in Asian and African countries), refugees (resettled from Africa and Asia to the US), and people with visual impairments in India. The remaining populations listed in Table I represent Western perspectives.

Researcher reflexivity. Though we do not know how the authors of dataset papers would identify their nationalities or describe their lived experiences, the majority appeared to be completed by teams at institutions in Western contexts. Similarly, we work in U.S.-based organizations and have backgrounds in HCI, computer science, security, and/or privacy. These contexts influenced the papers we sampled and the analysis lenses used—both by our team and the researchers who produced dataset papers. While this SoK establishes a framework, we advocate for additional perspectives to enrich our findings.