Security and Privacy Software Creators' Perspectives on Unintended Consequences

Harshini Sri Ramulu, *Paderborn University & The George Washington University;* Helen Schmitt, *Paderborn University;* Dominik Wermke, *North Carolina State University;* Yasemin Acar, *Paderborn University & The George Washington University*

https://www.usenix.org/conference/usenixsecurity24/presentation/ramulu

This paper is included in the Proceedings of the 33rd USENIX Security Symposium.

August 14–16, 2024 • Philadelphia, PA, USA

978-1-939133-44-1

Open access to the Proceedings of the 33rd USENIX Security Symposium is sponsored by USENIX.

Security and Privacy Software Creators' Perspectives on Unintended Consequences

Harshini Sri Ramulu[§†], Helen Schmitt[§], Dominik Wermke[*], Yasemin Acar [§†]
[§]*Paderborn University,* [†]*The George Washington University,*
[*]*North Carolina State University*

Abstract

Security & Privacy (S&P) software is created to have positive impacts on people: to protect them from surveillance and attacks, enhance their privacy, and keep them safe. Despite these positive intentions, S&P software can have unintended consequences, such as enabling and protecting criminals, misleading people into using the software with a false sense of security, and being inaccessible to users without strong technical backgrounds or with specific accessibility needs. In this study, through 14 semi-structured expert interviews with S&P software creators, we explore whether and how S&P software creators foresee and mitigate unintended consequences. We find that unintended consequences are often overlooked and ignored. When addressed, they are done in unstructured ways—often ad hoc and just based on user feedback—thereby shifting the burden to users. To reduce this burden on users and more effectively create positive change, we recommend S&P software creators to proactively consider and mitigate unintended consequences through increasing awareness and education, promoting accountability at the organizational level to mitigate issues, and using systematic toolkits for anticipating impacts.

1 Introduction

Security and Privacy (S&P) software is created to protect users from attacks, guard their privacy, and keep their communication confidential. However, technology may have unintended consequences that harm users, non-users, or society [1]. *Unintended consequences* are unforeseen outcomes of purposeful actions—classified as unexpected drawbacks, perverse results, and/or unforeseen benefits [2]. For security and privacy software, unintended consequences can include propagating misinformation [3], using the software for criminal activities [4] such as the proliferation of child sexual exploitation [5], and excluding certain users. For example, a lack of accessibility can exclude or harm users with disabilities, children, older adults, activists, non-Western populations, victims of intimate partner abuse, and other vulnerable groups [6]; privacy browsers not considering the privacy needs of users requiring assistive technologies can inadvertently put these users at risk [7], [8]. Identifying unintended consequences of software can be very complicated [1], and software creators frequently struggle to address the diverse needs of their users [9]. Researchers and communities have been working to effectively anticipate unintended consequences since before the spread of computers [2].

While the Human-Computer Interaction (HCI) community studies unintended consequences and harms in the broad context of software [1], [10], there has only been limited focus on S&P software [11], [12]. We argue that it is worth studying the unintended consequences of security and privacy software, particularly for the following reasons: Users explicitly trust these tools to protect and handle sensitive information and data. Even minor unintended harms can majorly impact users' privacy and security, which may risk their lives and safety, or they might abandon the tools and their benefits—putting them at more risk. Developers of this software typically exhibit strong values and care deeply about their users' safety. Therefore, the criticality of considering unintended consequences in security and privacy software cannot be overstated.

Ethics frameworks [13], toolkits [14]–[16], and checklists [17] have been proposed to help software teams systematically anticipate and consider unintended consequences during software development: In particular, the HCI and Artificial Intelligence (AI) communities have explored assessing unintended consequences through toolkits [14], [15], [18], [19]. These toolkits categorize consequences that may emerge from the use of technology and are supposed to help organizations analyze the potential for unintentional consequences arising from their software. Aspects including misinformation, inequalities, biases, addiction, distraction, accessibility, surveillance, and risks through harmful actors are addressed in these frameworks [13], [16], [18], [20]–[22]. While these toolkits seem promising, it is unclear whether they are effectively used to mitigate unintended consequences in S&P software.

We address a gap in our understanding of whether and how S&P software creators identify and mitigate unintended consequences and systematically address issues during development. For this, we conduct an interview study with 14 creators (we use this term throughout to include developers, CEOs, designers, industry researchers, and maintainers) involved in the development of user-facing and developer-facing security and privacy software (such as, e.g., VPNs and secure messengers) to understand if and how they consider and mitigate unintended consequences of their software. Furthermore, we investigate the attitudes and perceptions of S&P software creators toward their software's purposes and unintended consequences. For this research, we define security and privacy software as software developed specifically for security and privacy purposes, with core values centered around security and privacy, and software developed to replace existing software in a more secure and privacy-preserving way.

The research questions that guide this study are as follows:

RQ1: *What are the attitudes and practices of S&P software creators concerning unintended consequences of their software?* We investigate motivations for developing S&P software, the software's purpose, the users it serves, and how the creators engage with users to understand their needs and expectations.

RQ2: *What unintended consequence of S&P software do S&P creators anticipate, how do they reason about them, and how do they mitigate them?* We explore unintended consequences creators of end user-facing security and privacy software anticipate. We investigate their reasoning for what they consider important, which unintended consequences they address and how, and which they do not address and why not.

RQ3: *What are facilitators and blockers in (systematically) considering and addressing unintended consequences?* We explore how systematic assessment can be facilitated and how (lack of) knowledge and collaboration can impact the consideration and resolution of unintended consequences.

We find that S&P software creators are motivated to create a positive impact on their users, but they do not systematically anticipate and mitigate unintended consequences. Their primary focus is on privacy and security-related aspects, while other consequences that can harm or exclude some users are considered anecdotally based on personal interests and experiences within the development team, if at all. Some significant challenges that limit considerations of unintended consequences include a lack of organizational support, a lack of awareness and knowledge, and a lack of formal roles and responsibilities to consider unintended consequences.

2 Related Work

We discuss related work in three key areas: (1) unintended consequences of S&P software, (2) frameworks and tools helpful in anticipating and mitigating unintended consequences, and (3) studies with developers and software professionals. While our research focuses on unintended consequences of S&P software, the concept of considering unintended consequences is often discussed and addressed in the related work as "ethics" or "ethical impacts" [10].

2.1 Unintended consequences of S&P software

S&P software is designed to provide useful solutions to security and privacy problems but may have drawbacks caused by oversights in issues like accessibility or usability factors or due to misuse by malicious actors. Below, we provide instances of harm that can be caused by S&P software.

Access barriers. Traditional anti-virus companies offer anti-stalkerware software, but their high costs, ranging from $5 to $100, and compatibility issues with different operating systems make them inaccessible to many users [23]. While VPNs are promoted to increase privacy online, low-cost ones suffer from low adoption rates among non-tech-savvy users, and others have high costs that make them inaccessible to a broader audience [24].

Misleading marketing and non-disclosure. In a 2022 study, Akgul *et al.* found that VPN companies use misleading marketing strategies with social media influencers to attract users [25]. Their ads included exaggerated claims about online safety beyond what VPNs are capable of, misleading novice users into choosing their services based on false promises. Fassl *et al.* found that anti-stalkerware apps did not detect all vulnerabilities users expected based on the apps' presentation and may, therefore, deceive users and leave them unprotected [23]. False marketing and not disclosing software capabilities can lead users to trust the software too much, which may result in compromised security.

Accessibility issues. Studies highlight that security solutions sometimes do not consider disabilities. In particular, people with visual impairments find it hard to access websites that require authentication technologies [7], [8]. People with dyslexia often struggle with creating and remembering unique passwords [26]. This issue can be exacerbated if the password manager is incompatible across devices, preventing them from saving passwords on different devices [27]. A lack of consideration of disabilities in security & privacy software design might push people with disabilities to resort to less secure alternatives [8].

Usability issues. Research shows that usability problems can prevent using security and privacy technology effectively. For example, unsolved or undetected usability issues can cause an increase in cognitive load, frustration, and even security vulnerabilities like password exposures [28] or unencrypted communication [29]. A 2021 Huaman *et al.* study illustrates usability problems with the compatibility of password managers across different websites and devices that may impede adoption [30]. Another 2021 study demonstrated that cryptographically secure technologies come with usability issues for

users who do not have the necessary knowledge to use them, which can lead to a lack of adoption of such technologies [5].

Stalking, abuse, and harassment. Anti-theft tracking and child tracking software that can be downloaded from the app stores can be used for harmful and abusive purposes like spying and stalking people [31]. Previous literature also shows how technology can be used to perpetuate Intimate Partner Violence (IPV) through surveillance [32], doxxing [33], cyberstalking [34], and other means. A common approach taken by abusers is to install stalkerware on their victims' phones [23] or use AirTags to track their victims' location [35]. Privacy Enhancing Technologies (PETs) are sometimes used to perpetrate crimes like the distribution of imagery of child sexual abuse [4]. Wei *et al.* identified instances of advice being circulated on social media that suggest ways to use technology, which is created with positive intentions such as accessibility features, to secretly monitor individuals without their consent, thereby violating their privacy and security [36].

Harms to vulnerable groups. In some cases, members of marginalized communities are disproportionately impacted due to S&P violations [6], [37], [38]. Research also highlights the importance of designing with the needs of marginalized groups in mind, for example, considering factors like language barriers, digital literacy, and access to technology [39]. A growing area of research in the S&P community focuses on identifying the needs of vulnerable and/or marginalized populations, like IPV survivors [40], refugees [39], activists [5], and sex workers [41].

2.2 Considering Unintended Consequences

Prior work explains that neglecting considerations of unintended consequences during development can cause tragic failures [9]. It suggests considering the worst possible impact of the software during the development phase to deliver ethically sound systems. In the following, we discuss how this might be accomplished.

Tools used to anticipate unintended consequences. Toolkits (also called "frameworks"; we use the term "toolkit" throughout this paper for consistency) are commonly suggested to anticipate and address unintended consequences systematically. In the privacy space, privacy frameworks such as Privacy by Design [42], Fair Information Practices [43], and Privacy Impact Assessment [44] primarily focus on preserving the privacy of individuals. They generally do not address other unintended consequences for security and privacy software. For instance, enhancing privacy for a single group can harm others, as in the case of refugees who had difficulty answering security questions due to cultural differences, which led them to create less secure answers [39]. Similarly, the very possession of privacy browsers or downloading certain end-to-end encryption messengers may incriminate people in some countries [45]. To gain a holistic view of harms that can arise as a result of software use, ethics frameworks [13], toolkits [13], [16], [18], [20]–[22], and datasheets [20]–[22] have been proposed. These assessment tools cover a more comprehensive range of issues that can occur due to software use, including biases, discrimination, lack of accessibility, and more. While there are also other methods like value sensitive design [46], speculative design [47], and inclusive cards [48], these methods have been criticized for overclaiming universal values and for undermining user voices due to the assumed authority of the development team over users [49]. Concurrent work analyzed 27 AI-ethics toolkits, criticizing their technology-centered design and the neglect of organizational and power dynamics [50]. Our interview guide is informed by topics we extract from nine toolkits to identify common software harms addressed in S&P software. While we hypothesize that these toolkits can expand awareness of unintended consequences beyond privacy issues, we are also interested in impediments to addressing the unintended consequences creators discuss in our interview study.

Navigating unintended consequences in software development. Several factors affect how unintended consequences and software harms are handled during the development process. Organizational climate plays a significant role in developers' ethical awareness and behavior [51], and organizations need to foster an environment to identify and address unintended impacts [52]. Technologists may show little enthusiasm towards ethics [53] (a finding that our study does not corroborate), and software issues are often only addressed on a legal compliance level [54]. Developers may lack expertise and knowledge in areas surrounding their field of expertise [54] due to insufficient awareness, time, and resources [55], which may also contribute to unintended consequences. Team diversity also affects how unintended impacts are assessed [56]. Previous research highlights the lack of diversity in the vulnerability discovery community, noting that it is predominantly dominated by white and Asian men [57]. The study uncovers unique challenges marginalized populations face, revealing the need for more inclusivity in security. Our work extends this conversation to the role of diversity and other factors that influence anticipating unintended consequences during the development of S&P software.

Previous studies have highlighted the importance of education and awareness in developing ethical software and protecting end users from unintended consequences [58], [59]. However, research shows that software developers receive little to no ethical training in classrooms [56]. In a study conducted with students studying AI, students were not inclined to think about the implications of AI technologies. When they considered the impacts, it was centered on personal experiences. This study suggests that it is vital to consider integrating ethics education into the computer science course curriculum to support their future considerations of ethics in AI learning [56]. Further, Do and Pang showed that computer science researchers tend to overlook unintended consequences of their research for many reasons, including the lack of for-

mal processes, fast-paced academic publishing process, and reliance on Institutional Review Boards (IRB) to point out issues [10]. In our study, our objective is to determine attitudes and processes that help or hinder anticipating and mitigating unintended consequences during the development of S&P software.

2.3 Expert Studies

Here, we examine previous research conducted with software professionals and our methodology is informed by best practices from prior work.

Interviews with software experts. Interview studies have been used in the past to gain deep insights into the work, processes, and mental models of experts [52], [60]–[63]. For instance, several studies illustrate that interviews are an effective method to understand the processes and challenges of developers [64]and organizations [65]. Similarly, we will use interviews to identify S&P creators' considerations of the impacts of their software during the software development process. We focus on creators of S&P software, like developers, designers, researchers, and executives involved in the development, as these experts share a vision to create positive impacts by improving security and privacy.

3 Methods

We conducted 14 semi-structured expert interviews with creators involved in the development of S&P software; interviews lasted just under one hour on average. In this section, we explain the methodology of our study, including the development of the interview guide, recruitment process, data collection, analysis, ethical considerations, and limitations.

3.1 Instrument development

Analysis of toolkits to anticipate unintended consequences of software. We developed our interview guide by analyzing nine toolkits that identify themes related to potential negative consequences and ethical impacts of technology, which we call *unintended consequences*. We used the Affinity Diagramming approach [66] to identify common aspects covered in these toolkits, which we draw from previous work discussed in Section 2. These toolkits are generally lists of questions or checklists to help technology developers anticipate unintended consequences, such as misuse of technology and broader ethical issues. We identified and analyzed ethical technology assessment toolkits from Human-Computer Interaction (HCI) literature [19] such as the *Ethical OS Toolkit* [14], the *Digital Impact Toolkit* [15], and *Ethics and Algorithms Toolkit* [16]. We also analyzed frameworks and datasheets for software development, including the *Data Ethics Framework* from the UK government [13], the *European Union Guidelines for Trustworthy AI* [20], *ACM Code of Software Ethics* [22], *World Economic Forum Ethics for Responsible AI* [21], and *Japanese Society for AI Ethics Guidelines* [67]. We stopped analyzing toolkits when we reached topic saturation [50], [68]; concurrent work by Wong *et al.*
The first part includes topics around *ethical and societal consequences* such as barriers to access, biases, accessibility issues, disempowerment, misinformation, addiction, distraction, physical and mental health, environmental impacts, and human rights violations. The second part addresses *data privacy and security risks* such as security measures, protection of user data, data governance, tracking, and surveillance. We used these aspects as probing questions to guide participants in considering potential unintended consequences and whether they addressed them during development.

Development of interview guide. Two authors crafted the interview guide's initial version based on our research questions and toolkit analysis. We went through several iterations and brainstorming sessions to improve the quality of the interview questions. After receiving feedback from a third author, we refined the guide before conducting pilot interviews. The complete interview guide is available in Appendix A.

The final interview guide consists of questions in four key areas: (1) *About the software, motivations, and vision*: Here, we focused on details about the software; we were particularly interested in the core values of the software and the key problems creators intended to solve. (2) *Users and non-users:* Here, we asked participants about their primary users, bystanders, unintentional users, and non-users and how the software was designed to suit users' needs. (3) *Unintended consequences:* they anticipate through the use of their software: here we were specifically interested to learn about how participants thought their software impacted their users and non-users. We started this section with open questions, where participants mentioned the consequences they thought of or had considered previously, both positive and negative. Later, we probed about the themes identified in our toolkit analysis to understand whether participants considered these aspects in their development process. In this section, we also asked about their security and data privacy considerations. (4) *Processes for mitigating consequences:* Here, we asked about the structure of the development team and their organization, and who is responsible for mitigating unintended consequences. We were also focused on the current processes, tools, and resources that support or hinder them in identifying and mitigating unintended consequences.

3.2 Interview procedure

For our study, we interviewed a total of 18 participants, of which 14 were eligible; two participants were removed, as we identified their software as out of scope after the conclusion of the interviews, two were obvious scammers. Valid interviews lasted an average of 57.35 minutes.

Pilot study. We piloted our interviews to evaluate the interview questions' validity, comprehensibility, and language. We conducted four pilots: two with professional colleagues not involved in this project and two with software creators from another industry. We used the feedback to add examples and details to questions that were unclear to our pilot participants.

Recruitment. For this study, we recruited participants involved in the creation of user-facing and developer-facing security and privacy software. For example, we included privacy-preserving browsers and secure messengers. Our recruitment process targeted diverse roles in the development process, such as CEOs, researchers, designers, front-end and back-end developers, and engineers. We recruited S&P software creators through personal and professional networks; we additionally used snowball sampling to recruit other experts that our participants knew in other relevant companies. Once we exhausted professional contacts, we resorted to social media. We posted a recruitment message on then-Twitter seeking potential participants. We also cold-emailed developers who contributed to S&P projects on GitHub and in S&P groups on Slack. For social media recruiting, we initially asked participants to fill out an eligibility survey, which can be found in Appendix A, and invited those who qualified for the actual interview; both scammers were recruited through social media.

Interview process. We conducted all except two interviews on Zoom; two interviewees preferred to use our self-hosted Jitsi instance instead. We generally conducted the interviews in lead-interviewer configuration, in some cases joined by a backup interviewer. Participants were given an option to opt-in for an $80 Amazon voucher as compensation for their participation in the interviews, unless they waived it. Interviews were all conducted in English. All interviews except one were recorded and transcribed by a GDPR-compliant service. For accessibility reasons, one of the participants preferred to switch between chat and voice call, so we transcribed one interview ourselves using the chat messages and the notes we took during the interview. Participants received informed consent forms before the interviews, and we also obtained verbal consent to record before starting the interviews.

Ethical Considerations. This study was approved by our university's IRB. We de-identified any personally identifying data of our participants as much as possible. We only collected their email addresses if they opted-in to receive an $80 gift voucher for their participation, or if they wanted to be mailed a copy of our results which they will receive if requested. Apart from optionally provided email addresses, all the other personal data is associated with random identifiers. We also ensured that we masked details about their companies, coworkers, or direct competitors they mentioned in our transcripts. Given that we probed participants to think about the negative consequences of their software during the interviews, we ensured that they could refrain from answering questions if they were uncomfortable to avoid reputational harm to their employers. We also made it clear to them that they could decline to answer questions or withdraw from the interview at any time.

3.3 Data Analysis

We used open coding to analyze interview transcripts qualitatively [69]. Two researchers independently coded all the transcripts, regularly met, and iteratively built a codebook by combining their interpretations of the data, the codebook is attached to Appendix C. We also regularly discussed the codes with the rest of the research team for feedback and consensus on naming conventions. By double-coding all transcripts, we actively resolved disagreements and, therefore, did not calculate the intercoder agreement [70]. Finally, the entire research team collaboratively worked together to identify clusters and generate themes within the coded data; themes are available in Appendix B. We use these themes to report our results and guide discussion points and future research directions. We fully developed themes using the first 12 interviews—these themes remained stable for the last two interviews—for major themes, we reached thematic saturation [71]. Due to the nature of the differing software and experiences, interesting insights emerged with every new interview but did not change major themes.

3.4 Limitations

This paper presents the results of a qualitative study and therefore shares limitations inherent to qualitative research, such as limited generalizability. Participants were mostly recruited by purposive sampling, using personal and professional contacts. As a result, most participants have a similar background and are mostly from the US and the EU. Once we exhausted our contacts, we resorted to recruiting on social media. Recruiting from then-Twitter and Slack resulted in many scammers signing up for the study. To filter out scammers, we only chose participants who signed up with their professional emails. We discarded the data of two participants who provided generic responses throughout the interview, identifying them as scammers. Generally, recruitment was difficult and drawn out, as we were focused on an expert population of creators of S&P software. As in any study, participants over-reporting their ethical considerations due to social desirability could bias their answers—and, consequently, our results.

4 Results

We find that participants do anticipate some unintended consequences but do not *systematically* anticipate or mitigate them. We report our findings in four areas: First, we provide an overview of participants, the S&P software they develop, and who they consider (non) users (Section 4.1). Second, we present participants' (lack of) practices and motivations for

Table 1: Participants' Roles and Type of Software.

P.No	Role	Software
P01	Lead Developer	Secure Messaging
P02	CEO	Password Manager
P03	Backend Developer	Encrypted Email
P04	Sole Developer	Software Release Signing Tool
P05	Researcher	Stalkerware Detection
P06	Maintainer	Social Network
P07	Executive Director	Anonymity Network
P08	Core developer	VPN
P09	Developer	VPN
P10	Researcher	Browser
P11	Privacy engineer	Browser
P12	Software/UI developer	Secure Translation
P13	Design Lead	Encrypted Office Suite
P14	UX Designer	File Sharing and Collaboration

Table 2: Participants' Gender, Experience, and Organization Size.

Gender		Years of Experience		Size of Organization	
Women	3	< 5 years	1	< 10	5
Men	9	5 - 10 years	3	10 - 100	7
Non-binary	2	10 + years	10	100 - 1000	1
				1000 +	1

anticipating unintended consequences (Section 4.2). Third, we explore which unintended consequences are (not) considered (Section 4.3). Finally, we present the factors that help or hinder development teams from anticipating and mitigating unintended consequences (Section 4.4).

4.1 Participants, S&P software, and their users

Here, we describe our study participants, the S&P software they create, and the user population they report.

Participants. We interviewed 14 valid participants involved in developing 13 different S&P software projects, including VPN, secure messaging, and security- or privacy-focused browsers. An overview of the roles and software types is in Table 1. We interviewed participants involved in the development team in technical roles like developers, designers, engineers, researchers, and those responsible for overseeing the product in roles like executive directors and CEO. We attached our screening criteria in Appendix A. Most participants were men, three were women, and two were non-binary. Participants had varying years of on-the-job experience ranging from three to 30 years.

Organization sizes ranged from one-person teams and people who do volunteer work to large organizations with over a thousand employees. An overview of demographics and bucketed organization sizes is in Table 2. Many participants reported that their teams were not diverse and said that their teams constituted predominantly white men. Only one developer stated they had a process to diversify their team: "*[. . .]it's a continuous effort. I wouldn't say we are there yet.*" (P01) A few participants saw their team as relatively diverse concerning ethnicity but not gender. "*Compared to most other development groups, we are pretty diverse in terms of country origins.*" (P01)

Primary, Secondary, and Unintended Users. Most software is marketed towards and adopted by privacy-conscious users. Participants described primary users as privacy-conscious users trying to protect their identities, including special populations like activists, journalists, and people targeted by governments. For instance, Participant P07, who works on the anonymity network stated that their user base included, "*human rights defenders, activists, lawyers, journalists to just a dad who wants the kids to not have the information about what they search online being collected by search engines or a doctor who wants to protect the privacy of their patients as they are browsing[. . .].*" (P07) Some participants revealed that they make special efforts to cater to specific demographics to strengthen their privacy, e.g., the secure messaging software (P01) and a VPN (P08) are primarily used in the global south. Participants P05, P10, and P11 stated they do not target special populations, and their predominant users are in the global north. The password manager's users were mostly corporate employees (P02), and the software release signing software was mostly targeted toward open source developers (P04).

Some participants acknowledged that harmful actors are also among their primary user base. P07 said that their software is used for "*[. . .]illegal content or for illegal activity by selling drugs or other things in that sense,*" and P06 mentioned that their social network "*did have attraction, for example, from ISIS [. . .]*" and "*anti-vax groups.*" We elaborate on this in Section 4.3.3. Bystanders and non-users can also be significantly impacted by software; participants noted that they had not considered the threat models of bystanders and non-users in the development process. For instance, the encrypted email service has cumbersome processes for recipients not on their server, and thus their users resort to sending unencrypted emails: "*[. . .] because then what happens is we create a temporary mailbox for the external user, and they have to log in with a pre-shared password, and then the user can see the email.*" (P03)

Given that S&P software caters to a wide spectrum of users, ranging from regular users to criminals, harm can proliferate via multiple modes, from adversarial actors using the software to non-tech-savvy users being unable to use the software. In the next section, we will look at how participants currently foresee unintended consequences.

4.2 Practices and motivations for anticipating unintended consequences

Participants are motivated to have a positive impact with their software, especially on security and privacy. Their strong fo-

cus on security and privacy can overshadow consideration for unintended consequences. The focus on the inherent positive impact of the software and user feedback can be limiting, especially without a structured process to identify unintended consequences. In this subsection, we discuss participants' attitudes and current practices for anticipating unintended consequences.

4.2.1 Motivations: privacy, usability, and "doing good"

Privacy-centered development process. Protecting people's privacy was the participants' major focus and the key problem their software is trying to solve; their thoughts on unintended consequences also revolved around privacy issues. In addition to "privacy", "anonymity", "end-to-end encryption", and "security" were also mentioned as core values. Participants highlighted that they specifically work on S&P software to make a positive impact, contributing to a world where people's information is protected and untrackable. For instance, the participants developing the browser and VPN services highlighted that their software solves issues related to tracking and surveillance: "*The key problem is surveillance. Corporations and states do tend to spy upon their customers and citizens*" (P09) and "*Our mission is to advance human rights by building privacy, security, and ordinary technology.*" (P07) Thus, participants indicated that they were motivated to contribute to protecting their users' privacy, but they had not meaningfully considered the unintended consequences of their software.

Creating usable products. Some participants indicated following user-centered processes during development to make security easier for their target users. While striving for usability, some actively involved vulnerable user groups to make products less harmful for them. Some participants acknowledge that security software can be very difficult to use correctly and strive to make it easy for users: Participant P02 stated, "*[. . .] in comparison to the competition, we try to be the easiest password manager.*" In order to achieve usability, participants specified using techniques like testing new features with beta users (P10), using personas for development (P01), co-designing and testing with the target population of their software: "*[. . .]the same is true with other Global South countries. We also try to do our development physically in these environments as much as we can.*" (P01) Participant P11, who reached out to IPV survivors during development, emphasized that not involving affected users can cause more harm than good: "*[. . .]we're trying to talk to folks who actually work in cases of domestic abuse, there are some centers. I want to make sure that the solutions that we're coming up with are in collaboration with them so that it actually helps and doesn't end up harming even more because we have some insight or some folks on our team have some insight, but we need more data.*" (P11)

Personal interests and experiences drive them to create a positive impact. In addition to the privacy-protecting values of the software, some participants stated that they were intrinsically motivated to develop ethically sound software, have a personal interest in ethics, and think about how their technology can harm people. Participants P02 and P11 stated they are interested and driven by a political interest to do good for people, wanting to provide good software for their users. Their personal interest motivated them to learn about ethics (P02) and reach out to marginalized user groups for feedback when designing new features (P11). Participant P05 brings up personal interest:

> "*I often take it upon myself to raise concerns, especially when people start thinking about solving a problem for parents, like: 'Oh, I want to help parents to protect or to monitor their kids', then I have to be like, 'They've got to be careful because that could be used to monitor people that are not kids or things like that'.*" — P05, Stalkerware Detection

Additionally, a few participants mentioned that they adhere to their *inner moral framework*, as P04 explains: "*It's not a scientific endeavor for me. I just want to do what is right in my eyes.*"

In the aforementioned themes, we find that participants were enthusiastic and wanted to contribute positively to their users, however, they inadvertently overlook other harms that may occur due to their focus on mostly privacy and usability. In the following section, we explore their current software development practices and how they lead to an oversight in anticipating unintended consequences.

4.2.2 Current practices: reactive and unstructured

Current practices that S&P software creators follow to avoid unintended consequences currently focus on usability issues and reacting to issues raised by their users. Additionally, there is a lack of formal processes to anticipate and mitigate unintended harms.

Reliance on user complaints rather than being proactive. Participants expected users to bring unintended consequences to their attention, for example, through GitHub issues, product reviews, and customer support. Some monitor product reviews and feedback surveys to identify problems their users encounter: "*[. . .]users continue to contact customer support, they leave reviews for our apps, and we monitor those to see if there are any issues with false positives or any complaints as we release the products and update them and things like that.*" (P05) Participant P04 also mentioned that if users encounter challenges, they can open issues on GitHub. Some participants mentioned monitoring social media channels that they use to interact with users to identify user problems. While user feedback is regularly monitored, sometimes issues may take a long time to fix or the feedback is obtained after harm has already occurred: "*through feedback collection, where*

like we learn about the main pain points our users have and we have. Sometimes they are not easy problems to fix. They may take like a year for us to get some solution out there." (P07) Participant P05 recalled a past incident where they received an uptake in user complaints after releasing a feature that provided warnings in their stalkerware detection software where, in some cases, perpetrators were potentially notified, and users deleted the software.

Fixing usability issues and bugs is the main focus. When talking about unintended consequences that can be detrimental to their users, participants stated that they have anticipated bugs in their code that might affect usability and are well-equipped to tackle usability issues. While usability is a critical aspect that can lead to issues with software adoption and in some cases abandonment, we think that participants overstated usability as the most critical negative impact. Participant P11, who works for a browser, mentioned that users might abandon the product if a website does not work on their browser, which may compromise their privacy: "*If you're browsing your favorite website and in Google Chrome, it works fine because Chrome lets the website fingerprint you and do whatever it wants with your data. But if that same website doesn't work in [browser] now, you will not use [browser], and then you lose out on a bunch of other additional privacy protection.*" Some had dedicated resources and processes to mitigate usability issues:

> "*Of course, there might be bugs in the software, like we might have bugs so that people cannot properly receive messages. We have some reports of those; this can be dangerous, like if you send a message that doesn't arrive on the other side for whatever reason.*" — P01, Decentralized Messenger

While usability was prioritized, unintended harms were sometimes overlooked, which was majorly due to lack of knowledge and resources to proactively consider them.

Lack of formal processes to anticipate and mitigate unintended consequences. A majority of the participants stated that their teams do not have formal processes or toolkits they follow to anticipate unintended consequences. The teams decide on processes depending on what fits them best, like involving collaborators and experts, allowing collaborators to open issues on GitHub, and facilitating discussions within teams. "*Other than us continuously discussing this, there's no procedure in place. Also, there's no checklist that we go through and so on.*" (P01) Participants mentioned that they take a reactive approach. They mitigate issues if users flag potential threats presented by harmful actors, "*[. . .]other than closing abusive accounts, we don't have any other processes in place.*" (P03) Participants mentioned that there were either no formalized processes in place or they were personally unaware of their existence, thereby not accounting for them in their work: "*We don't really use any kind of ethical framework. We don't think about the number of users impacted versus the harm or anything more formalized like that.*" (P11)

4.3 Reasons unintended consequences are or are not considered

4.3.1 Aspects that are ingrained in the software vision

Security. All participants specified that they followed processes to ensure security and keep their software free from vulnerabilities. Most participants stated that they have regular security reviews and follow industry standards to ensure security. Other security mechanisms were embedded in software development: P01 pointed out using Rust to develop the secure messenger and conducting usable security research regularly, P02 uses encryption in their password manager, and P04 also took additional security measures. Participant P03 highlighted that they use a threat model whenever developing a new feature.

Privacy & data protection practices. Most participants report collecting some sort of user data, few collect none. Participant P06 stated that users can use dummy information to signup on their social networking platform: "*if they ask them for like [sic] the name or we ask them for the location, if they want to enter dummy information, that's perfectly fine with us, you know, for the location, you don't do any validation there.*" (P06) Participant P03 mentioned that they allow anonymous signup so that the users need not use a phone number or email to sign up. However, they collect and store IP addresses for a short time. Few participants mentioned trying to collect only the bare minimum amount of data required to use their services. This included the password manager (P02) collecting metadata like "*how many people logged in today,*" a privacy browser (P11) collected minimal data on usage statistics: "*we try to collect the least amount of data and we are pretty good about updating our privacy policy to show all the kinds of data that we collect, so transparency is a goal.*" (P11) All participants mentioned a data protection practice; P02 mentioned that the password manager has a data deletion process, P02 and P10 mentioned they follow GDPR practices, four participants mentioned that they have a review process to ensure data governance procedures, and four participants also mentioned having data privacy officers and legal personnel to conduct reviews. The teams who did not follow constant review practices followed GDPR (2) and got in touch with external experts (1) to ensure good data protection practices.

Some software inherently has a positive impact on users' mental health. Many participants emphasized that their software positively impacts their users' physical and mental health. For instance, the password manager does not require people to remember complex passwords; encrypted email and secure messaging software help keep information secure, and VPNs make people feel "*less watched or surveilled, which can have a positive impact.*" (P09) Participant P11 mentioned

that their browser's ad-blocking could have a positive impact on their users' mental health by blocking disturbing ads: "*it benefits users' mental well-being positively is that we block all ads by default[...]I think users just don't have to see a bunch of really nasty ads across the web.*" (P11)

4.3.2 Prior awareness does not guarantee mitigation

Software is mostly not designed for non-western and non-technical users. While most participants could identify access barriers and biases of their software, many did not consider mitigating them during the development process, and only a few were working on actively making the software accessible to more audiences. Common barriers to access that participants pointed out were: "*low awareness*" (P01) of the product because these products are mostly marketed to the "*sphere of users who are already seeking privacy software or a Western audience,*" (P03) software being too "*technical for some users,*" (P02) high costs for certain populations, OS compatibility, lack of stable internet connections, and language barriers. Some participants were actively making the product available in more languages and/or countries and in more "*precarious situations than the well-resourced Western countries.*" (P01) Additionally, P12 remarked that mistranslations of some phrases in certain languages could harm users.

Making software accessible is not a priority. Few participants had processes in place to design for accessibility. On the lack of processes, they remarked that they either lack resources or have not considered prioritizing accessibility: "*It is not accessibility first in general. Mainly, I would say because of a lack of specific training in the developer community towards these needs.*" (P09) Participant P01 acknowledged that the desktop version of their decentralized messenger has accessibility issues and is not usable for users with visual impairments. A few participants pointed out that they followed best practices for accessibility, but they did not test whether the products worked well. For example, the privacy browser's interoperability with accessibility tools was not tested, e.g., whether the privacy browsers worked with accessibility tools. However, they cannot control how website developers operate. "*Then there are many limited things that we have there in that as a browser we could provide all of these properties that enhance accessibility, but is the responsibility of the website operators to actually use them.*" (P01) Only a few participants indicated that they have specific design teams responsible for accessibility: "*Internally, we have UX researchers and UX designers who are experts on accessibility.*" (P11)

Distraction and addiction are sometimes addressed through design. A few participants (browsers and communication tools) stated that their software was prone to maximizing users' attention. However, to mitigate this problem to a certain extent, the browser and the social network did not display ads, nor were their business models designed to maximize attention. Therefore, P06 mentioned that it depends on how users manage their time, and they "*have no incentive for maximizing your attention on the product*" because there are no ads on the platform.

The encrypted email, secure messaging, and browser participants indicated that their software could distract users. P10 had strategies to limit distractions by limiting ads on their browser. The secure messaging software limits distraction by allowing users to disable notifications. Participant P03, the encrypted email developer, mentioned that mitigation depends on users; however, they default to "non-intrusive defaults."

One participant mentioned that stalkerware warnings can make people anxious:

> "*[...]so if it's not giving any warnings, it gives them peace of mind, if there are warnings, though the warnings could make them more anxious but hopefully for a good reason.*" — P05, Stalkerware Detection

4.3.3 Aspects participants have no control over

Some issues may surface due to lack of control and missing content moderation. Participant P06 said that their social media platform could be used to share content that can violate human rights: "*[...]of course you can share content. Right? I guess you literally generate content and [sic] can be of course violation of human rights.*" (P06) Few participants mentioned that it was hard to mitigate the spread of misinformation through software because they do not monitor or control messages because "*it's difficult when you think about private messaging to have end-to-end encryption, and also the fact that we are not the mediators of the messages.*" (P01) Furthermore, participants P01, P10, and P07 mentioned that distinguishing misinformation is the responsibility of their users. Participant P07 additionally highlighted that their platform could be used to "*store, to distribute illegal content, or for illegal activity by selling drugs or other things,*" but it is difficult to "*control who uses it.*" Participant P04 mentioned that their software could be used to create fake signatures. However, a few participants were also actively working on mitigating misinformation; P01 enables marking messages as spam and blocking users in their messenger, and P11 mentioned their browser has strict guidelines to block problematic and misleading ads. Finally, two participants (P01, P07) mentioned that more research and strategies are required to handle misinformation on communication platforms.

Responsibility is shifted to administrators and moderators. Some participants mentioned that the software could be disempowering to their users: "*[...]the future of giving out the login without the password moves to power up in the hierarchy,*" (P02) if corporations enforce the code signing tool. Further, P02 mentioned that they will implement a feature that lets managers see which websites employees are logging into, and this may be misused: "*[...] it's not visible for us*

as a provider, but it will be visible for the project managers or CEOs. So that may be misused for surveillance of employees." (P02) Additionally, P14 stated that the responsibility of protecting the users shifts away from the people creating the tools to people who self-host instances, thus emphasizing challenges that come with autonomy granted to administrators:

> "*Again, there is a lot of responsibility and accountability taken away from the company itself building a tool that people can set up and use however they would like. If they choose to violate human rights with the tool, they can't do anything about it.*" — P14, File Sharing and Collaboration

Legal obligations versus user privacy. One participant highlighted a concern that they are compelled to surrender user data to law enforcement when they have a court order: "*We do have to provide the police with the data that they request, and they sometimes request data of specific users. Well, that has to be with a court order, but in case they are surveilled, and we have to deliver every information that we have about them, and that's mostly encrypted data.*" (P03)

Harmful actors and criminals can leverage the software. Participants said some consequences are inevitable and depend on how their users use the software, and they cannot control or moderate content on their software. A majority of them acknowledged that adversarial actors could use software to cause harm. Few participants (Browser, Stalkerware Detection) suggested that the use of their software might aggravate domestic abuse or IPV further. Because using the software "*may look like you're hiding something*" (P05) if the abuser finds the software installed on the victim's device. To mitigate this, they explained that they left the decision to delete an application to the users.

> "*It is possible, for example, someone gets a warning from our app, decides to remove the app, and then [sic] beaten or something right by an abusive partner, so our app doesn't automatically remove the app, it gives them the ability and the information that they need to make that decision.*" — P05, Stalkerware detection

P02 mentioned that the software could be used for employee surveillance. Cyberbullying was also mentioned by P03, who said that abusers and hackers use their encrypted E-mail service. These accounts can be reported through an abuse reporting channel and are disabled by the service. P03 also mentioned that since they enable "*anonymous signup*", a lot of spammers use their servers to send spam emails; therefore, they "*have several measures like a waiting period for people to be able to send emails; we collect and we maintain for some time the IP addresses that were used to register with the account. So we can detect suspicious activity there.*" Harmful actors use social networking sites to spread hateful messages; P06, a developer of a social networking site, stated, "*[. . .]being open source technology, privacy-first technology, it does attract certain groups which feel like anonymity gives them a space to do maybe not always lawful things right, so we did have attraction, for example, from ISIS, we do have the attraction from these groups.*" Similarly, P07 mentioned that software used to distribute illegal content or for illegal activities like selling drugs, and cyberwarfare. They further stated:

> "*I feel that it is a little difficult to pinpoint what exactly they are using it for. But it would be naive to think that it is not used in that (harmful) way.*" — P07, Anonymity Network and browser

Downloading the software can be harmful to users in hostile environments. Some participants mentioned that in some cases, the very use of their software or the download of it could be incriminating in certain countries and circumstances. Participant P01 emphasized this by stating the example of installing a secure messaging tool in Turkey: "*There have been cases in Istanbul and Ankara where people have been put into prison temporarily just for having installed a secure messaging tool, even without any content, so just the fact that you have it.*" Participant P09 who works for a VPN said "*I haven't directly, we haven't directly heard this*", but mentioned that they are aware that the use of VPN can be "*incriminating in some places.*" Participant P11 mentioned that using their browser might be banned on some networks and workplaces.

4.3.4 Aspects that do not affect all S&P software

Some aspects do not apply to all software. Addiction as a factor was discussed as irrelevant for many software types, including VPN, code signing, password manager, and secure translation. Many participants mentioned that their software could not be used to spread **misinformation**. These tools included VPN, translation, password manager, and stalkerware detection. This was because "*there's no way for people to communicate to our app or anything like that.*" (P05) Similarly, participants felt that software to accomplish very specific tasks like code signing tool and translation tool were **not distracting**. "*My products are not distracting. All you need to do is just get it on your phone and then tap on it whenever you want to use it and then just send a request to the software and it's going to respond to you, it's not going to distract you anyway.*" (P08) A majority of the participants mentioned that they see no difference in **negative environmental impacts** compared to any other software. Participants P02 and P03 mentioned that they try to have a positive impact by using renewable or green energy sources. Participants also could not think of how their software could **violate human rights**

and were unaware of **laws in different countries**. Participants P01 and P05 mentioned that their software contributes to strengthening human rights.

Most participants had at least considered **disempowerment** and the question of whether technology can impede users' decision-making. Conversely, a few participants mentioned that they had never considered disempowerment and how their technology can influence users' decision-making. The participants also indicated that it is quite hard to think about this question and had nothing to say about it: "*I couldn't say, I didn't have discussions about this. If I had to judge as much as any other communication platform, you will get challenges there.*" (P06) Participant P07 also never considered **physical and mental health** impacts of their technology on users.

4.4 Factors that help or hinder anticipating and mitigating consequences

4.4.1 Factors that hinder anticipating and mitigating unintended consequences

Shared accountability is assumed, and no one is accountable for anticipating unintended consequences. Many participants mentioned that there was no specific person in their team to look into unintended consequences. A few participants mentioned that the CEO, CTO, and leadership teams must be responsible for creating a product vision to lower negative impacts. Participant P11, who was from a larger organization, mentioned that there is no specific person assigned to look into it, "*[...]think it's not defined, to be very honest. The ethical impacts get rolled into privacy, so me and our DPO (Data Privacy Officer). But it's often quite tricky because that's not our mandate.*" (P11) Additionally, participants mentioned that everyone on the team should be held accountable. Participant P04, a volunteer one-person team, mentioned that they do not have the capacity to work on it and expect collaborators on GitHub to open issues. In addition, smaller teams faced more issues because they lacked the resources to conduct user research activities. Participant P13 mentioned that it is a collective responsibility of everyone on the team:

> "*I would say it's a very much a collective effort. Everyone does it. When we discuss features or how something is going to be implemented, there is always a high awareness of how it could be misused or how do we protect the users' data in a more secure way, etc.* " — P13, Encrypted Office Suite

Some participants believe that there are no negative consequences because their intentions are good. Participants stated that they did not observe any direct negative impacts due to the positive nature of their software. Participant P08 could not imagine any negative consequences of downloading the software and emphasized that they protected their users: "*No, no, no, not at all. What we're doing here is trying to protect your system. We're trying to protect your phone from harm.*" (P08) In addition, participants mentioned that their products positively impact people by empowering them and encrypting their messages. Some also mentioned that PETs like VPN could be incriminating, but in comparison, they have a positive impact:

> "*I think in comparison to other privacy-enhancing tools, it's a pretty small impact. I mean, basically [password manager] is a tool for being more productive. Saving time. Storing passwords.*" — P02, Password Manager

Participants also stated that certain consequences are not a direct consequence of downloading their software but the problem of third-party applications and providers. Participant P12, mentioned, "*[...]if some text offers offensive translations for certain languages, it is the fault of google translate and not my software.*"

Lack of awareness and knowledge at an individual and an organizational level. A lack of organizational support hindered participants from systematically anticipating unintended consequences. Participant P11 stated: "*I think it's more that we don't have the have the mandate to do it. We don't really have organizational buy-in to stop a feature from rolling out because of a perceived ethical risk. There is no one really who full-time has the job title to think about ethical issues.*" (P11) Furthermore, P11 specified that they should have an ethical overview similar to privacy overviews. Additionally, some participants did not receive formal training in ethics or unintended consequences of digital software. Few participants mentioned that they have read about ethics and try to educate themselves through books and other media. Few participants mentioned having some sort of ethics training via universities and company training. Additionally, the reported lack of team diversity may also hinder considerations of diverse users and their needs and unintended consequences in general [56].

Privacy trade-offs can impede the identification of certain unintended consequences. Prioritizing user privacy introduces challenges while identifying unintended consequences: S&P software creators are limited in collecting user feedback that might identify their users and also limited in their capacity to moderate content. This highlights challenges with balancing user autonomy and identifying issues that might harm them. For instance, when it comes to topics like misinformation and communication between their users, participants mentioned that since they are a "*privacy first technology*" (P06), they "*don't have central control*" (P14) to oversee conversations or content. Further, P01 stated that, since they are a private messaging software and their core value is end-to-end encryption, they are unsure how to solve the spread of

disinformation: "*to be honest, I think in the space of private messaging, how to prevent the spread of dubious information is difficult. I'm not sure that anyone has the answer to that.*" Participants reported that the privacy-first focus also interferes with collecting user input during development because they refrain from collecting any personal information about their users, which might reveal the demographics of their users: "*The first thing is that the feedback that we have or any kind of data that we get from the users is anonymous because we don't want to reveal, identify information that can pinpoint specific users.*" (P10) P11 also mentioned that though they can think of proactively involving users to get input through measurement studies, they still "*[. . .]don't want to just push for having as much measurements as we can because that will hit our core values, which is just preserving privacy.*"

4.4.2 Systematic assessments help to be proactive

When asked about what unintended consequences they anticipate without using systematic probing, participants initially had a narrow view of unintended consequences. Probing the aspects we extracted from toolkits helped them think through broader consequences, many of which they had not previously considered.

5 Discussion

In this paper, we explored how S&P software creators assess and mitigate unintended consequences through 14 semi-structured expert interviews. We also identified whether they use any toolkits and investigated their current processes. We found that these experts are highly motivated to create a positive impact by protecting people's privacy and security. However, they do not systematically anticipate and mitigate unintended consequences. Most participants mentioned that they hoped users would identify and report issues. We now discuss the implications of waiting for users to identify issues and provide recommendations for better impact assessment processes.

Relying on user feedback burdens them and puts them at risk. As discussed in Section 4.2.2, participants highlighted reliance on reactive approaches—acting upon user feedback or reports about issues they face. To receive feedback they use methods like tracking user feedback, and opening issues on GitHub. The term *technical debt* was coined by Ward Cunningham in 1992 [72], which describes the price software teams pay for prioritizing speedy delivery over perfect code. Similarly, *ethical debt* is the idea that when software development teams lack of oversight for the ethical and societal implications of their work, they may create harm that will have to be addressed later [73], [74]. While waiting for users to bring up and then resolving the issues works in some cases, this approach may be harmful, especially for vulnerable populations who might be experiencing threats of exacerbated quantities and qualities.

We think reactivity can be problematic because of the additional responsibility on marginalized users, expecting them to report issues they encounter during usage in addition to enduring partially usable tools. McDonald *et al*. state that there is too much onus on users in expecting them to *make do* with tools they have rather than soliciting users' feedback right at the beginning [75]. Negative experiences may be detrimental to users and put their privacy and security at risk, as they may abandon S&P tools and resort to less secure practices. Previous work has shown that negative experiences such as bad usability and inconvenience lead to the abandonment of privacy tools [76].

Proactively considering user threat models during the initial phases of software design can identify vulnerabilities quickly and thereby avoid adverse impacts on end users. We recommend technology designers be proactive in anticipating unintended consequences and negative impacts on users, during the initial stages of development. They may consider involving users and user input at the design stage.

Create incentives to avoid the shift of risk to users: a moral hazard problem. Chua *et al.* discuss the *moral hazard problem*, where software creators do not have an incentive to take action to avoid unintended harm because it does not affect them negatively [12]. Similar to our discussion above, they also emphasize that when harms are not addressed by software creators who have the power to do so, the risks are shifted to users who might be unprepared to bear the consequences. For instance, in case of negative experiences users might abandon the software, resorting to less secure means. Chua *et al.* further emphasize that software creators may not be incentivized to avoid unintended harms that do not directly impact them [12]. Therefore, creating incentives might be crucial. Software creators may consider making prevention of unintended harms a strategic priority. Harm prevention must be a fundamental part of the planning and designing of the software. Additionally, proactive harm mitigation strategies might incentivize software creators in attracting and retaining users, reinforcing trust of users, and maintaining a competitive advantage.

Systematic toolkits help in uniting the team and providing a more holistic view of potential unintended consequences. Participants had no systematic process to anticipate adverse outcomes; instead, they relied on user feedback or focused only on privacy and security aspects. Their current processes may not be sufficient to anticipate unintended consequences; systematic toolkits provide a comprehensive overview of possible harms. Currently, many participants reported not considering how their products may be misused in specific ways, as reported in Section 4.3. While previous work criticizes the limited view of systematic ethical toolkits and checklists [77]–[79], they can still serve as excellent probes and evocative heuristics that elicit thinking about harms that

may be otherwise overlooked [49], [70]. One major drawback of toolkits and checklists is that they are not prescriptive in how to mitigate harms [50]. We find them a good starting point to think about broader ethical and societal issues; due to the overwhelming number of available toolkits and checklists, their generality, and lack of prescriptive support, none seemed like in itself, it would effectively help mitigate harm. We implore that future studies explore effective preventative strategies for mitigating negative consequences of S&P software.

Anticipating harm is only a part of the solution; software professionals must be empowered to fix problems. Anticipating or identifying unintended consequences is required, but often insufficient to mitigate them. S&P software creators must be equipped to mitigate the problems they identify. We find that stakeholders not in leadership roles may not be empowered to bring up issues in their development process. A lack of resources and authority may result in an inability to mitigate unintended consequences. As discussed in Section 4.4.1, a majority of participants expected their management teams to take up the responsibility of mitigating unintended consequences. In this line, Widder *et al*. highlight that there is a disparity between anticipating and mitigating ethical issues [80]. Even if software creators have the knowledge and awareness to anticipate ethical issues, they often lack the power to act on them. We recommend that S&P organizations incorporate ethics into the core job responsibilities of their employees and foster an organizational culture that provides them with the resources and power to raise and mitigate unintended consequences.

Experts are more focused on their roles, and there is no accountability to address unintended consequences. Our study finds that participants are motivated to create a positive impact—within the confines of current knowledge and capacity—they fulfill the obligations of their official roles and responsibilities. Most often, they balance different roles and priorities. Even if they may recognize some ethical dilemmas, they may be prevented from addressing those consequences, not due to ignorance, but by the formal specifications of their roles (see Section 4.4.1). We found a lack of shared team or organizational goals to reduce or mitigate harm, as opposed to shared goals of privacy. Some participants mentioned that leadership should be responsible for making this a vision of the organization. We implore organizations to establish new roles or expand on formal role responsibilities and create a shared mission for the entire organization to uphold ethical values beyond privacy.

Without dedicated resources, accountability, and organizational backing, negative issues cannot be appropriately addressed through design changes. Organizations should actively play a role in dedicating resources to anticipating ethical impacts and issues right from the projects' inception and follow through with mitigating issues as they are identified. Furthermore, defining harm reduction and fixing ethical issues as business metrics like Objectives and Key Results (OKRs) might allow teams to do their due diligence in systematically anticipating and mitigating unintended consequences. Additionally, organizations should consider forming governance bodies for ethical review of software similar to privacy overviews; this would help to place equal significance on software impacts as other governance aspects, as discussed in Section 4.4.

Create resources and materials for smaller teams and teams where everyone volunteers. Large organizations may readily afford to invest time and resources into forming governance bodies and invest in large-scale efforts to look into unintended consequences. In contrast, smaller teams and individuals often lack the luxury to devote their time and resources to solving ethical issues. As these approaches can be time-consuming and resource-intensive, smaller organizations may look into building support communities with users and stakeholders in different organizations with aligned goals. Open source communities have historically helped foster support and faster iterations of products; therefore, similar communities can also be used to echo software impacts. The security community should also investigate ways to build software with ethics in mind and focus on unintended consequences within and outside the community. Importantly, when educators teach security, they should include teaching unintended consequences.

Develop formal guidelines and training to raise awareness. We also identified that most participants lacked formal education in ethics. While ethics frameworks, toolkits, and checklists are pervasive, they are not widely adopted or used in software development teams, likely due to the lack of ethics in technology education [81], [82]. As discussed in Section 4.4.1, S&P software organizations may want to consider creating a review process for examining ethical considerations and potential software harms before a feature is released to the broader public. Existing review processes for security and privacy may inform these review processes for software harms. Ethics discussions must be a priority and not an afterthought contingent on available time. Fiesler *et al.* suggest integrating ethics into technical classes beyond stand-alone ethics classes. McDonald *et al.* show that students' thinking about social justice improves over time if they are introduced to ethics over a longer time, therefore advocating for more time and exposure to ethics topics [83]. Kohno *et al.* emphasized using ethical scenarios to reason about security research implications [84]. Further, not everyone in software development has access to formal education; therefore, organizations should consider company training on ethics and software harm.

Privacy-moderation trade-off remains a constraint—posing many ethical dilemmas like user data collection and content moderation. Although collecting user feedback is essential, monitoring usage and gathering personal information or feedback can conflict with privacy principles. Since this remains inevitable, we recommend that future research develop guide-

lines and resources to balance the privacy-moderation trade-off.

Training and guidelines for self-hosters and moderators. As mentioned in Section 4.3.3, moderators and self-hosters often directly interact with end users. We find that developers usually create tools and hand over the responsibility of moderating to self-hosters, and they may most often encounter ethical dilemmas in the software instances they create and moderate. Moderators and self-hosters need to be empowered and provided with training, guidelines, and support on handling negative impacts and ethical dilemmas they may encounter. Further, future work can also look into how self-hosters deal with negative and challenging scenarios—identifying which resources should be provided to equip them better to mitigate harm or unintended consequences that cannot be prevented through software design or that clash with their privacy values.

6 Conclusion

We explored how developers of security and privacy software consider the unintended impacts their software can have, finding that participants want to create beneficial software but lack the resources to address unintended consequences systematically. We find that S&P projects often rely on user feedback to identify problems and mitigate risks and often also face ethical dilemmas due to conflicts stemming from balancing trade-offs between privacy values and moderation. We argue that systematically assessing risks, as we modeled in our interviews, can lessen the burden on users. We make recommendations for organizations, educators, and the S&P research community to take action to assess and mitigate unintended consequences of S&P software.

7 Acknowledgment

We want to express our deepest appreciation to Anna Lena Rotthaler, Kelly Almon Mumba, Anastassija Kostan, Marcel(le) Fourné, and Swar Joshi for their feedback and interest in this research. We thank our participants for participating in the study and providing valuable input. We also thank all the anonymous reviewers and shepherds for their thoughtful comments and feedback.

References

[1] A. Bruckman, "'Have you thought about ...' talking about ethical implications of research," *Communications of the ACM*, vol. 63, no. 9, pp. 38–40, 2020.

[2] R. K. Merton, "The unanticipated consequences of purposive social action," *American sociological review*, vol. 1, no. 6, pp. 894–904, 1936.

[3] C. Machado, B. Kira, V. Narayanan, B. Kollanyi, and P. Howard, "A study of misinformation in WhatsApp groups with a focus on the brazilian presidential elections.," in *Companion proceedings of the 2019 World Wide Web conference*, 2019, pp. 1013–1019.

[4] B. Levine, "Shining light on internet-based crimes against children," Santa Clara, CA: USENIX Association, Aug. 2019.

[5] A. Daffalla, L. Simko, T. Kohno, and A. G. Bardas, "Defensive technology use by political activists during the Sudanese revolution," in *S&P 2021: IEEE Symposium on Security and Privacy*, 2021.

[6] Y. Wang, "Inclusive security and privacy," *IEEE Security & Privacy*, vol. 16, no. 4, pp. 82–87, 2018.

[7] N. Saxena and J. H. Watt, "Authentication technologies for the blind or visually impaired," in *Proceedings of the USENIX Workshop on Hot Topics in Security (HotSec)*, vol. 9, 2009, p. 130.

[8] B. Dosono, J. Hayes, and Y. Wang, ""I'm stuck!": A contextual inquiry of people with visual impairments in authentication," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, 2015, pp. 151–168.

[9] M. Dorin and S. Montenegro, "Ethical lapses create complicated and problematic software," in *2021 IEEE/ACM 2nd International Workshop on Ethics in Software Engineering Research and Practice (SEthics)*, 2021, pp. 1–4.

[10] K. Do, R. Y. Pang, J. Jiang, and K. Reinecke, ""that's important, but...:" how computer science researchers anticipate unintended consequences of their research innovations," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–16.

[11] R. Wash, E. Rader, K. Vaniea, and M. Rizor, "Out of the loop: How automated software updates cause unintended security consequences," in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, 2014, pp. 89–104.

[12] Y. T. Chua, S. Parkin, M. Edwards, *et al.*, "Identifying unintended harms of cybersecurity countermeasures," in *2019 APWG Symposium on Electronic Crime Research (eCrime)*, IEEE, 2019, pp. 1–15.

[13] GOV.UK, *Data Ethics Framework*, https://www.gov.uk/government/publications/data-ethics-framework, Online; accessed Nov 8, 2021.

[14] I. for the Future and O. Network, *Ethical OS toolkit*, https://ethicalos.org/.

[15] D. C. S. L. (DCSL), *Digital Impact Toolkit*, https://digitalimpact.io/toolkit/.

[16] D. Anderson, J. Bonaguro, M. McKinney, A. Nicklin, and J. Wiseman, *Ethics and Algortithms toolkit*, https://ethicstoolkit.ai/.

[17] M. A. Madaio, L. Stark, J. Wortman Vaughan, and H. Wallach, "Co-designing checklists to understand organizational challenges and opportunities around fairness in ai," in *Proceedings of the 2020 CHI conference on human factors in computing systems*, 2020, pp. 1–14.

[18] CIGREF, *Digital Ethics Guide for professionals of digital age*, Online; accessed Nov 8, 2021.

[19] A. G. Pillai, A. Baki Kocaballi, T. Wah Leong, *et al.*, "Co-designing resources for ethics education in hci," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–5.

[20] E. Commission, *European Commission Ethics guidelines for trustworthy AI*, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[21] N. Aayog, *World Economic forum ethics for responsible AI*, https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.

[22] ACM Code 2018 Task Force, *ACM code of ethics and professional conduct*, https://www.acm.org/code-of-ethics, 2018.

[23] M. Fassl, S. Anell, S. Houy, M. Lindorfer, and K. Krombholz, "Comparing user perceptions of anti-stalkerware apps with the technical reality," in *Proceedings of the Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, USENIX Association, 2022.

[24] M. Namara, D. Wilkinson, K. Caine, and B. P. Knijnenburg, "Emotional and practical considerations towards the adoption and abandonment of VPNs as a privacy-enhancing technology," *Proceedings on Privacy Enhancing Technologies*, vol. 1, pp. 83–102, 2020.

[25] O. Akgul, R. Roberts, M. Namara, D. Levin, and M. L. Mazurek, "Investigating influencer VPN ads on youtube," in *2022 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2022, pp. 876–892.

[26] K. Renaud, G. Johnson, and J. Ophoff, "Dyslexia and password usage: Accessibility in authentication design," in *Human Aspects of Information Security and Assurance: 14th IFIP WG 11.12 International Symposium, HAISA 2020, Mytilene, Lesbos, Greece, July 8–10, 2020, Proceedings 14*, Springer, 2020, pp. 259–268.

[27] N. Huaman, S. Amft, M. Oltrogge, Y. Acar, and S. Fahl, "They would do better if they worked together: Interaction problems between password managers and the web," *IEEE Security & Privacy*, vol. 20, no. 2, pp. 49–60, 2022.

[28] S. Chiasson, P. C. van Oorschot, and R. Biddle, "A usability study and critique of two password managers," in *USENIX Security Symposium*, vol. 15, 2006, pp. 1–16.

[29] S. Sheng, L. Broderick, C. A. Koranda, and J. J. Hyland, "Why johnny still can't encrypt: Evaluating the usability of email encryption software," in *Symposium on usable privacy and security*, ACM, 2006, pp. 3–4.

[30] N. Huaman, S. Amft, M. Oltrogge, Y. Acar, and S. Fahl, "They would do better if they worked together: The case of interaction problems between password managers and websites," in *42nd IEEE Symposium on Security and Privacy, IEEE S&P 2021, May 24-27, 2021*, IEEE Computer Society, 2021.

[31] A. Gallardo, H. Kim, T. Li, L. Bauer, and L. Cranor, "Detecting {iphone} security compromise in simulated stalking scenarios: Strategies and obstacles," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, 2022, pp. 291–312.

[32] D. Woodlock, "The abuse of technology in domestic violence and stalking," *Violence against women*, vol. 23, no. 5, pp. 584–602, 2017.

[33] D. M. Douglas, "Doxing: A conceptual analysis," *Ethics and information technology*, vol. 18, no. 3, pp. 199–210, 2016.

[34] A. Shimizu, "Domestic violence in the digital age: Towards the creation of a comprehensive cyberstalking statute," *Berkeley J. Gender L. & Just.*, vol. 28, p. 116, 2013.

[35] B. Eterovic-Soric, K.-K. R. Choo, H. Ashman, and S. Mubarak, "Stalking the stalkers – detecting and deterring stalking behaviours using technology: A review," *Computers & Security*, vol. 70, pp. 278–289, 2017.

[36] M. Wei, E. Zeng, T. Kohno, and F. Roesner, "Anti-Privacy and Anti-Security advice on TikTok: Case studies of Technology-Enabled surveillance and control in intimate partner and Parent-Child relationships," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, 2022, pp. 447–462.

[37] N. McDonald, K. Badillo-Urquiola, M. G. Ames, *et al.*, "Privacy and power: Acknowledging the importance of privacy research and design for vulnerable populations," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '20, Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–8.

[38] S. Sannon and A. Forte, "Privacy research with marginalized groups: What we know, what's needed, and what's next," *arXiv preprint arXiv:2206.15037*, 2022.

[39] L. Simko, A. Lerner, S. Ibtasam, F. Roesner, and T. Kohno, "Computer security and privacy for refugees in the United States," in *S&P 2018: IEEE Symposium on Security and Privacy*, 2018.

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

[41] A. McDonald, C. Barwulor, M. L. Mazurek, F. Schaub, and E. M. Redmiles, ""it's stressful having all these phones": Investigating sex workers' safety goals, risks, and practices online," in *Security 2021: USENIX Security Symposium*, 2021.

[42] A. Cavoukian *et al.*, "Privacy by design: The 7 foundational principles," *Information and privacy commissioner of Ontario, Canada*, vol. 5, p. 2009, 2009.

[43] M. Rotenberg, "Fair information practices and the architecture of privacy (what Larry doesn't get)," *Stan. Tech. L. Rev.*, p. 1, 2001.

[44] D. Wright and P. De Hert, *Privacy impact assessment*. Springer, 2012, vol. 6.

[45] N. Desk, *Turkey criminalizes Signal messaging app*, https://www.weeklyblitz.net/tech/turkey-criminalizes-signal-messaging-app/, Online; accessed September 24, 2022.

[46] B. Friedman, "Value-sensitive design," *interactions*, vol. 3, no. 6, pp. 16–23, 1996.

[47] J. Auger, "Speculative design: Crafting the speculation," *Digital Creativity*, vol. 24, no. 1, pp. 11–35, 2013.

[48] S. Elsayed-Ali, S. E. Berger, V. F. D. Santana, and J. C. Becerra Sandoval, "Responsible & inclusive cards: An online card tool to promote critical reflection in technology industry work practices," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–14.

[49] A. Borning and M. Muller, "Next steps for value sensitive design," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2012, pp. 1125–1134.

[50] R. Y. Wong, M. A. Madaio, and N. Merrill, "Seeing like a toolkit: How toolkits envision the work of ai ethics," *Proceedings of the ACM on Human-Computer Interaction*, vol. 7, no. CSCW1, pp. 1–27, 2023.

[51] I. Hadar, T. Hasson, O. Ayalon, *et al.*, "Privacy by designers: Software developers' privacy mindset," in *Proceedings of the 40th International Conference on Software Engineering*, ser. ICSE '18, Gothenburg, Sweden: Association for Computing Machinery, 2018, p. 396.

[52] M. Gutfleisch, J. H. Klemmer, N. Busch, Y. Acar, M. A. Sasse, and S. Fahl, "How does usable security (not) end up in software products? results from a qualitative interview study," in *43rd IEEE Symposium on Security and Privacy, IEEE S&P*, 2022, pp. 22–26.

[53] J. D. Bustard, "Improving student engagement in the study of professional ethics: Concepts and an example in cyber security," *Science and Engineering Ethics*, vol. 24, pp. 683–698, 2018.

[54] R. Balebako, A. Marsh, J. Lin, J. I. Hong, and L. F. Cranor, "The privacy and security behaviors of smartphone app developers," *DOI: http://dx. doi. org/10.1184*, vol. 1, 2014.

[55] D. Adams, A. Bah, C. Barwulor, N. Musaby, K. Pitkin, and E. M. Redmiles, "Ethics emerging: The story of privacy and security perceptions in virtual reality," in *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, 2018, pp. 427–442.

[56] N. McDonald and S. Pan, "Intersectional AI: A study of how information science students think about ethics and their impact," *Proc. ACM Hum.-Comput. Interact.*, vol. 4, no. CSCW2, 2020.

[57] K. R. Fulton, S. Katcher, K. Song, *et al.*, "Vulnerability discovery for all: Experiences of marginalization in vulnerability discovery," in *32nd USENIX Security Symposium (USENIX Security). USENIX Association*, 2023.

[58] A. Radermacher and G. Walia, "Gaps between industry expectations and the abilities of graduates," in *Proceeding of the 44th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '13, Denver, Colorado, USA: Association for Computing Machinery, 2013, pp. 525–530.

[59] W. Groeneveld, J. Vennekens, and K. Aerts, "Software engineering education beyond the technical: A systematic literature review," *arXiv preprint arXiv:1910.09865*, 2019.

[60] D. Botta, R. Werlinger, A. Gagné, *et al.*, "Towards understanding IT security professionals and their tools," in *Proceedings of the 3rd Symposium on Usable Privacy and Security*, ser. SOUPS '07, Pittsburgh, Pennsylvania, USA: Association for Computing Machinery, 2007, pp. 100–111.

[61] L. Bauer, L. F. Cranor, R. W. Reeder, M. K. Reiter, and K. Vaniea, "Real life challenges in access-control management," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 899–908.

[62] M. Silic and A. Back, "Information security and open source dual use security software: Trust paradox," in *Open Source Software: Quality Verification*, E. Petrinja, G. Succi, N. El Ioini, and A. Sillitti, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 194–206.

[63] R. A. Bridges, M. D. Iannacone, J. R. Goodall, and J. M. Beaver, "How do information security workers use host data? a summary of interviews with security analysts," *arXiv preprint arXiv:1812.02867*, 2018.

[64] D. Wermke, N. Wöhler, J. H. Klemmer, M. Fourné, Y. Acar, and S. Fahl, "Committed to trust: A qualitative study on security & trust in open source software projects," in *Proceedings of the 43rd IEEE Symposium on Security and Privacy*, IEEE Computer Society, 2022.

[65] D. Wermke, J. H. Klemmer, N. Wöhler, *et al.*, "" always contribute back": A qualitative study on security challenges of the open source supply chain," in *2023 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2023, pp. 1545–1560.

[66] S. Takai and K. Ishii, "A use of subjective clustering to support affinity diagram results in customer needs analysis," *Concurrent Engineering*, vol. 18, no. 2, pp. 101–109, 2010.

[67] J. S. for AI, *Japanse Society for AI ethics guidelines*, http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf.

[68] A. Lowe, A. C. Norris, A. J. Farris, and D. R. Babbage, "Quantifying thematic saturation in qualitative data analysis," *Field methods*, vol. 30, no. 3, pp. 191–207, 2018.

[69] M. A. Cascio, E. Lee, N. Vaudrin, and D. A. Freedman, "A team-based approach to open coding: Considerations for creating intercoder consensus," *Field Methods*, vol. 31, no. 2, pp. 116–130, 2019.

[70] N. McDonald, S. Schoenebeck, and A. Forte, "Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice," *Proc. ACM Hum.-Comput. Interact.*, vol. 3, no. CSCW, Nov. 2019.

[71] N. Thorogood and J. Green, "Qualitative methods for health research," *Qualitative methods for health research*, pp. 1–440, 2018.

[72] F. Buschmann, "To pay or not to pay technical debt," *IEEE software*, vol. 28, no. 6, pp. 29–31, 2011.

[73] N. Cristianini, *Shortcuts to artificial intelligence*, 2019.

[74] C. Fiesler, *Ethical tech starts with addressing ethical debt*, 2020.

[75] N. McDonald, A. K. Massey, and F. Hamidi, "Ai-enhanced adaptive assistive technologies: Methods for ai design justice.," *IEEE Data Eng. Bull.*, vol. 44, no. 4, pp. 3–13, 2021.

[76] Y. Zou, K. Roundy, A. Tamersoy, S. Shintre, J. Roturier, and F. Schaub, "Examining the adoption and abandonment of security, privacy, and identity theft protection practices," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–15.

[77] A. H. Kiran, N. Oudshoorn, and P.-P. Verbeek, "Beyond checklists: Toward an ethical-constructive technology assessment," *Journal of responsible innovation*, vol. 2, no. 1, pp. 5–19, 2015.

[78] F. Allhoff, P. Lin, J. Moor, and J. Weckert, "Ethics of human enhancement: 25 questions & answers," *Studies in Ethics, Law, and Technology*, vol. 4, no. 1, 2010.

[79] C. A. Le Dantec, E. S. Poole, and S. P. Wyche, "Values as lived experience: Evolving value sensitive design in support of value discovery," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2009, pp. 1141–1150.

[80] D. G. Widder, D. Zhen, L. Dabbish, and J. Herbsleb, "It's about power: What ethical concerns do software engineers have, and what do they (feel they can) do about them?" In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, ser. FAccT '23, Chicago, IL, USA: Association for Computing Machinery, 2023, pp. 467–479.

[81] P. Kumar, J. Vitak, M. Chetty, *et al.*, "Co-designing online privacy-related games and stories with children," in *CHI 2018: ACM Conference on Human Factors in Computing Systems*, 2018.

[82] N. Garrett, N. Beard, and C. Fiesler, "More than "if time allows" the role of ethics in ai education," in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2020, pp. 272–278.

[83] N. McDonald, A. Akinsiku, J. Hunter-Cevera, *et al.*, "Responsible computing: A longitudinal study of a peer-led ethics learning framework," *ACM Transactions on Computing Education (TOCE)*, vol. 22, no. 4, pp. 1–21, 2022.

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

A Interview Guide and Screening Questions

Screening questions

1. Do you work on developing security and privacy-enhancing software like VPN, TOR, Password Managers, and Secure messaging?
2. Is the core Value or mission of your software or organization to enhance privacy and/or security?
3. Please describe the kind of software you work with. (VPN, Secure Messaging, TOR.. etc)
4. What gender do you identify as (Man, Woman, Non-binary, Prefer not to answer, Self- describe)
5. How many years of development experience do you have? (self-described)
6. What is your role within the development team or your role in the organization? (self-described)
7. Do you have previous experience with ethics? (Yes, No, Prefer not to answer, other(please specify))
8. Was ethics a part of your training or studies? (Yes, No, Prefer not to answer, other(please specify)
9. Do you wish to get a $80 gift card as a thank you for participating in this study?

Introductory questions

1. What product are you working on?

2. Could you describe the core value of your product? What is the key problem/ need you are trying to address?
3. What is your role within the development team?
4. Who are the primary users of your product? Who benefits from your product? And how?
5. Were users consulted or involved in the development process? If so, how was their feedback incorporated into the product development?
    (a) How did you ensure that a good representative sample of your users was involved in the development process?
    (b) Can you think of other secondary users, bystanders, or unintentional users who might also use your product?
    (c) How did you factor in primary, secondary, bystanders, and unintentional users within the development process?
6. How did you ensure that the product fits the problems/ needs of the users?
    (a) *Probe if needed:* How did you measure if the product fits the needs of the users? (eg. using metrics and user tests)
    (b) How have you communicated your understanding of the users' needs with them?

Negative and Unintended consequences
For each question, ask for mitigation strategies in place.

1. **Anticipated negative impacts:** Have you or your company observed your product having any negative impacts? *(negatively)*? *Who is impacted, for instance:* individuals, businesses, environment, groups of people, operators, etc.
2. **Potential harm:** Can you imagine people being harmed by using your software? (eg. issues with the government, partner abuse, etc.)
3. **Barriers to access:** Are there specific groups or communities who do not have access to the product? What are the different barriers to access?
4. **Biases:** How did you ensure that your software is unbiased and treats all your users fairly?
    (a) *If mitigation is not mentioned:* How were risks of biases identified and mitigated? (Selection bias, Historic Bias, Individual biases)
    (b) *Probe:* What about paid/ free users?
5. **Accessibility:** How do you design with the accessibility needs of people with disabilities in mind?
6. **Disempowerment:** Have you thought about disempowerment? Does the technology replace or weaken users' authority in decision-making? (eg. Blackbox AI Technology, Technology replacing people/ business)
7. **Misinformation:** Do you think your tech be used to generate or spread misinformation to create political distrust or social unrest?
    (a) *If they say no/ if it is a social tool then:* How do you think someone could use this technology to spread misinformation?
8. **Addiction:** How was addiction factored into the design of the software? Do you think your software could be addicting for users?
9. **Distraction:** Do you think your software fits into the natural workflow of your users? How do you think your software could be distracting to users?
10. **Physical and Mental health impact:** How does your software affect the physical and mental well-being of your users?
11. **Environment Impact:** What do you think are the environmental impacts of the solution?
12. **Human Rights Violation:** Do you think your software could violate the respect for the human rights of your users?

Data Privacy and Security Risks

1. **Security:** How did you assess potential forms of attacks and security threats to which the solution could be vulnerable?
2. **Data Protection:** How does your software protect users' personal data?
    (a) *Consent:* How did you communicate to your users about what data is collected and how it is used?
    (b) *If data is collected:* How do you minimize data collection?
    (c) Will the raw data be shared with additional external partners?
    (d) *Privacy impact:* How did you evaluate the privacy impact of the project? Does a Data Privacy Officer (DPO) or similar exist? Did you involve this person at an early stage in the process?
3. **Data Governance:** What protocols, processes, and procedures did you follow to manage and ensure proper data governance?
4. **Tracking and Surveillance:** How might your product be used for tracking and surveillance of targeted individuals?
    (a) Whom would they track, why, and do you want your tech to be used in this way?

Organizational factors, Ethics frameworks, and tools

1. What does your development team look like? What are the different roles of people within the development team? (e.g., are there designers, legal, etc.)
2. Is the development team a representative of the social and gender diversity of society?
3. Who takes the responsibility to assess ethical and societal impacts? Who do you think should be accountable?
4. How did you identify and address ethical risks? Are there any measures in place to assess and mitigate negative consequences?
    (a) *If frameworks are not mentioned:* Do you use any frameworks or tools to assess ethical risks?
    (b) *If not mentioned:* Have you engaged external domain experts in your project to assess the ethical considerations of your project (e.g., academics, ethicists, researchers)?
    (c) *If not done currently:* How do you think ethical risks can be mitigated in your software?
5. Did you have any training or education on ethics in creating digital tools?

B Themes

- **Theme 1: Current practices and attitudes for identifying unintended consequences** *privacy-first focus, create usable products, personal interest and moral compass, positive software can have no negative impacts, reactivity over proactivity, lack of formal processes and no frameworks/ responsible people*
- **Theme 2: Reasons unintended consequences considered and not considered** *unintended consequences ingrained in software vision, unintended consequences accounted through design, Inevitable Unintended consequences, unintended consequences that are not applicable, Software creators are unaware of unintended consequences*
- **Theme 3: Factors that hinder working on unintended consequences** *lack of accountability, good intentions have no negative consequences, lack of awareness individual level, lack of awareness organizational level, privacy vs collecting user feedback/ content moderation*
- **Theme 4: Systematic frameworks help with identifying unintended consequences**

C Codebook

Category	Code groups	Sub-codes
Primary users	Regular Users	Everyday users, regular unt-argetted users, un-specific
	Privacy Concerned users	Users avoiding ads, avoiding surveillance
	Special and Vulnerable populations	Journalists, Activists, IPV survivors
	Businesses	Mid-small sized businesses, corporates, employers
Un-intentional(non-users)	One time users	Out of necessity or crisis, downloaded by accident/ unaware of functions
	Unintentional(non-users)	People communication with users, users of product developed, employees
Anticipated negative impact	No direct harm	No consequences, only positive impacts
	Usability/ bugs	Bugs affecting usage
	Harmful actors	Spammers, abusers, spread misinformation
	Incriminating	downloading software can be incriminating, illegal in some countries
	Inevitable harms	Use by harmful actors, No moderation
Barriers to Access	Does not apply	-
	Applies - not mitigating	Skills needed (technology/ language), internet needed
	Applies - mitigating	Marketing to more users, reducing barriers (play store, language)
	Positive	Everyone can access
Biases	Does not apply	-
	Applies - not mitigating	Global north/ west, language
	Applies - mitigating	Global south, adding language
Accessibility	Does not apply	Just a library, no interface
	Applies - not mitigating	Not prioritized/ no process, no resources
	Applies - mitigating	Works with accessibility tools, designed with accessibility
Addiction	Does not apply	Possible in future with sufficient gamification
	Applies - not mitigating	Not considered, no incentive for maximizing attention
	Applies - mitigating	Working on mitigation: Offline capability
Distraction	Does not apply	Not distracting
	Applies - not mitigating	Difficult to integrate into workflow, users are responsible
	Applies - mitigating	Limit notifications through UI design, no ads
Disempowerment	Does not apply	No/No direct impact, empowering
	Applies - not mitigating	Disempowering when corporates force the use of this tool
	Applies - mitigating	Consent for data sharing
Environmental impacts	Does not apply	Not different from other software
	Applies - not mitigating	Software used by bitcoin, end device consumes energy
	Applies - mitigating	Use renewable engery
Human rights	Does not apply	no violation, strengthens human rights
	Applies - not mitigating	Social media content may be problematic, may have issues in future
	Applies - mitigating	-
Physical & mental health impacts	Does not apply	No/ No impact, positive impact
	Applies - not mitigating	Warnings from stalkerware can make people anxious
	Applies - mitigating	-
Misinformation	Does not apply	Not a communication/ content software
	Applies - not mitigating	Mitigation is with users, hard to mitigate, cannot moderate
	Applies - mitigating	Marking as spam/blocking strategies present
Security	Process	Security review, threat modeling, research
Data protection, surveillance & governance)	Processes	Minimize data collection, no sharing with external partners, consent processes
User involvement	Proactive practices	User research at the start, constant feedback, co-developing
	Reactive practices	Raise issues, fix issues/ react on feedback, impacts through feedback
	Challenges	No user data(privacy), small samples, lack of resources
Impact assessment	Current practices	Privacy & not ethical impacts, no processes, personal interest, Diversity
	Accountability	Leadership/bosses should be responsible, everyone in the team, no assigned person
	Framework/ Process followed	No framework used, user-centered design, Moral compass
	Training/ education ethics	No education, education in ethics (university/ company training), personal interest
	Own process	Proactive, reactive(git issues/reviews), experts(Legal, ethics, compliance)
Oganizational factors	Size of organization	Micro, Small, medium, large
	External Collaborations	lack of external collaborations (researchers/ ethicists)
	Diversity	No diversity (all white/ male), gender diversity, diversities of ethnicities