# Duty to Respond: The Challenges Social Service Providers Face When Charged With Keeping Youth Safe Online

XAVIER V. CADDLE, University of Central Florida, U.S.A
NURUN NAHER, University of Central Florida, U.S.A
ZACHARY P. MILLER, University of Central Florida, U.S.A
KARLA BADILLO-URQUIOLA, University of Notre Dame, U.S.A
PAMELA J. WISNIEWSKI, Vanderbilt University, U.S.A

Social service providers play a vital role in the developmental outcomes of underprivileged youth as they transition into adulthood. Educators, mental health professionals, juvenile justice officers, and child welfare caseworkers often have first-hand knowledge of the trials uniquely faced by these vulnerable youth and are charged with mitigating harmful risks, such as mental health challenges, child abuse, drug use, and sex trafficking. Yet, less is known about whether or how social service providers assess and mitigate the *online* risk experiences of youth under their care. Therefore, as part of the National Science Foundation (NSF) I-Corps program, we conducted interviews with 37 social service providers (SSPs) who work with underprivileged youth to determine what (if any) online risks are most concerning to them given their role in youth protection, how they assess or become aware of these online risk experiences, and whether they see value in the possibility of using artificial intelligence (AI) as a potential solution for online risk detection. Overall, online sexual risks (e.g., sexual grooming and abuse) and cyberbullying were the most salient concern across all social service domains, especially when these experiences crossed the boundary between the digital and the physical worlds. Yet, SSPs had to rely heavily on youth self-reports to know whether and when online risks occurred, which required building a trusting relationship with youth; otherwise, SSPs became aware only after a formal investigation had been launched. Therefore, most SSPs found value in the potential for using AI as an early detection system and to monitor youth, but they were concerned that such a solution would not be feasible due to a lack of resources to adequately respond to online incidences, access to the necessary digital trace data (e.g., social media), context, and concerns about violating the trust relationships they built with youth. Thus, such automated risk detection systems should be designed and deployed with caution, as their implementation could cause youth to mistrust adults, thereby limiting the receipt of necessary guidance and support. We add to the bodies of research on adolescent online safety and the benefits and challenges of leveraging algorithmic systems in the public sector.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: Online safety, Underprivileged Youth, Human-Centered Artificial Intelligence

6

## 1  INTRODUCTION

Social service providers (SSPs) are one of the most important and necessary stakeholders in
providing care services for underprivileged[1] youth [78]. Social work subsumes multiple professions
(e.g., juvenile justice officers, educators, mental health professionals, and child welfare caseworkers),
which promote the ideals of social justice, human rights, empowerment, and actively creating
structures to better the lives of people [66, 69]. Disadvantaged youth that lacks strong support
systems at home rely on SSPs to provide wrap-around services and supplemental support to
thrive [68]. These youth face numerous life challenges, such as drug addiction, violence, self-harm,
substance abuse, serious crimes, and mental health issues, which makes them more vulnerable than
privileged youth [59, 78]. Research has consistently confirmed that SSPs are critical to increasing
positive outcomes for underprivileged youth in terms of offline risks [33, 45, 57]. Meanwhile,
research on how SSPs serve to prevent and mitigate *online* risks of underprivileged youth is fairly
nascent [6, 26] and warrants further inquiry.

Researchers who have studied the online risk experiences of underprivileged youth often high-
light the severity of these risks and the lack of support and resources parents have for risk mitigation.
For instance, El Asam and Katz [26] found that youth who were more vulnerable offline (e.g., youth
experiencing family instability, who had disabilities or special education needs, or those with
significant mental health challenges) also experienced more high-risk experiences online. Badillo-
Urquiola et al. [7] surfaced the concerns of foster parents in mediating technology use in the
home to protect foster children from further abusive situations, including sex trafficking that was
facilitated through online communications. They found that foster parents felt unsupported by
SSPs, lacked the technological skills to manage technology in their homes, and thus were at a loss
for how to mitigate online risks. This often resorted to technology restriction as a means of risk
prevention. As such, the researchers called for more sociotechnical solutions to help foster parents
mediate the technology use, as well as more support mechanisms to help foster parents protect the
underprivileged youth in their care.

Due to their training and close daily interactions with many underprivileged youths, SSPs may
be more equipped, have a more nuanced view, and have better insights on how to prevent and
mitigate online risks than many parents. On the other hand, numerous studies [43, 55] have also
highlighted the significant lack of resources within the public sector to be able to consistently
provide the needed social services underprivileged youth require to achieve positive outcomes.
Abaquita [1] interviewed child welfare caseworkers and found that they faced their own challenges
trying to manage the online safety of foster youth. For example, case managers do not receive the
appropriate training to properly help youth mediate online risks, and they are overburdened with
large caseloads and other tasks to the extent that online safety is not prioritized. As such, researchers
in the SIGCHI community [29, 80] have begun to study whether leveraging algorithmic decision
support systems could potentially bolster social services provided within the public sector. Most
relevant to the current work, Saxena et al. [79, 80] found that the child welfare system frequently
deploys algorithmic decision-support systems to assess offline risk outcomes (e.g., mental health

---

[1]Literature has used deficit-based labels such as "at-risk" [27, 68, 75] to describe youth who are *"considered to be "at-risk" of
failing to make the transition to adulthood successfully"* [84]. Such labels tend to negatively describe the youth as opposed to
focusing on the challenges faced by the population [59]. Therefore, we chose instead to describe youth as "underprivileged"
rather than "at-risk" to emphasize their lack of ecological systems of support as opposed to their proclivity towards risks.

challenges, sex trafficking risks) of foster youth, so that caseworkers can proactively mitigate risk outcomes and support the needs of the children placed in the child welfare system. They found that while some of these algorithmic decision support systems reduced caseworker burden and improved youth outcomes, others inhibited caseworkers' ability to leverage their expertise in doing what was in the best interest of the child. Therefore, they concluded that more human-centered approaches are needed to clearly identify the benefits and drawbacks of how artificial intelligence (AI) systems should (or should not) be deployed in high-stakes decisions affecting the lives of youth.

Algorithmic risk detection has been leveraged across multiple public domains, including law enforcement [63], mental health [41, 48], education [42], and child welfare [32]. Yet, the deployment of such systems, if not done carefully, could lead to unintended consequences and harm, particularly to vulnerable populations, such as systems that have unintentionally committed privacy violations [58, 85] and propagated racial discrimination [37]. As such, this work is the first to interrogate the idea of leveraging AI risk detection as a means to help SSPs identify and mitigate the online risk experiences of underprivileged youth. First, however, we set out to identify the challenges and currents ways in which SSPs tackle the problem of online risks, which leads us to the framing of our high-level research questions:

- **RQ1**: *What online risks do social service providers of underprivileged youth find most concerning in terms of the positive youth outcomes they are trying to achieve?*
- **RQ2**: *What are the current ways in which youth social service providers become aware of such online risks?*
- **RQ3**: *According to the social service providers, what would be the advantages and disadvantages of an automated risk detection system that notified providers of online risks encountered by the youth they serve?*

To answer these research questions, we conducted 37 semi-structured interviews with SSPs of underprivileged youth. These interviews were conducted as part of the National Science Foundation (NSF) I-Corps program [56], which is designed to train researchers on how to conduct customer discovery to bring cutting-edge technologies to market. Therefore, while the initial purpose of the interviews was to find a viable business model for deploying human-centered AI risk detection for youth online safety (RQ3), the primary focus of this research (RQ1 & RQ2) was first on understanding SSPs as stakeholders in the space of online safety for underprivileged youth, then to identify the benefits and disadvantages of leveraging AI risk detection in this context.

Overall, we found that across all SSP groups, sexual risks and cyberbullying were the most prevalent online risks to which underprivileged youth were exposed. SSPs found out about these incidents primarily through youth self-reports via direct questioning in conversations with youth but also through formal risk-assessment tools, such as surveys and questionnaires. In the absence of good relationships with SSPs, youth often did not report these incidents, which when missed, exposed the youth to potentially harmful offline dangers, such as human trafficking, molestation, self-harm, suicide, and/or substance abuse. SSPs thought the prospect of using technology as a means to detect online risks in a more timely fashion would be helpful since it would allow them to enact interventions sooner. However, they had concerns regarding the ability of such solutions to gain access to the data necessary to make well-informed decisions about youth circumstances, the liability they may face when using such solutions, and the resource constraints they currently have that might limit their ability to respond to incidents once recognized. The contributions of our study are thus as follows:

- *An in-depth understanding of how SSPs currently identify online risk experiences that are concerning to the safety of underprivileged youth under their care.*

- *A comparison across different types of SSPs as to how risk awareness and mitigation strategies varied based on their unique job roles.*
- *A better understanding of the role AI risk detection may play in helping SSPs mitigate online risks, as well as potential pitfalls. We provide implications for the future design of automated risk detection systems based on the unique perspective of SSPs who work directly with underprivileged youth.*

The key contribution of this work is in highlighting the sociotechnical problems related to SSP awareness and digital literacy in mitigating the online risks faced by underprivileged youth. Further, our work is the first to take a human-centered approach to understanding whether SSPs see potential in the use of artificial intelligence (AI) and algorithmic decision support for automated detection and/or mitigation of online risks. Taking into account the perspectives of SSPs is important as they are often the first line of defense in protecting the safety and well-being of underprivileged youth – both online and offline.

## 2 RELATED WORK

The sections below situate our research by positioning SSPs as integral players within underprivileged youths' socio-ecological systems of support and who are directly and/or tangentially charged with ensuring their online safety.

### 2.1 The Role of Social Service Providers in Protecting Underprivileged Youth

SSPs are either directly or indirectly connected with underprivileged youth, including youth from low-income households, homeless youth, foster care youth, and teen moms [5]. As such, SSPs are considered to be part of the third-level support (i.e., Exosystem) of Bronfenbrenner's five ecological system framework, which outlines social systems that have an impact on a child's development [18]. Previous literature highlights how underprivileged youth require support from many of the services offered by different social service providers [57]. SSPs may include, but are not limited to, law enforcement officers, mental health professionals, and child welfare caseworkers [14, 47, 50, 55]. The research presented acknowledges that underprivileged youth have challenges that can be addressed by SSPs. This research has shown patterns between social service needs and service use, in that unmet service needs and/or prior service involvement bear some influence on youth's entry and prolonged involvement with the juvenile justice system [57]. This is backed up by findings that show that relationships between child welfare caseworkers or mental health service providers and underprivileged youth are associated with more positive youth outcomes [45]. A recent study surveyed a large group of underprivileged adolescents aged 12–17 years and discovered that youth that engaged in two or more service systems (child welfare, juvenile justice, additional education, mental health) showed an increase in youth resilience and better well-being outcomes [78]. In this regard, youth resilience is not an outcome of the individual alone but involves the help of multiple positive service settings, resilience helps youth to cope with risk themselves, hence they can achieve better outcomes [34]. It is therefore essential to identify the SSPs concerns as it relates to underprivileged youth and the risks they encounter both offline and online. Our research adds to this body of knowledge by focusing on the online risk experiences of underprivileged youth through the lens of SSPs.

### 2.2 Underprivileged Youth and Online Safety

Most youths are heavy consumers of technology and social media. While social media platforms provide youth with benefits such as social connections, learning, and creativity [4, 76], they also expose youth to online risks, such as cyberbullying [12, 60, 83], sexual risks, and exposure to

inappropriate material [8, 36, 73]. An increasing amount of literature has expanded on the negative consequences of social media on youth, such as mental health issues, self-harm, and suicidal tendencies [13, 28, 52]. To date, most of this literature has focused primarily on online safety for general adolescent populations [35, 38, 71, 94] with a focus on parental mediation of online risks [54, 82] and the first-hand risk experiences of youth themselves [21, 95]. Thus, the focus has primarily been at the microsystem level of Bronfenbrenner's ecological framework. Yet, many underprivileged youths lack strong microsystems of social support and must rely instead on SSPs to help them with these unmet needs [45, 57]. While prior work has confirmed that underprivileged youth are more vulnerable to online risk and harm [26, 90] than their counterparts, less work has examined the role of SSPs in identifying and mitigating these risks. We aim to fill this gap by identifying what SSPs consider the most concerning online risks for underprivileged teens (RQ1) and how they identify those risks (RQ2). In the next section, we move towards the potential for using AI for the automated detection of online risks.

Past research provides invaluable insights and a broad overview of the current state of the online safety literature regarding underprivileged teens. For instance, previous work has shown that foster youth are at higher risk for offline risks, such as mental health, self-harm, and suicidal ideation, and a majority of these offline risks are directly or indirectly influenced by the online encounter of risks [6]. Good et. al [34] interviewed 25 service providers from youth in residential treatment to understand digital media use effects on highly vulnerable populations. They found that youths in residential treatment have varying degrees of online behavior issues related to impulse control, judgment, boundaries, and self-esteem, which can increase both online and offline risk [34]. This is backed up by findings that show that underprivileged teens are likely to experience a higher level of online risk with greater impact due to their vulnerable situations and may lack parental guidance on staying safe in online or offline life thereby feeling isolated and motivated to find compensatory friends/relationships online [26, 53]. SSPs can instead fill that void and are motivated towards improving youth outcomes as studies that investigated online risks faced by underprivileged youth in foster care have shown [10].

## 2.3 Towards Automated Risk Assessment to Protect Underprivileged Youth

Since SSPs are often under-resourced, governments and non-organizations have increasingly turned to automated tools for algorithmic decision support [75]. These tools include different forms of risk assessments, such as those used in the domain of child welfare to help in child placement, referral processes, reducing workload [24, 32], and increasing efficiency [87]. Efficiency in this context is gained through automation and embedding machine learning algorithms in Decision Support Systems (DSS). Being able to move resources to areas of high need for underprivileged youth necessitates identifying or predicting the risk factors that they face which is especially important when the resources given to youth social service providers are constrained [75]. However, there are debates surrounding the use of algorithmic decisions in place of the caseworkers with the argument being made that caseworkers would be more contextual since they would know more about the circumstances of the specific cases under scrutiny [44, 79]. In recent years there has been a push towards having more transparent algorithmic processes and explainable AI even in the Child Welfare context as exemplified in [80]. It has also been suggested that major stakeholders in the Child Welfare DSSs such as investigators, researchers, designers, and public service agencies, engage in participatory design to produce more socially acceptable systems [81]. To do this in the context of automated detection of online risks for underprivileged youth, we engaged directly with SSPs regarding their role in youth protection as it relates to online risks. It is imperative to examine the views of SSPs in this context. Of particular interest are the online risks they view as

most salient in the execution of their roles, as well as how they identify and mitigate these online risks to promote positive outcomes for the underprivileged youth they serve.

## 3  METHODS

In this section, we provide an overview of our study, detailing our process for data collection, data analysis, and participant recruitment.

### 3.1  Study Overview

We conducted 37 semi-structured interviews with SSPs over a two-year period from July 2019 to August 2021 in the United States. We undertook this work as part of a larger project sponsored by the National Science Foundation (NSF), which gave us the opportunity to partake in the NSF Innovation Corps (I-Corps$^{TM}$) program [56]. The NSF I-Corp program provides researchers with the training and tools necessary to become entrepreneurs by supporting customer discovery toward commercialization. One of the requirements of the process was to interview individuals who are subject matter experts in the researchers' area of interest (i.e., adolescent online safety and risks) to determine what the current needs are in the domain and how those needs could be best addressed. As part of the NSF grant itself, the over-arching goal was to produce a viable commercial product based on the core research aims. Further, we were tasked with identifying a customer base for such products. As such, SSPs were only one subset of potential users (i.e., customers) that were interviewed as part of the I-Corps program. We conducted interviews with other stakeholders, such as social media companies, parental control companies, online safety educators, and researchers. However, those interviews were held outside of the scope of the current paper as they had different needs and faced challenges that SSPs who served underprivileged youth. Also, this is why a subset of the interview questions focused on understanding the role technology, particularly algorithms trained to detect online risk behaviors and experiences of adolescents, may play in the job roles of key stakeholders in the space of adolescent online safety, in this case, SSPs.

### 3.2  Semi-Structured Interview Design

We used a semi-structured interview method to elicit context-rich responses from SSP participants. Interview questions were chosen to understand the participants' interaction with underprivileged youth, what metrics they used to track youth outcomes, and the role technology played in these youth outcomes. There were four main parts to the semi-structured interviews. First, we asked general background questions to understand the occupational role of SSPs and the youth populations in which they engaged. We then asked questions related to the online risks they felt were most relevant and concerning for their role in youth protection. We also included questions on the use of technology to identify online risk experiences among underprivileged youth. This section was introduced with a high-level description of a technology capable of showing the participant the online risks experienced by particular youth was described. Questions included examples of online risk experiences of youth they worked with, how they became aware of such risks, and the current approaches for risk mitigation. Next, we introduced the idea of using automated tools for online risk detection. The description given to participants was kept simple to limit any bias related to how the researchers thought the technology should operate and also to limit jargon. We described the scenario of a web-based dashboard, where SSPs could review risk-flagged data from youths' social media accounts. No technology jargon or technical operation specifics were used by the interviewer unless asked for clarification. For example, participants often asked how the system would gain access to such information. The researcher would offer possible solutions, such as the youth or their legal guardians providing the necessary account information for authentication.

We asked SSPs to consider the potential benefits, drawbacks, and design considerations for such a system. The semi-structured interview questions are included in Appendix A.

## 3.3 Participant Recruitment and Procedure

We recruited SSPs living in the United States that were 18 years of age or older and had previously or were currently working with youth between 13 to 17 years old. We used non-probabilistic purposive sampling to recruit participants [51, 86]. Purposive sampling is a widely used and accepted technique employed for participant inclusion in which researchers select participants due to the knowledge they have on the topic of interest [86]. To reach our target population, we used three different approaches: 1) advertisements were placed on social media via the researcher's social media accounts on Facebook, Instagram, LinkedIn, and Twitter, 2) cold calls using phone numbers from online searches for SSPs such as *"Youth Social Service Providers in [City]"*, and 3) through word-of-mouth propagated through the researchers' online and offline social networks.

Before starting each interview, participants were provided an IRB approved explanation of the research and given the option to withdraw from the study (but continue with the interview for I-Corps purposes) if desired. Out of the 37 interviews, 13 were conducted in person. The remaining 24 interviews were conducted online using Zoom video conferencing to accommodate for COVID restrictions. Each interview was approximately 30 minutes to an hour and a half. Two interviewers were present in most of the interviews with one designated as the interview lead with the other taking notes. At the beginning of the interview session, all participants were asked for consent to audio and/or video record the interview session. Due to the constraints of the I-Corps program, participants were thanked for their time but not given a monetary incentive.

We had a wide variety of SSPs, including Juvenile Justice and Law Enforcement Officers, Educators (including Teachers, Researchers, and participants whose primary focus is training), Mental Health Professionals, and Child Welfare Workers. We used the following definitions to categorize the types of underprivileged youth's social service providers:

- Judicial System (N=11): *Law enforcement, or works with juvenile offenders to decrease the rate of recidivism.*
- Education (N=10): *Works with youth within an educational setting or trains people to encourage risk prevention and academic researcher who works with teen and online risk.*
- Mental Health (N=9): *Provide psychiatric treatment and prevention, either through behavioral therapy or medication, directly to teenagers.*
- Child Welfare (N=7): *Child welfare and nonprofit organizations provide policy guidelines for community building, outreach, problem analysis, and rectification.*

SSPs also held different roles within their organizations from ground-level social service support to director-level/managerial roles. Participant characteristics are included in Appendix B.

In drawing from such a diverse group of SSPs who work with youth, our goal was to obtain a more holistic view of issues affecting underprivileged youth at varying stages in their development and interaction with the social structures which provide them with support and direction [18].

## 3.4 Data Analysis Approach

In-person interviews were audio-recorded, while Zoom interviews were audio and video recorded and transcribed using live transcription in Zoom. Two research assistants watched each video and/or listened to the audio recordings to correct the transcriptions for errors. For qualitative analyses, we followed the thematic analysis process as described by Braun et al. [17]. Based on their familiarity with the data from conducting the interviews, the first and last authors worked together to identify the key dimensions for which to code the data, an initial codebook of emergent

codes, and formulated the formal research questions to frame the analysis. To answer our research questions, the interviews were first coded for the types of online risks youth encounter and are perceived to be the most concerning by the SSPs (RQ1). Then, they were coded for how these online risks were discovered by the SSPs and the limitations that exist with these approaches (RQ2). Finally, the transcripts were coded for the benefits and limitations of an automatic online risk discovery technology (RQ3). The interview transcripts were evenly divided and coded interdependently (i.e., through constant consultation with one another via Skype) by the first three authors based on the initial codebook, but we allowed for new codes to emerge during this process. All new codes were agreed upon and discussed among all authors. After initial coding was completed for each RQ, the coders presented their initial results to the last two authors for feedback, discussion, and iteration. At this time, codes were clarified, merged, changed as needed, and organized into themes. Conflicts between coders were resolved through consensus building. Tables 1, 2, and 3 present the final themes aligned with our three over-arching research questions. In our results which follow, illustrative quotes from participants are used to describe the findings and each quote is identified by the participant's unique ID (P=Participant) followed by SSP type (e.g., PX, Education).

## 4 RESULTS

In the following sections, we present our findings. First, we present the themes found across all SSPs related to the online risks youth experience (RQ1), and how SSPs found out about these situations (RQ2). Finally, we present the benefits and limitations of using automated risk detection technologies for these situations (RQ3).

### 4.1 Most Salient Online Risks (RQ1)

Our results indicate there are three major types of risks among youth that all the SSPs found to be concerning: sexual predation/pornography-related risks, cyberbullying, and internet addiction. Yet, the manner in which the different SSPs discussed these risks was markedly different. These differences mostly resulted from the SSPs focusing on different success outcomes and goals related to their job duties. In this section, we discuss the risks SSPs found most concerning in relation to the positive youth outcomes they try to achieve.

*4.1.1 Underprivileged Youth are Frequently Exposed to Sexual Predation and Subjects of Child Pornography.* Overall, 68% (N=25/37) of SSPs stated that sexual predation and pornography-related risks are the most salient among their youth. Many of the stories reported by SSPs were about youth being coerced by online adult predators to engage in sexual activities. However, SSPs also raised concerns about youth engaging in pornography by sharing intimate photographs or videos of themselves with people they think are trustworthy who in turn share that media with others. Among each of the SSPs groups, 71% of child welfare workers, 70% of educators, 67% of mental health professionals, and 64% of juvenile justice officers reported on sexual-related risks. Yet, each of these providers focused on different concerns.

Child welfare workers mostly focused on human trafficking concerns. They said the youth they interact with tend to already have histories of involvement in sex trafficking or histories of trauma that make them more susceptible to these sexual risks.

"*I think our kids are already really vulnerable. So if someone friends them on Facebook, or Instagram or whatever, and is promising them these things, they can easily coerce them into getting or doing something.*" - P24, Child Welfare - Youth Engagement Services

Table 1. Emerging themes for RQ1

| RQ | Theme | Definition | Quote |
|---|---|---|---|
| **RQ1: Most Salient Online Risks** | Youth exposed to sexual predation and pornography | Taking advantage of children by sexual coercion, grooming, or any other sexual act, as well as online exposure to illicit media (pornography) and/or sharing ("sexting") of indecent messages and films to other children or strangers. Sexual solicitation can arise when people post too much personal data on social media sites. | *"Part of the biggest concern for parents that we have is contacting predators it's probably number one number two is exposure to things that you know that they're not really ready to see..."* |
| | Cyberbullying impacts youth mental health | Through targeting vulnerable youth or making threats online, has a significant impact on psychological well-being. | *"...we actually had a youth that was being cyberbullied, and that impacted their mental health, quite negatively and that child committed suicide..."* |
| | Internet addiction inhibits youths ability to develop relationships | Internet addiction (mobile, online gaming, social media) and/or participation in online trends encourages youths to engage in social media disputes, physical fights, gang allusions, and other forms of violence. On social media, kids download and utilize copyright content, i.e. piracy, and promote drug use, photos with weapons, money, and gang recruiting. | *"I would say addiction. Because it doesn't matter if it's online gaming or cyberbullying, or on Pinterest or whatever the media platform is, they have disengaged completely from reality and each other and that's a problem because communication is almost breaking down"* |

Educators and mental health providers tended to share stories about youth engaging with other youth in sexting or pornography. A lot of these stories tended to be about romantic interactions, that is, interactions between boyfriends and girlfriends. However, there were also accounts of these experiences occurring between friends and peers as well.

*"I am getting more and more calls with parents who say my seven-year-old was playing with her friend who was seven, they had their iPads, and I walked in, and they were watching porn on accident,*

Table 2. Emerging themes for RQ2

| RQ | Theme | Definition | Quote |
|---|---|---|---|
| **RQ2: Risk Discovery Approach** | Relationship with youth most important | Having direct contact with youth, such as through direct conversation or purposefully building a connection, discloses online risk. | *"...That's the end goal, like building those relationships but it's never an explicit part of what we like around our programming....most of the time when you find out the teens are experiencing some kind risks online is because it's through the relationships you've built."* |
| | Formal investigations uncovering risk incidents | Risk is discovered when law enforcement conducts investigations. | *"The officer will then turn the phone into the digital forensics lab and then turn in some type of consent form or a search warrant for the phone and then we will use different tools to conduct what's called an instruction on the phone so it's when we're actually pulling that data out of the phone."* |
| | Risk assessment tools uncovering risk incidents | Online risk was discovered through a tool (e.g. Questionnaire) that asses risks, taken either by both victims or guardians or just the victims. | *"Through the assessments. With the open cases, we do a clinical assessment, and we have sections on assessing for risk"* |

*or their friend told them or someone else told them."* - P4, Mental Health - Therapist

Educators felt that these risks are a result of youth having easy access to technologies in which pornography can be shared. For example, educators stated that students use school computers and tablets to do classwork and homework. These devices typically have programs like Dropbox where students can store, share, and access files online. Educators said students take advantage of these platforms as ways to share pornography:

"*Students were using Dropbox as a way to basically share pornographic material of students so girls would take pictures they would send to boyfriends, boyfriends were uploading into the Dropbox and then sharing it around.*" - P37, Education - School Administrator

Even though juvenile justice officers also talked about youth sharing intimate sexual related content online, they mostly focused on the risks of these behaviors becoming criminal offenses such as child exploitation, pornography, and the online solicitation of minors for sexual acts. They emphasized sexual predation and pornography experiences between youth and adults.

Table 3. Emerging themes for RQ3

| RQ | Theme | Definition | Quote |
|---|---|---|---|
| **RQ3a: Technology Benefits** | Monitor youth online experience | Technology can automatically provide metrics to SSP, Parent, or Guardian. | *"...you have it connected to your phone or your computer, so that when certain things come up so you can kind of put a range of things that you want your attention brought to you. So there are a lot of curse words or sexual words or search histories and things like that it can send alert."* |
| | Early discovery of potential youth issues | Technology can be used to find out about potential challenges facing youth sooner | *"...we can't control what they do outside of it as much as we can inside the school, but if we're at least aware of it, we can know that that's something we need to work towards and be aware of."* |
| | Better resource management | Technology can be used to help SSP filter and prioritize cases | *"..it would be helpful if I was getting the cyber tip like... You know username unicorn123 is whatever percentage more susceptible you know to be bullied online or sexually exploited, then, that would be just more helpful for me to try to determine which cases are more of a priority."* |
| **RQ3b: Limitations of Technology Use** | Lack of nuanced contextuality | Inability of technology to make contextual determinations about youth circumstance as would a human | *"I think in that respect, every risk tool, or every tool that's needed, needs to be looked at with worth taking into consideration that each person is an individual, and it's not going to be the exact same that you can't use that as the end all be all."* |
| | Inability to access the necessary data | Solution providers don't get access to the data they need to provide quality products | *"I think the first barrier that comes to mind is privacy."* |
| | Liability and increased complexity of duties | SSPs already have too much to deal with and don't have room for assessing online risk | *"I think one of the potential liabilities is let's say that we were using this tool,... It could potentially put us in a really bad spot because if we weren't monitoring it close enough, ... Potentially, we could be liable because when we are part of our charges, if we're going to assess for something, or if we're going to screen or we're going to monitor something, then we have a duty to respond to it.* |

"*One of the things that we tend to see the most when it comes to juveniles is sending images of themselves... We do a lot of time looking for people that they've sent these pictures to or communication that they're having with usually older people.*" - P33, Juvenile Justice - Law Enforcement Officer

Child welfare workers and juvenile justice officers are similar in the concept of sexual risk as they mostly dealt with online sex predation or criminal offenses where youths were the victims of sex trafficking and child pornography. It is related to the success outcome they want to achieve which are reducing recidivism and reducing truancy. When they discover abstinence of any youth, they will investigate to find out if that youth is missing or became the victim of sex trafficking. Both educators and mental health providers directly interact with youth in their daily job. Educators focus on teaching youth knowledge and skills pertaining to their personal development. Likewise, mental health providers support youth in achieving mental health stability.

Participants in all SSP groups identified youth online sexual exposure as being associated with other offline and online risks such as increased sexual activity, increasingly risky behavior (e.g. meeting strangers in person for sex), becoming reclusive, the possibility of human trafficking, and prostitution.

"*...kids that are starting to engage in high risk sexual activity online, to kids who are engaged in really high risk behaviors talking to videos, so you know sexually explicit videos, photographs, attempting to meet our meeting with adults online for sex, and actually what it is to be sexually abused by adults, to trafficking.*" - P34, Education Advocacy Services

Concern was also expressed by participants about how the online sexual practices of youth led to cases of cyberbullying, which at times had extremely negative outcomes. This online risk is discussed in the next section.

*4.1.2  Cyberbullying Impacts Youth Mental Health.* Another prevalent risk among the youth that 42% of SSPs (N=15/37) reported was cyberbullying. This was shared among 56% of mental health providers, 45% of juvenile justice officers, 30% of educators, and 29% of child welfare workers. Participants tended to focus more on the effects of cyberbullying than they did on the reasons why it happened. Those who addressed the reasons elaborated that it stemmed from negative influences in the life of youth, from the perspective of both the bully and victim. In the case of the bully, their involvement in peer pressure groups such as gangs and online groups which encourage or egg on negative behavior came to the fore. Peer pressure, reward, and/or clout, were put forth as reasons bullies direct their attacks on victims. From the perspective of youth victims, the inability to escape from online attackers and females being involved with abusive partners or exes was presented as the reasons behind cyberbullying.

"*Drug use. Gang involvement. Negative peer influences. Group fights. Group instigations. Cyberbullying.*" - P11, Juvenile Justice - Probation Officer

Generally, participants shared stories about peer pressure leading to cyberbullying. These online behaviors and risks often led to offline consequences such as mental health issues, injuries through self-harm, and suicide.

"*...we actually had a youth that was being cyberbullied, and that, unfortunately, impacted their mental health, quite negatively... COVID hits, there was more of an in home online environments, and*

*cyberbullying took place, and that child committed suicide..."* - P15, Child Welfare - Child and Family Services

When mental health professionals spoke of cyberbullying, they spoke more about its effects on youth based on interactions between the youth, their families, and the professionals. The effects of cyberbullying they mentioned were signs of depression and anxiety in youth who are targeted. They also mentioned threats being made through digital mediums, and fights being recorded for social media upload.

For educators, they discussed the inability of youth to escape cyberbullying due to the persistent nature of online media and ubiquitous online access. Peer pressure through the consumption of social media which sets unrealistic lifestyle demands in youth leading to negative views of self was also mentioned. The interplay between peer pressure and cyberbullying was discussed when SSPs shared instances where youth started fights online, met in person to record themselves fighting offline, and then posted it on social media sites. Concern about the impact cyberbullying has on youth well-being and increased suicide rates was the underlying theme.

*"... I know that kids at certain schools will have like these school themed pages where they basically put up like, really harsh photos and information about kids at that school and it becomes like this, like focal point for a lot of these kids on social media..."* - P19, Education - Language Arts Teacher

Juvenile justice officers often talked about cyberbullying risks with offline risks or consequences. For instance, suicide was mentioned in conjunction with cyberbullying, and law enforcement officers stated that these were highly requested to speak on these topics by parents and concern groups. They also often linked consistent online peer pressure with the concept of cyberbullying. The concept of *cancel culture* being used in peer pressure to an extent that it is linked to cyberbullying was also a concept raised by law enforcement officers.

*"Cyberbullying is something that we see quite a bit. And it's cyberbullying, but it's also this. I want to say it, like, you know, one kid who's kind of egging people on, so to speak."* - P2, Juvenile Justice - Prevention and Intervention

In speaking about cyberbullying, child welfare officers discussed it effects on youth in the form of truancy. They stated that youth tried to avoid online bullies in real life even if that meant not going to school, which the child welfare officer would eventually find out. This SSP group also raised concern about the toxic banter used in gaming culture, which in some cases is seen as closer to harassment to those on the receiving end of the comments.

*"I'm engaging in online activity, like late into the night, and I can't get up early, or, like something like that, like, there's a cyberbully, you know, I'm afraid of someone because of online interaction I've had."* - P3, Child Welfare - Children, Youth and Family

Each SSP group therefore identified cyberbullying as a major problem and highlighted the negative and often extremely harmful offline consequences that result from its prevalence.

*4.1.3 Internet Addiction Inhibits Youths Ability to Develop Relationships.* The other prominent risk among 24% (N=9/37) of the SSPs was internet addiction. It was identified as a risk by 33% of mental health professionals, 29% of child welfare workers, 27% of juvenile justice officers, and

10% of educators. This risk often manifested itself as staying up late (past bedtime) using online applications. Internet addition not only affected school attendance, according to the SSPs, but it also deteriorates the communication and interactions youth have with one another. Child welfare officers and educators both spoke of youth becoming anti-social or otherwise detached from communicating in real life, instead opting for digital forms of communication with peers. This was said to be associated with depression in youth and truancy.

"*Yeah, I would say addiction. ...they have disengaged completely from reality and each other and that's a problem because communication is almost breaking down*" - P13, Child Welfare

Taking a closer look at the specific roles that mentioned this type of youth behavior, juvenile justice officers explained that the youth exhibited non-compliance with program requirements or court orders regarding restricted mobile app usage. The concern raised was that youth would often use a device owned by someone else or choose harsher punishments such as going to jail rather than complying.

"*...if I say give me your phone, I'm confiscating that, no matter if they're going to jail for two years, in that point in time they will take the two years in jail, they're not giving up their phone.*" - P11, Juvenile Justice - Probation Officer

Mental health providers shared concerns about the internet being the only place youth have learned to connect with people, and where that could possibly lead with respect to their mental health. In this way, youth are seen as developing a dependency on social constructs form through internet usage thereby having problems interacting in with people in the physical world. In speaking about addictive internet usage, mental health SSPs also shared the issue of youth developing behavioral issues due to conflicts with guardians over spending too much time online (such as with gaming or social media use). SSPs here also mentioned that youth also purposely put themselves in danger for online "*likes*" or "*clout*". One example shared was that some youth try to get "*Baker Acted*" [62] by being purposely institutionalized and streaming their experience online for views.

"*They're doing it [purposefully], because they want to get like seen, and some clout and things like that.*" - P20, Mental Health - Therapist

In this way, while SSPs identified several different risks affecting youth, sexual risk and cyberbullying were identified as the most concerning online risks the youth face. In the next section, we discuss how these risks are discovered by the SSPs.

## 4.2 The Often Haphazard Discovery of Online Risks (RQ2)

We also asked SSPs to share the ways they identify when youth have been exposed to online risks. Overall, the SSPs discovered risks through youth self-reports, formal investigations, and risk assessments. However, we found that there is no consistency in methods. This results in SSPs' inability to properly safeguard their clients. We discuss these findings further next.

*4.2.1 SSPs Rely Heavily on Youth Self-reports.* Overall, self-reports were the most common way SSPs (51%) found out about youth's online risk experiences. Self-reports were described as instances in which the youth disclosed the risk experience directly to the provider. It was the primary method for educators (60%) and child welfare workers (57%). Yet it was still prominent among mental health

providers (44%) and juvenile justice officers (45%).

Educators used their relationship with youth to observe their online interaction in classroom settings directly or through the use of school-owned technology (online conferencing software, information technology security software). They also used their relationships that have been built over time to overcome barriers of discourse.

"*...most of the time when you find out the teens are experiencing some kind risks online is because it's through the relationships you [have] built.*" - P10, Education - Public Library Administrator

Child welfare SSPs talked about befriending youth on social media apps thereby being exposed to their online life and learning about areas that could impact youth negatively. Youth being comfortable with the SSPs such that they speak openly about their experiences was also mentioned.

"*...kids kind of just like, let that information out. It's not like a huge secret because it's almost like it's nothing to them like it's just everyday stuff.*" - P24, Child Welfare - Youth Engagement Services

Youth self-reports to mental health professionals occurred mainly in discussions with youth during therapy sessions. Mental health professionals intentionally formed relationships with youth in which they gained their trust thereby laying the foundation necessary for the youth to feel comfortable enough to disclose when they experienced risks.

"*.. one of the biggest things that we do is to spend a lot of time engaging with the child to really build the rapport in that relationship with the child to help put them at ease. ...but the biggest pieces that we spend a lot of time just engaging with them, just talking getting to know them, and really listening to the child.*" - P20, Mental Health - Therapist

When speaking of youth self-reports, juvenile justice officers elaborate on youth disclosing information about their activities or experiences on social media platforms in the form of posts. In some cases, it was stated that youth portray themselves differently online than they do in person to the officers. In other instances, juvenile justice officers stated that teachers, parents, and guardians having good relationships with youth such that they could speak to them about changes in behavior or new acquaintances they had, allowed for the uncovering of risk exposures.

"*But the kids will come in and present one way with us and then you find a different persona online because you see the pictures to prove it; gang signs, weapons, like he said, peers.*" - P2, Juvenile Justice - Prevention and Intervention

A major limitation to this discovery method, however, is distrust. This arose as the major risk discover in the educator (100%), mental health (66.67%), and child welfare (50%) SSP groups, and as the second highest risk discovery limitation in the juvenile justice SSP group. It also dominated the entire risk discovery limitations space capturing 48% of the total limitations identified. Mental health SSPs said that youth withhold information from adults for fear of being punished or upsetting their parents/guardians. They also stated that youth at times provide inaccurate or misleading information for the same reason. Youth not being comfortable enough or feeling safe enough to share their experience with adults was the underpinning theme that arose from the three SSP groups for which lack of trust was the major discovery limitation and was the major discovery limitation theme found within the educator and child welfare groups. Juvenile justice SSPs focused

more on the actions that youth take to cover up their deeds such as deleting applications from their mobile phones or borrowing devices from friends to use instead in attempts to circumvent judicial restrictions. This SSP grouping also identified youth withholding information due to fear of repercussions as the major contributor to risk discovery.

"*I'm relying on the report of someone who may be providing me limited information. Maybe providing inaccurate information may account for varying reasons. So I may not really know, the level of risk that they're experiencing due to those limitations.*" - P17, Mental Health - Youth and Family Counselor

Overall, providers felt that having a close and trustful relationship with the youth allowed the youth to self-disclose risk experiences and find help from the provider.

*4.2.2  To serve and protect: formal investigations uncover risk.*  Formal investigations were the primary method by juvenile justice officers (72%) to discover online risk experiences of youth. Discoveries were said to be made at varying points in the investigative process such as through initial reports made by individuals, operations deliberately set to capture persons committing illegal activity, forensic analysis of devices used by youth, and testimony given during court cases. Mention was also made of discoveries being made through tips submitted to local law enforcement from national or worldwide agencies dedicated to protecting children online. Internet Service Providers (ISP) and social media companies also alert law enforcement when they recognize illegal activity related to youth.

"*...some of the Internet service providers or the social media sites capture that. They would then be required to report it to [the] National Center for Missing Exploited Children...that's how we get the cases to come in.*" - P28, Juvenile Justice - Law Enforcement Officer

The major limitation put forward by juvenile justice offers was the constant changing of technology. In speaking about this, juvenile justice offers spoke about the parents not knowing how to use technology properly, criminals using sophisticated means to hide their activity from authorities, and law enforcement not being able to use social media to monitor youth. They recalled cases where officers were caught by the youth using online systems to check what the youth they were assigned to was doing.

"*...only one in our department out of 42 probation officers that has successfully been able to monitor clients through Facebook. We had another officer that was doing on our sex offender team and totally got caught by the client.*" - P11, Juvenile Justice - Probation Officer

While self reports were the primary discovery method for child welfare workers, 43% of child welfare workers also said formal investigations were common. Investigative methods here were processes related to court cases in which the youth were involved. Sometimes they were also discovered when historical reports were shared between government agencies such as reports being shared with child welfare officers when youth were transitioned to work with them. Generally, child welfare officers shared that they would prefer if this discovery method was a last resort as it often meant the youth was in some sort of trouble with authorities.

"*...sometimes it comes up legally, and that's where it can be even more difficult. So in court someone will say, here's your Instagram account, so I know you're doing illegal activity...*" - P3, Child Welfare -

Children, Youth and Family Services

In discussing online risk exposure discovery limitations, in addition to youth lacking trust in adults, adults not knowing how to use popular applications used by youth was also stated as a limitation. In expressing this concept, child welfare SSPs shared that parents at times don't even know some applications exist or the content that can be accessed by youth who use the application. Adding another perspective to this issue of adults lacking the technological capabilities to recognize online risk possibilities, child welfare SSPs also noted that we tend to be technologically behind those looking to harm youth online.

*"...parents don't even have an understanding of what their teens are using, like the apps or the, whatever they're doing online, they don't even understand the concept of it much less like what their kids are accessing."* - P7, Child Welfare - Youth Development Services

Few participants in the educators SSP grouping (10%) identified formal investigations as an online risk discovery method. In making this identification they spoke of finding out of risk exposures tangentially through the judicial system such as through Child Protective Services (CPS), or law enforcement. These participants tended to be from organizations providing rehabilitative services to youth but who focus on helping the youth obtain an education. Educators did not identify any limitations related to risk discovery through formal investigation.

*"Most... either come directly from law enforcement from children's division which in some places is children's protector's, child protective services and family court..."* - P34, Education Advocacy Services

Mental health professionals did not identify formal investigation as a risk discovery method. They instead identified risk assessment tools as the method by which they discovered risk exposures. This will be discussed in the next section.

*4.2.3   Risk assessment tools: don't ask, don't tell.* On the topic of discovering online risk exposure, mental health SSPs (44%, N=4/9) shared that this was done through the use of risk assessment tools, which on analysis, was found to be equally as important as having a good relationship with youth for this provider group. The risk assessment tools referred to here are case intake forms, survey forms used during meetings with youth, or self-reports that youth fill out so that SSPs can ascertain adherence to program guidelines. Specific professional tools such as the Patient Health Questionnaire-9 (PHQ-9) [49] or Children's Functional Assessment Rating Scale (CFARS) [89] were mentioned as being used to assess the youth in areas such as their feeling of well-being and how they get along with others. SSPs in this group noted that youth at times refuse to answer some sections, leaving them blank, when they don't want to talk about the particular area of their lives thereby limiting the discovery of risk exposure.

*"...we use that to start our research when we have our assessment, with our intake. So with the open cases, we do a clinical assessment, and we have sections on assessing for risk."* - P22, Mental Health - Therapist

This matches what mental health SSPs denoted as the major online risk discovery limitation in the undertaking of their duties, in that youth lack trust in adults to feel comfortable enough to disclose their online experiences. Lack of trust accounted for 67% of the discovery limitations

space in this group. In articulating how this lack of trust is displayed, mental health providers spoke about information being purposely withheld as well as inaccurate information being provided.

"*The biggest limitation would be just honesty. I can, you know, I only get as much information as they're willing to share.*" - P21, Mental Health - Therapist

The juvenile justice SSP group (18%, N=2/11) also identified risk assessment tools as a means of discovering youth risk exposure. Participants spoke about administering these tools as polls when visiting schools, a human trafficking assessment tool [64], and also through a set Community Assessment Tool (CAT) [65] which captures information about youth when they enter the judicial system. Youth withholding information by choosing not to answer questions on or during the administering of these assessments also factors in as a limitation to this group using risk assessment tools for the discovery of youth risk exposure.

"*...every school that we've been in and then we've taken polls, when you talk to these kids say you know who sent dirty pictures nude pictures to somebody, there's a whole lot of kids in that class that will admit to it, way more than you ever anticipate.*" - P26, Retired law enforcement officer

The use of risk assessment tools was not mentioned by the remaining SSP groups. The use of a human trafficking assessment tool to discover risks was mentioned by 14% (N=1/7) of child welfare participants. This tool was said to be used during investigative processes in cases relating to youth that enter the judicial system. None of the limitations identified by child welfare participants were linked to the using risk assessment tools for risk exposure discovery. The use of PHQ-9 assessment tool was mentioned by 10% (N=1/10) of educators. Distrust, and the consequence of youth withholding information due to it, were the only limitations identified by educators. These limitations could affect the quality of data collected using the assessment tools, thereby hampering risk exposure discovery.

This preceding section on risk assessment tools highlights the concept of "*don't ask, don't tell*" that came up in discussions, as youth refrain from talking about situations or experiences if they are not directly asked about them.

### 4.3 The Potential Benefits and Drawbacks of Automated Online Risk Detection (RQ3)

We asked participants about their views on the advantages and disadvantages of technology designed to support automated online risk detection to assist in the identification and prevention of youth online risk exposures. This section synthesizes our findings related to this question grouped by SSP. First, we discuss the prevailing theme that came out in the responses to this research questions: the interplay of online risk identification versus the ability to adequately address them once known. We then discuss the perceived benefits as well as perceived limitations of using such a system as realized by each SSP group.

*4.3.1 If they know, then they have the duty to respond.* A recurring theme highlighted by SSPs was the need to respond to online and offline risks once they have been identified. This was not spoken of in a negative light but they shared that in some cases they do not know who should take responsibility to address risks or threats once they have been identified. For instance, participants shared that schools that have implemented some form of risk identification process often question who should handle threats identified by the processes during vacation periods. One participant shared that schools with this question should be the entity that takes responsibility for such

matters. This raises the issue of liability for some SSPs as given they would rather not be responsible for any negative consequences that may arise from their usage of technology to make decisions.

"*Potentially, we could be liable, because when we are part of our charges, if we're going to assess for something, or if we're going to screen or we're going to monitor something, then we have a duty to respond to it.*" - P2, Juvenile Justice - Prevention and Intervention

Another concern they raised was the lack of resources to properly address online risks and threats that would be identified should technology be used in this area. SSPs spoke about this by mentioning that they are already heavily overwhelmed with other priorities that using another system would not be seen as helpful. The issue raised was that while the SSPs want to do the best job they can, some risks go undetected, or even neglected, due to resource constraints.

"*...we see kids who need more intervention who need more help than what we're able to offer and that the system that is meant to serve and protect them often fails them. ...sometimes our level of service is still not enough. And there's not really an option of where to go...*" - P1, Mental Health - Therapist

Closely linked to the previous issue of liability, they shared concerns that if the system weren't being monitored closely enough and something serious happened with youth, the SSPs would be held responsible, even though constant monitoring is perceived by them as an exhaustive or otherwise labor-intensive pursuit. SSPs raised being understaffed as an issue that leads to youth receiving less help than they need. The result is that youth are left feeling that the system has failed them.

*4.3.2 Technology is useful, but humans are needed.* SSPs involved in child welfare saw monitoring as the most beneficial use of the proposed automatic risk detection system. The monitoring category captured 44% of the possible technology benefits identified by providers in this group. They favored technology that would be able to provide them with more information than they currently get in their formal assessment tools. Recall formal assessment was found to be the second most prevalent online risk discovery method for this provider group. For these providers, having more data about the youth's experience, both online and offline, was viewed as beneficial as it would allow them to possibly get a clearer picture than what is now available. For example, they spoke of using technology to identify when youth create social media posts depicting drug use or other illegal activity which they may later deny or otherwise answer untruthfully during formal assessments. The ability to use such technology on school devices to identify when youth violate policy by using them to access restricted content was also viewed favorably, but with the added concept of alerting the youth towards potential danger. The next major potential benefit of the system as seen by child welfare workers was the better management of resources capturing 33% of the perceived benefits. SSPs who contributed to this potential benefit pointed to the issue of having a limited amount of time to spend with individual youth and welcomed the idea of technology that would reduce the amount of time spent searching for issues that might affect the youth. This would lead to more time spent with the youth to help and support them in problematic areas. Responses from child welfare providers on using automated online risk discovery technology had the undertone of early issue discovery to support quicker interventions as the potential benefit of technology use.

"*I mean, it's gonna flag specific items that would be helpful to see and it would definitely cut down and cut back on time spent on looking at those things.*" - P23, Child Welfare - Youth Engagement

Services

When speaking about the possible limitations of using such technology, the inability of algorithmic solutions to take nuanced views by utilizing the contextual environment of individual youth into consideration was found to be the major limitation identified by child welfare providers. This captured 33% of the limitations identified by child welfare providers. In discussing this they exhibited concern over the usage of broad categorizations that may lead to decisions being made that cause youth to be placed in worse situations than they already exist. Concerns were also raised about the existence and validity of differing parenting styles versus the suggestions made by algorithmic solutions. It was shared that child welfare providers are to be flexible with respect to the different parenting styles rather than require strict adherence to some set of requirements, so there was concern about the utilization of technology output in this regard. Child welfare providers spoke from the perspective of the provider having access to the technology and not the parent/guardian.

"*...we need to be flexible with that, because that's how everybody's raised.*" - P23, Child Welfare - Youth Engagement Services

Child welfare SSPs tended to look at the issue from the perspective of how the technology could best support the youth rather than cause the youth to be implicated in wrongdoing.

*4.3.3 We might be able to make a difference, if we know about the issues.* Providers in the education category identified the early discovery of issues as the major benefit of using technology to automatically discover online risk. This benefit made up 70% of the discovery benefits provided by this provider group. When speaking about this benefit they mostly linked it to possible offline situations it could help mitigate such as the prevention of violence at schools and teen suicide. They expounded on the fact that they can only control or see what happens with the youth within the school setting which does not give them a complete picture of what is affecting the youth. While they could more effectively mitigate what happens within the school, they have little to no insight into what happens outside of school. Using technology to become aware of this gray area was seen in a favorable light. Being able to effectively monitor youth made up the remaining 30% of potential technology benefits space for this SSP group. This benefit was closely related to the early discovery of issues but spoken about more from the perspective of security controls to prevent access to restricted content or methods of communication.

"*... we can't control what they do outside of it as much as we can inside the school. But if we're at least aware of it, we can know that that's something we need to work towards and be aware of.*" - P19, Education - Language Arts Teacher

Like the child welfare workers, educators also identified algorithmic lack of context (38%) to be the major system limitation but they provided different underlying reasons in doing so. Predominately educators thought that technology does not, and should not, parental and teacher relationships with youth. Concern was raised over youth being banned from using technology when people are given more insights into their online experience, a course of action that educators were not in favor of. They acknowledge that youth have legitimate reasons for using technology such as reaching out for support when needed. Educators also identified the possibility of the solution adding complexity to their duties and the solution having access to the data necessary to work as described were also raised limitations by this provider group.

"...*I guess, as far as like electronic safety things in place as your strategy is just not effective*" - P34, Education Advocacy Services

Education SSPs favored a holistic approach to mitigating youth risk with an emphasis on strength-based approaches rather than punishment or criminalization.

*4.3.4 Protection of youth through monitoring requires access to data.* Judicial SSPs thought an automated risk assessment tool would be beneficial in assisting them with monitoring youth online activities (46%). When speaking of the potential benefits, judicial system providers gave examples of monitoring adolescents and online risk victims, assisting parents in monitoring, assisting in the generation of an automatic risk report, predator protection, and identifying illegal conduct. The notion of parents not understanding technology and being able to identify the possible dangers youth are exposed to online was raised by providers here. Providing a means of easily distilling this information to parents was viewed favorably. These benefits were mostly put forward by judicial service providers with current or past work as law enforcement officers. They also viewed knowing what youth experience online as beneficial towards case prioritization. Utilizing technology for investigative purposes on cases intrigued judicial service providers as exemplified in the following participant quote:

"*I think it would be helpful if I was getting the cyber tip if there was some sort of software that had already said yeah this, username unicorn123 is whatever percentage more susceptible to being bullied online or sexually exploited or just based on whatever algorithms you guys are talking about putting together. Then, that would be like just more helpful for me to try to determine which cases are more of a priority*" - P29, Juvenile Justice - Law Enforcement Officer

Being able to discuss with youth topics directly related to their online experience was another perceived benefit of system use by judicial service providers with a more juvenile justice leaning coming in at 23% of the potential benefits space. In this light, it was evident that the monitoring was not to get youth in trouble but instead to provide focused assistance where needed. The monitoring benefits identified by this provider group were for the purpose of assisting in early discovery and better management of resources. SSPs in the judicial system were the only participant grouping to identify driving engagement with the youth as a perceived benefit of the solution.

"*I see it in a way as an opportunity to engage kids a little bit more thoughtfully about what's going on in their lives.*" - P2, Juvenile Justice - Prevention and Intervention

Even though there are perceived benefits to using an automated risk detection system, judicial system service providers also identified drawbacks to using such a system. Concern was raised over the access to the data necessary for the system to function as required. Fears over privacy concerns, internet companies not being willing to share data, as well as parents and teens not being comfortable with sharing data with providers, were raised when providers spoke of this potential limitation. These data access concerns accounted for 33% of the identified limitations for this category of provider:

"*The challenge is going to be them allowing you or your software or whatever access to their world.*" - P35, Juvenile Justice - Retired law enforcement officer currently working with agency that serve underprivileged youth

Concerns were also raised about the exclusion of humans in the decision-making process should such a system be utilized. This made up 22% of the limitations space and was captured as the system making decisions that currently can only be done well with human involvement. Judicial service providers mentioned that technology should not replace humans in this area as relationships with the youth are vital. The following quote from the participant summarizes this concept:

"*once you've met with the family, with the kids, you know, talked to other people that are involved in that kid's life, you can draw a better conclusion as to what actually is going to help that kid because you are not just one part, you're the sum of your parts*" - P16, Juvenile Justice - Government Consultant

Judicial service providers also raised concerns over the system adding complexity to their current duties. The potential for this was raised they reflected on the possible increase of cases to manage, given that the technology identifies more instances of risk exposure from online mediums. They also noted that youth may find ways to outsmart technology thereby rendering the technology useless as in the past when youth realized they were being monitored on their own devices, they would borrow devices from others to engage in risky activity.

"*Potentially, we could be liable, because when we are part of our charges, if we're going to assess for something, or if we're going to screen or we're going to monitor something, then we have a duty to respond to it.*" - P2, Juvenile Justice - Prevention and Intervention

The potential of automatic risk detection to detect risks that might not be responded to due to lack of resources was a cause of concern to judicial SSPs, especially from the perspective of possible liability should something harmful occur.

*4.3.5 Monitoring supports earlier interventions.* Mental Health service providers identified early discovery (50%) as the major perceived benefit of using technology to automatically identify youth risk exposure. They spoke of being able to recognize trends in youth behavior quicker, and this being able to provide the needed support in a more timely manner. They also spoke of it being used to help providers who are not active on social media to get insights into youth experiences. They also raised a similar point with respect to parents as it was noted that not all youth share their negative experiences (such as being cyberbullied), so the technology could be used by parents to get some idea of the challenges being faced by their youth. Monitoring (40%) was another perceived benefit identified by this group. This came across as assistive with respect to early discovery and not as a goal in itself as the prevailing idea from mental health service providers was that the solution would be used by those to gain insights on youth under their care.

"*...I mean, it would give me the ability to intervene sooner, so that significant damage could potentially be could potentially prevented.*" - P17, Mental Health - Youth and Family Therapist

SSPs in mental health viewed the solution as not having access to the data necessary to perform as described to be a major limitation. In talking about data access limitations, mental health therapists raised factors such as social norms, parenting styles, and youth not trusting that their intimate information is kept private even if they agree with the idea of their data being checked for risky material. Regarding the latter, the fear of the youth being punished rather than helped was shared as a concern. Data access limitations captured 55% of the perceived limitations in this group. Algorithmic solutions lacking context (18%) were another limitation category identified by this group. This was shared within discussions about balance being needed between the social structures

youth grow up in and that technology should not be used to replace any of them. Mental health providers also raised the issue of algorithmic solutions encoding bias towards a subset of youth populations. They stated that such systems are not useful as they serve a diverse youth population that would be misrepresented, or under-represented in the biased solution.

## 4.4 SSP Experiences and Needs Differed Based on their Job Roles

Across the SSP groups it was found that the online risks they identified aligned with their job role. In this section, we highlight the interplay between SSP job roles, the risks they identified, and the proactive or reactive nature through which SSPs respond to youth risk exposure. These findings should not be taken as one SSP group performs better or does more for youth than the other. Rather, the focus is brought to the differing points in the lives of youth at which these SSP groups may interact with them, the risks the SSPs focus on, and pre/post risk identification model through which the SSP job function necessitates.

SSPs in the judicial system (juvenile justice and law enforcers) were more concerned with risks that have, or have the potential to be associated with, criminal components. Sexual risks such as prostitution, child sexploitation, child pornography, and human trafficking dominated their concerns. Cyberbullying and its consequences related to physical harm were also of major concern to SSPs in the judicial system. The SSPs closer with duties closer to law enforcement take a more reactive approach to risk management in that they become involved after the exposure has caused an incident, or through investigating other issues which cause the risks to become known. When they took a proactive role it was more investigative, looking for instances where the law is being broken, or actively looking to catch adults attempting to solicit youth for sex. They also took a proactive role in reaching out to schools and parents to raise awareness about online risks, but this is not the main way they currently handle online risks. Ex-law enforcement officers tended to take this more proactive role in terms of speaking to focus groups, church groups, and schools when invited. Comparing this with what judicial SSPs consider as the main success metric related to online risk, minimizing sexual risk, it can be seen that mitigating and handling risks with criminal components was their main focus.

"*To combat people trying to solicit minors for sex and we also do active downloads with software of people who are sharing or downloading child pornography on the Internet.*" - P29, Law Enforcement Officer

SSPs in the child welfare group preferred to catch youth risk exposure, or actions related to risks, to mitigate or eliminate them where possible, *before* law enforcement was involved or before further interaction with the judicial system. They preferred this to avoid youth being implicated and being placed in situations that may not be in their favor (i.e. not being able to receive the help they need). For child welfare SSPs, part of this is also accounting for the different parenting styles of parents/guardians of the youth they interact with. For instance, mention was made of the different levels of leniency that parents/guardians have (i.e. on the topic of youth phone ownership) and the fact that child welfare SSP officers should not override the decisions of parents. This being the case, child welfare SSPs play an active role in the day-to-day lives of youth under their care and therefore took a more proactive approach in the lives of youth. This more active role was to ensure that youth obtain an education and find employment to sustain themselves after school is finished. This aligns with the main success metric identified by child welfare SSPs, which was educating youth with the supporting success metric, and engaging more with youth to build relationships

being second.

"...we want to make sure that our kids are graduating high school and on a timely basis, that they're enrolled in college in some kind of post-secondary or college or technical program, that they're having some kind of exposure to employment before they turn 18." - P23, Child Welfare - Youth Engagement Services

Mental health SSPs become part of the lives of youth through direct engagement by parents/guardians, or through referrals made by the judicial system or other agencies. They remain part of the life of youth until treatment has ended which may be due to various reasons. Due to this they initially play a more reactive role in the identification and mitigation of online risks which is also seen through their dependence on risk assessment tools to discover risk exposure. Once identified, they attempt to mitigate the *consequences* of risk exposure by helping youth through the negative symptoms they experience such as anxiety, and depression. This aligns with the main success goals of mental health SSPs, the improvement of the mental health of youth, and helping them to not re-enter the judicial system.

"So, you know, are the youth recidivating, you know, ...the whole purpose of us being a provider for the department is to prevent [youth] from catching the charges and being placed in residential programs." - P23, Mental Health - Youth and Family Therapist

Through the education system, education SSPs are introduced into the lives of youth from an early age and continue to be part of their lives for as long as youth remain in school. They play both proactive and reactive roles as it pertains to risk discovery and mitigation in the lives of youth. They proactively try to mitigate risks, including online risks, by teaching youth life skills in a holistic manner. Reactively, they are also able to respond to changes in youth, such as their behavioral patterns or recognizing instances of physical harm, to identify when something significant has taken place in the life of the youth. The education of youth was the main success metric of the educator SSP group. The risks they identified were spoken about from the perspective of what causes harm to youth, and what causes them to not be present physically, or in the frame of mind to receive, a sound education.

"Oh, yeah, we teach health, ...we start off with the basics of what health actually means. So what that means to your body, so we go through mental health, we go through physical health, and then we go through social health." - P14, Education - Health and Wellness Teacher

The findings show that underprivileged youth, as seen through the lens of SSPs, are affected by various online risks which have can be of a sexual nature, expose them to cyberbullying, and can result in addictive internet usage. There exists some interplay between these risks which SSP can recognize, but across SSP groups, it is done in an inconsistent manner. In the next section, we will present a discussion about these findings, and their implications.

## 5 DISCUSSION

In this section, we discuss the major takeaways of our findings and provide design implications for future online risk detection systems. We conclude with a discussion of the limitations of our work and future research directions.

## 5.1 Online Risks Negatively Impact SSP's Ability Protect Youth Offline, But Are Not A Focus for Risk Prevention Efforts

For RQ1, we uncovered several ways in which online risks elevate offline risks that put youth in harm's way and negatively impact various youth outcomes. SSPs linked online sexual risks with youth meeting strangers, prostitution, and human trafficking, all offline risks stemming from interactions youth have online with adults. The identification of the *online* risks is where our findings exposed challenges. For RQ2 we found that online risk factors were usually not included in youth risk assessment tools and mechanisms currently used by SSPs when evaluating teen risk factors. We also found that SSPs rely predominately on youth self-disclosure to uncover online risk exposure. While SSPs address issues in individuals which arise from human trafficking, they lack the knowledge of how technology is used to facilitate the business of human trafficking, and also the language used online by traffickers to lure their victims [39], even though these interactions are predominately initiated by means of online communication [11]. SSPs therefore mostly retroactively address the issues which begin online, or otherwise have online components which are exposed. However, it is accepted that the online environment is part of the expanded ecological environment [19] within which youth reside and influences them [61]. This risk identification gap inhibits SSPs from effectively performing their duties thereby making it harder to improve youth outcomes. One way to bridge this divide is to incorporate online risk assessment tools, and/or their outputs, in the risk assessment evaluations used by SSPs. Since SSPs examine the ecological environment within which youth reside in-order to improve youth outcomes, and the online environment is part of that ecological environment, more attention should be paid to how the online environment can be used to safeguard youth online interactions through the lens of SSPs. As a community, we will have to consider including online technology in the network of support available to youth so that it can be appropriately strengthened through research and development. The identification of risks youth face online is one such area. We will discuss this in the following section.

## 5.2 Early Discovery of Online Risks Is Key, But It Is Easier Said Than Done

In RQ2 it was found that youth risk exposure was mostly identified through self-disclosure aided by SSPs having good relationships with youth. Care should be taken that outputs from automatic risk detection systems are part of fact-finding processes utilized by those who use the outputs in decision making. The output from such systems should be used as factors or points of interest, rather than absolute truth since the context within which youth experience and react to incidents matters [40]. The outputs from algorithmic risk detection systems should therefore be designed to nudge youth towards healthier interactions whilst providing discussion points to SSPs and parents. Technology should not replace the relationship between youth, SSPs, and parents. It should be used to promote conversation with youth. Rather than be designed to get youth in trouble by use as a surveillance tool, risk detection technology should be used to promote discussion with youth, thereby building and re-enforcing relationships through which risk exposures are discovered. In order to achieve this, algorithmic risk detection systems need to allow SSPs to focus on the risks pertinent to their duties or areas of concern in a particular time frame.

Mitigating youth online risk exposure in a timely, strength-based manner requires automatic online risk detection frameworks to notify SSPs of online risk exposure as promptly as possible. Our findings show that SSPs heavily rely on youth self-disclosure after exposure to risk, which can be left undisclosed for lengthy periods. Technology can help them to discover youth risk exposure closer to incident occurrence and aid in the more timely implementation of interventions. Reactive exposure-response chronology sometimes can occur long after initial risk exposure and, sadly at times, long after the damage has been done. Being able to identify risk exposures sooner and see

trends in youth behavior would allow SSPs to use technology as an early warning system to drive conversations with youth, or keep them out of trouble by providing direction. Earlier discovery of issues affecting youth leads to better outcomes [40]. A possible step in this direction is to integrate automated risk detection technologies into solutions currently used by SSPs to receive incident tips. This additional information can then be used to prioritize cases such that more concerning cases can be dealt with in a timely manner. Similarly, risk assessment tools can be integrated with automatic risk detection technology, allowing risks to incidents to be found sooner, thereby decreasing the time it takes to discover risk.

## 5.3  AI Risk Detection Could Help, But It Could Also Cause More Problems

In RQ3, we found that SSPs believed that AI risk detection could potentially assist them in the early discovery of potential issues, monitoring youth interactions, and enabling better resource management. However, while the utilization of automatic risk detection technology to discover youth risk exposure was seen as a possible benefit to assist in the proactive mitigation of youth online risks, concerns such as SSP liability, replacement of humans in the decision-making process, and the use of technology to implicate youth, rather than help them manage online risks, needs to be addressed prior to such technology being deployed in the real-world. In particular, most SSPs felt strongly that these tools should not be used to replace their interactions with youth and that caution should be taken to ensure that such tools are not utilized in a big brother-style surveillance scenario. Doing so could erode trust and undermine their relationships with youth. This was a particularly salient theme among educators, mental health, clinicians, and educators who relied heavily on building trusting relationships with youth to optimally perform their job duties. This was less pertinent for law enforcement officers who were less reliant on building long-term relationships with underprivileged youth to perform their job duties effectively.

Our findings reflect some of the same sentiments that have recently surfaced more broadly from the literature on adolescent online safety and algorithmic decision-making in the public sector. First, children, particularly teens, value their personal privacy [93] and resent surveillance-based technology designed to monitor and restrict their online behaviors [31, 91]. Instead, technologies that promote increased risk awareness, collaboration, and open communication between youth and those charged with their safety are recommended [2, 77]. In regard to the use of algorithmic risk detection, SSPs preferred a tool that would assist with risk assessments but allowed them the discretion on how best to respond, rather than a prescriptive tool that required them to escalate to measures that could potentially harm youth long-term (e.g., mandated reporting, probation violations, etc.). Most SSPs, including law enforcement officers, felt that a tool that did not allow them some discretion ran the risk of punishing underprivileged youth instead of protecting them. These concerns align closely with Saxena et al.'s [80] guidelines for designing high-stakes algorithmic decision-making systems for the public sector where trade-offs must be made with respect to human discretion, bureaucratic processes, and algorithmic decision-making. Emphasis should be placed on strength-based approaches in which algorithmic outputs would be a component to help SSPs and underprivileged youth manage risks together, rather than to police them or get them in trouble. This view builds upon the growing body of adolescent online safety research that advocates for resilience, rather than abstinence-only based approaches, to allow youth to self-regulate their online activities but with guidance towards safer online interactions [25, 91, 92].

In order for algorithmic risk detection solutions to work effectively, it necessitates access to data needed to make correct risk assessments. In addition, to maintain up-to-date risk assessments, the data used by solutions must also be kept current, or otherwise be brought in alignment with the current content accessed by, or created by youth, as exemplified in recent research [72]. This raises further ethical concerns regarding the amount of access, and volume of data, needed for

these types of solutions to work optimally. There are also concerns regarding *who* has access to the data once it is collected, the *persistence* (the length of time stored before being discarded, if at all) of that data, and *how* the data will be used. The latter is a growing concern with companies selling or sharing personal data [31], the judicial system sequestering data from companies [16], and countries making international data handling laws for their citizens [30, 67, 96]. For instance, Brown et al. [20] cites the collection of negative data-points and the exclusion of positive data-points as a major criticism of data being used to in algorithmic decision making for underprivileged youth. These concerns are further compounded by SSPs that raised the concern of parents/guardians not being comfortable disclosing private information about youth under their care to others as articulated in our results. These are all serious ethical and privacy challenges that face creators of automatic risk detection solutions, and those who wish to use them. As such, creators of such systems should design their systems with these concerns in mind, possibly allowing opt-in/out functionality, and clearly addressing the concerns (who, how, and for how long) to the users of such systems. One potential way to do this well is through using participatory methods for the design of AI risk detection systems that directly engage users and/or key stakeholders [9, 97], including SSPs, underprivileged youth, and their parents and/or legal guardians.

## 5.4 Design Implications

In this section, we address the design implications uncovered through our study. These implications should be taken under consideration by solution designers of youth automatic online risk detection systems.

*5.4.1 Implications for Risk Prioritization: Sexual Risk and Cyberbullying.* Of particular concern, the identification of exposure to sexual risk and cyberbullying should be prioritized by algorithmic systems. The research community has recently started to pay closer attention to the identification of these types of risks and improving the mechanisms by which identification is done by taking human-centered approaches. [46] [74]. Youth with addictive internet tendencies are more likely to be involved in online and offline sexually risky behavior, and also are more likely to live in less than ideal home circumstances [15]. Studies also point to online sexploitation, a sexual risk, being the fastest growing online risk experienced by youth in the US [70]. Youth who exhibit internet addiction also suffer higher rates of both offline bullying and cyberbullying [88] leading to stress, anxiety, and the display of traits related to low self-esteem [3]. SSPs in our study identified cyberbullying cases that led to serious cases of self-harm and even suicide, showing the importance of being able to detect this risk and implement appropriate interventions. While studies have shown that digital interventions have a relatively low impact on the reduction of cyberbullying [23], the implication for automated risk detection systems is that they should be used to facilitate strength-based approaches. In addition, to effectively handle the cases of most concern, automated risk detection systems need to support the filtering of risk cases based on different risk levels as needed by SSPs. We will address this in the next section.

*5.4.2 Implications for Thresholds: Support Differing Views and Risk Level Settings for SSPs.* To support timely youth risk exposure interventions, automatic online risk detection systems should show information to SSPs in a manner that is curated to support the duties as defined for each target SSP group. Since risk factors can be possibly aggregated from many sources, some of which may contain sensitive information, technology that targets multiple user groups should distill the information and support the display of only what is necessary for the group to perform their task or reach their desired goal. For example, judicial system users need views that support investigation, prioritization, and incident response. They care about the detection and prevention of illegal behavior. In contrast, educators and child welfare officers need a more day-by-day visualization of

the individual experience of specific teens. Clinicians, therapists, and mental health practitioners, mainly want to see assessments of their patients to correlate findings with their current established means of patient assessment towards the identification of factors that indicate poor mental health. Designers of algorithmic solutions would do well to understand the interplay of what they can detect and how it correlates with risk assessment tools such as the Patient Health Questionnaire-9 (PHQ-9) [49] or Children's Functional Assessment Rating Scale (CFAR) [89], which were mentioned by SSPS in our study. Other use cases which necessitate different views are providing useful nudges to patients towards healthier directions in just-in-time solutions and providing useful talking points to parents that can assist them in conversations with youth about their online experiences. In addition to providing differing views to the SSP groups, algorithmic solutions should also support the setting of risk severity thresholds. This would allow SSPs in various groups to prioritize cases leading to better resource management and timely intervention implementation. It is our recommendation that designers of automatic online risk detection solutions utilize participatory design and work with SSPs to capture, process, and display the required information for the different views and thresholds needed by different SSP groups. Members of the SIGCHI community have started work in this direction [22].

### 5.5 Limitations and Future Work

There are several limitations to our study that present areas for future research. First, our interview questions were initially designed to address the core purpose of the NSF I-Corps program, which was to identify viable a business model for commercializing a technological solution (i.e, risk detection algorithms for adolescent online safety) that would address customers' (broadly defined) needs. Therefore, RQ3, which examined the use of AI risk detection for youth protection was *a priori*, rather than an emergent finding from the research. However, examining the potential for AI risk detection within the unique context of SSPs and the online safety of underprivileged youth was a novel area of inquiry that allowed us to identify potential benefits and pitfalls prior to bringing such a technology to market. Future work should build on our findings and heed the concerns identified by SSPs as warnings on how not to design such technologies. Furthermore, this study was limited to a subset of SSPs in the U.S. and therefore does not represent the entire worldview. SSPs in other countries may have different challenges, cultural norms, and/or regulations that may not generalize to the U.S. context. Finally, this study did not investigate the effectiveness, technical, or subjective user satisfaction of any automated risk detection solutions. Instead, we presented the abstract idea of an AI risk detection dashboard, which we allowed SSPs to iterate on as a potential solution for their own context. Therefore, in future work, we plan to develop a prototype of an actual risk detection dashboard to be evaluated by SSPs and underprivileged youth, so that they can provide more tangible feedback.

### 6 CONCLUSION

Exposure to online risks is an issue that SSPs identify as important in the youth populations they serve. SSPs are part of the support system youths depend on but often find themselves under-resourced thereby making it difficult to adequately perform their duty to assist youth. Automation through the use of technology such as decision support systems has been used to address some aspects of this such as reducing workload. These are at times not designed based on the specific needs of social service providers as they experience in the field. With respect to online risk, this is of particular interest as SSPs currently predominantly rely on some form of human disclosure through conversation or tip-offs. In this study, we have found that SSPs view risk detection technology as useful in providing insights into the youth online experience, but such technology should be used mainly to sponsor conversations with youth rather than to perform surveillance on them. Risk

detection technology should therefore support different levels of end-user views to address the differing needs of the distinct types of social service providers.

## ACKNOWLEDGMENTS

## REFERENCES

[1] Denielle Kirk L Abaquita. 2020. Understanding the Challenges Child Welfare Workers Encounter Related to Promoting the Online Safety of Foster Youth. (2020).

[2] Mamtaj Akter, Amy J. Godfrey, Jess Kropczynski, Heather R. Lipford, and Pamela J. Wisniewski. 2022. From Parental Control to Joint Family Oversight: Can Parents and Teens Manage Mobile Online Safety and Privacy as Equals? *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 57 (apr 2022), 28 pages. https://doi.org/10.1145/3512904

[3] Yara Al Tavash, Nofe Al Mussa, Alhanouf AlAjaji, and Nor Shahida Mohd Jamail. 2022. Effects Cyberbullying of Children and Prevention Methods. In *2022 Fifth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU)*. IEEE, 133–137.

[4] Monica Anderson and Jingjing Jiang. 2018. Teens, Social Media & Technology 2018 | Pew Research Center. http://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/

[5] Karla Badillo-Urquiola, Jaclyn Abraham, Arup Kumar Ghosh, and Pamela Wisniewski. 2018. A Stakeholders' Analysis of the Systems that Support Foster Care. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork*. 158–161.

[6] Karla Badillo-Urquiola, Scott Harpin, and Pamela Wisniewski. 2017. Abandoned but not forgotten: Providing access while protecting foster youth from online risks. In *Proceedings of the 2017 Conference on Interaction Design and Children*. 17–26.

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

[8] Karla Badillo-Urquiola, Afsaneh Razi, Jan Edwards, and Pamela Wisniewski. 2020. Children's Perspectives on Human Sex Trafficking Prevention Education. In *Companion of the 2020 ACM International Conference on Supporting Group Work*. 123–126.

[9] Karla Badillo-Urquiola, Zachary Shea, Zainab Agha, Irina Lediaeva, and Pamela Wisniewski. 2021. Conducting risky research with teens: co-designing for the ethical treatment and protection of adolescents. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–46.

[10] Karla Badillo-Urquiola and Pamela Wisniewski. 2017. For Those Who Need it Most: Helping'At-Risk'Youth Manage Online Risks. In *Badillo-Urquiola, K.\*, Wisniewski, P.†(2017)"For Those Who Need It Most: Helping "At-Risk" Youth Manage Online Risks," Extended Abstract presented at the Workshop on Equity and Inclusivity at the 2017 ACM Conference on Interaction Design and Children (IDC 2017), Stanford, CA.*

[11] Kyla Baird and Jennifer Connolly. 0. Recruitment and Entrapment Pathways of Minors into Sex Trafficking in Canada and the United States: A Systematic Review. *Trauma, Violence, & Abuse* 0, 0 (0), 15248380211025241. https://doi.org/10.1177/15248380211025241 arXiv:https://doi.org/10.1177/15248380211025241 PMID: 34184579.

[12] Anna Costanza Baldry, David P Farrington, and Anna Sorrentino. 2017. School bullying and cyberbullying among boys and girls: Roles and overlap. *Journal of Aggression, Maltreatment & Trauma* 26, 9 (2017), 937–951.

[13] Fanni Bányai, Ágnes Zsila, Orsolya Király, Aniko Maraz, Zsuzsanna Elekes, Mark D Griffiths, Cecilie Schou Andreassen, and Zsolt Demetrovics. 2017. Problematic social media use: Results from a large-scale nationally representative adolescent sample. *PloS one* 12, 1 (2017), e0169839.

[14] Burt S Barnow, Amy Buck, Kirk O'Brien, Peter Pecora, Mei Ling Ellis, and Eric Steiner. 2015. Effective services for improving education and employment outcomes for children and alumni of foster care service: Correlates and educational and employment outcomes. *Child & Family Social Work* 20, 2 (2015), 159–170.

[15] Susanne E Baumgartner, Sindy R Sumter, Jochen Peter, and Patti M Valkenburg. 2012. Identifying teens at risk: Developmental pathways of online and offline sexual risk behavior. *Pediatrics* 130, 6 (2012), e1489–e1496.

[16] Johana Bhuiyan. 2022. Facebook gave police their private data. Now, this duo face abortion charges. https://www.theguardian.com/us-news/2022/aug/10/facebook-user-data-abortion-nebraska-police (Accessed on 08/15/2022).

[17] Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry. 2019. *Thematic Analysis*. Springer Singapore, Singapore, 843–860. https://doi.org/10.1007/978-981-10-5251-4_103

[18] Urie Bronfenbrenner. 1975. Reality and research in the ecology of human development. *Proceedings of the American Philosophical Society* 119, 6 (1975), 439–469.

[19] Urie Bronfenbrenner. 1979. *The ecology of human development: Experiments by nature and design*. Harvard university press.

[20] Anna Brown, Alexandra Chouldechova, Emily Putnam-Hornstein, Andrew Tobin, and Rhema Vaithianathan. 2019. Toward algorithmic accountability in public services: A qualitative study of affected community perspectives on algorithmic decision-making in child welfare services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[21] Jo Bryce and James Fraser. 2014. The role of disclosure of personal information in the evaluation of risk and trust in young peoples' online interactions. *Comput. Hum. Behav.* 30 (2014), 299–306.

[22] Xavier V Caddle, Afsaneh Razi, Seunghyun Kim, Shiza Ali, Temi Popo, Gianluca Stringhini, Munmun De Choudhury, and Pamela J. Wisniewski. 2021. MOSafely: Building an Open-Source HCAI Community to Make the Internet a Safer Place for Youth. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (Virtual Event, USA) *(CSCW '21)*. Association for Computing Machinery, New York, NY, USA, 315–318. https://doi.org/10.1145/3462204.3481731

[23] Qiqi Chen, Ko Ling Chan, Shaolingyun Guo, Mengtong Chen, Camilla Kin-Ming Lo, and Patrick Ip. 2022. Effectiveness of digital health interventions in reducing bullying and cyberbullying: a meta-analysis. *Trauma, Violence, & Abuse* (2022), 15248380221082090.

[24] Stephanie Cuccaro-Alamin, Regan Foust, Rhema Vaithianathan, and Emily Putnam-Hornstein. 2017. Risk assessment and decision making in child protective services: Predictive risk modeling in context. *Children and Youth Services Review* 79 (2017), 291–298.

[25] Leen d'Haenens, Sofie Vandoninck, and Verónica Donoso. 2013. How to cope and build online resilience? (2013).

[26] Aiman El Asam and Adrienne Katz. 2018. Vulnerable young people and their experience of online risks. *Human–Computer Interaction* 33, 4 (2018), 281–304.

[27] Dafna Etzion and Shlomo Romi. 2015. Typology of youth at risk. *Children and Youth Services Review* 59 (2015), 184–195. https://doi.org/10.1016/j.childyouth.2015.10.017

[28] Helen L Fisher, Terrie E Moffitt, Renate M Houts, Daniel W Belsky, Louise Arseneault, and Avshalom Caspi. 2012. Bullying victimisation and risk of self harm in early adolescence: longitudinal cohort study. *Bmj* 344 (2012), e2683.

[29] Asbjørn William Ammitzbøll Flügge, Thomas Hildebrandt, and Naja Holten Møller. 2020. Algorithmic decision making in public services: A CSCW-perspective. In *Companion of the 2020 ACM International Conference on Supporting Group Work*. 111–114.

[30] Werner Geyser. 2021. GDPR and Social Media: What Data Protection and Privacy Mean for Social Media Marketers. https://influencermarketinghub.com/gdpr-social-media/ (Accessed on 08/15/2022).

[31] Arpita Ghosh and Aaron Roth. 2015. Selling privacy at auction. *Games and Economic Behavior* 91 (2015), 334–346. https://doi.org/10.1016/j.geb.2013.06.013

[32] Jeremy D Goldhaber-Fiebert and Lea Prince. 2019. Impact evaluation of a predictive risk modeling tool for Allegheny county's child welfare office. *Pittsburgh: Allegheny County* (2019).

[33] Simon Gonsoulin and NW Read. 2011. Improving Educational Outcomes for Youth in the Juvenile Justice and Child Welfare Systems through Interagency Communication and Collaboration. Practice Guide. *National Evaluation and Technical Assistance Center for the Education of Children and Youth Who Are Neglected, Delinquent, or At-Risk* (2011).

[34] Bethany Good and Faye Mishna. 2021. "Double-edged sword" of digital media use among youth in residential treatment: Perspectives of service providers. *Residential Treatment for Children & Youth* 38, 2 (2021), 198–222.

[35] Heidi Hartikainen, Afsaneh Razi, and Pamela Wisniewski. 2021. Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–31.

[36] Nicola Henry and Anastasia Powell. 2018. Technology-facilitated sexual violence: A literature review of empirical research. *Trauma, violence, & abuse* 19, 2 (2018), 195–208.

[37] Kashmir Hill. 2020. Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match. https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html

[38] Alexis Hiniker, Sarita Y Schoenebeck, and Julie A Kientz. 2016. Not at the dinner table: Parents' and children's perspectives on family technology rules. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 1376–1389.

[39] Raquel Monique Holguin and Athena Noel Barber. 2018. Social Service Workers Knowledge on the Use of Technology for Human Trafficking. (2018).

[40] Cheryl Honeyman. 2007. Recognising mental health problems in children and young people. *Nursing Children and Young People* 19, 8 (2007).

[41] Charlotte Jee and William Douglas Heaven. 2021. The therapists using AI to make therapy better. https://www.technologyreview.com/2021/12/06/1041345/ai-nlp-mental-health-better-therapists-psychology-cbt/

[42] Mark Jones. 2022. The Next Evolution of Campus Security. https://campuslifesecurity.com/articles/2022/05/23/the-next-evolution-of-campus-security.aspx

[43] Bruce Kamradt. 2002. Funding Mental Health Services for Youth in the Juvenile Justice System: Challenges and Opportunities.

[44] Emily Keddell. 2019. Algorithmic justice in child protection: Statistical fairness, social justice and the implications for practice. *Social Sciences* 8, 10 (2019), 281.

[45] Suzanne EU Kerns, Michael D Pullmann, Barbara Putnam, Anne Buher, Sarah Holland, Lucy Berliner, Ellen Silverman, Lin Payton, LaRessa Fourre, Dae Shogren, et al. 2014. Child welfare and mental health: Facilitators of and barriers to connecting children and youths in out-of-home care with effective mental health treatment. *Children and Youth Services Review* 46 (2014), 315–324.

[46] Seunghyun Kim, Afsaneh Razi, Gianluca Stringhini, Pamela J Wisniewski, and Munmun De Choudhury. 2021. A Human-Centered Systematic Literature Review of Cyberbullying Detection Algorithms. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–34.

[47] Susan J Ko, Julian D Ford, Nancy Kassam-Adams, Steven J Berkowitz, Charles Wilson, Marleen Wong, Melissa J Brymer, and Christopher M Layne. 2008. Creating trauma-informed systems: Child welfare, education, first responders, health care, juvenile justice. *Professional psychology: Research and practice* 39, 4 (2008), 396.

[48] John Koetsier. 2022. Doctors Using AI, Supercomputer To Predict And Prevent 50% Of Mental Illness. https://www.forbes.com/sites/johnkoetsier/2022/03/02/doctors-using-ai-supercomputer-to-predict-and-prevent-50-of-mental-illness/?sh=51b3dbf26bb7

[49] K Kroenke, RL Spitzer, and JBW Williams. 2001. The patient health questionnaire (phq-9)–overview. *J. Gen. Intern. Med* 16 (2001), 606–616.

[50] Joan Laird and Ann Hartman. 1985. *A handbook of child welfare: Context, knowledge, and practice.* Simon and Schuster.

[51] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research Methods in Human-Computer Interaction* (2 ed.). Morgan Kaufmann, Cambridge, MA. https://www.safaribooksonline.com/library/view/research-methods-in/9780128093436/

[52] Suzet Tanya Lereya, Catherine Winsper, Jon Heron, Glyn Lewis, David Gunnell, Helen L Fisher, and Dieter Wolke. 2013. Being bullied during childhood and the prospective pathways to self-harm in late adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry* 52, 6 (2013), 608–618.

[53] Sonia Livingstone, Magdalena Bober, and Ellen Helsper. 2005. Internet literacy among children and young people: Findings from the UK Children Go Online Project. (2005).

[54] Sonia Livingstone and Ellen Johanna Helsper. 2008. Parental Mediation of Children's Internet Use. *Journal of Broadcasting & Electronic Media* 52 (2008), 581 – 599.

[55] Jackie Loffell, Merle Allsopp, Eric Atmore, and Jo Monson. 2008. Human resources needed to give effect to children's right to social services.

[56] Jennifer A Nekuda Malik. 2016. US administration makes efforts to increase inclusiveness in STEM industries: www.whitehouse.gov/demo-day. *MRS Bulletin* 41, 5 (2016), 355–356. https://doi.org/10.1557/mrs.2016.103

[57] Tina Maschi, Schnavia Smith Hatcher, Craig S Schwalbe, and Nancy Scotto Rosato. 2008. Mapping the social service pathways of youth to and through the juvenile justice system: A comprehensive review. *Children and Youth Services Review* 30, 12 (2008), 1376–1385.

[58] Nora McDonald and Andrea Forte. 2022. *Privacy and Vulnerable Populations.* Springer International Publishing, Cham, 337–363. https://doi.org/10.1007/978-3-030-82786-1_15

[59] J Jeffries McWhirter, Benedict T McWhirter, Ellen Hawley McWhirter, and Anna C McWhirter. 2016. *At risk youth.* Cengage Learning.

[60] Diana J Meter and Sheri Bauman. 2015. When sharing is a bad idea: The effects of online social network engagement and sharing passwords with friends on cyberbullying involvement. *Cyberpsychology, Behavior, and Social Networking* 18, 8 (2015), 437–442.

[61] Faye Mishna, Jane E. Sanders, Karen M. Sewell, and Elizabeth Milne. 2021. Teaching Note—Preparing Social Workers for the Digital Future of Social Work Practice. *Journal of Social Work Education* 57, sup1 (2021), 19–26. https://doi.org/10.1080/10437797.2021.1912676 arXiv:https://doi.org/10.1080/10437797.2021.1912676

[62] FL: Department of Children and Families. 2014. Baker Act. https://www.myflfamilies.com/service-programs/samh/crisis-services/laws/BakerActManual.pdf

[63] US. Department of Justice. 2014. Predictive Analytics in Law Enforcement: A Report by the Department of Justice. https://epic.org/wp-content/uploads/foia/doj/criminal-justice-algorithms/EPIC-16-06-15-DOJ-FOIA-20200319-

Settlement-Production-pt1.pdf

[64] FL: Department of Juvenile Justice. 2021. Community Assessment Tool. https://www.djj.state.fl.us/content/download/22467/file/human-trafficking-screening-tool.pdf

[65] FL: Department of Juvenile Justice. 2021. Community Assessment Tool. https://www.djj.state.fl.us/partners-providers-staff/our-approach/risk-assessment-tools/community-assessment-tool

[66] International Federation of Social Workers. 2014. Global Definitions of Social Work. https://www.ifsw.org/what-is-social-work/global-definition-of-social-work/

[67] Office of the Attorney General. 2022. California Consumer Privacy Act (CCPA). https://oag.ca.gov/privacy/ccpa (Accessed on 08/15/2022).

[68] Olin Oldeide, Ingrid Holsen, and Elisabeth Fosse. 2020. Youth perspective on outreach service: A safety net for at-risk youth in a municipality. *Children and Youth Services Review* 116 (2020), 105234. https://doi.org/10.1016/j.childyouth.2020.105234

[69] Abigail Ornellas, Gary Spolander, and Lambert K Engelbrecht. 2018. The global social work definition: Ontology, implications and challenges. *Journal of Social Work* 18, 2 (2018), 222–240.

[70] Justin W Patchin and Sameer Hinduja. 2020. Sextortion among adolescents: Results from a national survey of US youth. *Sexual Abuse* 32, 1 (2020), 30–54.

[71] Anthony T Pinter, Pamela J Wisniewski, Heng Xu, Mary Beth Rosson, and Jack M Caroll. 2017. Adolescent online safety: Moving beyond formative evaluations to designing solutions for the future. In *Proceedings of the 2017 Conference on Interaction Design and Children*. 352–357.

[72] Afsaneh Razi, Ashwaq AlSoubai, Seunghyun Kim, Nurun Naher, Shiza Ali, Gianluca Stringhini, Munmun De Choudhury, and Pamela J Wisniewski. 2022. Instagram Data Donation: A Case Study on Collecting Ecologically Valid Social Media Data for the Purpose of Adolescent Online Risk Detection. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–9.

[73] Afsaneh Razi, Karla Badillo-Urquiola, and Pamela J. Wisniewski. 2020. Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, Honolulu, HI, USA, 1–13. https://doi.org/10.1145/3313831.3376400

[74] Afsaneh Razi, Seunghyun Kim, Ashwaq Alsoubai, Gianluca Stringhini, Thamar Solorio, Munmun De Choudhury, and Pamela J Wisniewski. 2021. A Human-Centered Systematic Literature Review of the Computational Approaches for Online Sexual Risk Detection. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–38.

[75] Gary Resnick and Martha R. Burt. 1996. Youth at risk: Definitions and implications for service delivery. *American Journal of Orthopsychiatry* 66, 2 (1996), 172–188. https://doi.org/10.1037/h0080169 Place: US Publisher: American Orthopsychiatric Association, Inc..

[76] Michael J Rosenfeld, Reuben J Thomas, and Sonia Hausen. 2019. Disintermediating your friends: How online dating in the United States displaces other ways of meeting. *Proceedings of the National Academy of Sciences* 116, 36 (2019), 17753–17758.

[77] Tara L. Rutkowski, Heidi Hartikainen, Kirsten E. Richards, and Pamela J. Wisniewski. 2021. Family Communication: Examining the Differing Perceptions of Parents and Teens Regarding Online Safety Communication. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 373 (oct 2021), 23 pages. https://doi.org/10.1145/3479517

[78] Jackie Sanders, Robyn Munford, Tewaporn Thimasarn-Anwar, Linda Liebenberg, and Michael Ungar. 2015. The role of positive youth development practices in building resilience and enhancing wellbeing for at-risk youth. *Child abuse & neglect* 42 (2015), 40–53.

[79] Devansh Saxena, Karla Badillo-Urquiola, Pamela J Wisniewski, and Shion Guha. 2020. A Human-Centered Review of Algorithms used within the US Child Welfare System. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

[80] Devansh Saxena, Karla Badillo-Urquiola, Pamela J. Wisniewski, and Shion Guha. 2021. A Framework of High-Stakes Algorithmic Decision-Making for the Public Sector Developed through a Case Study of Child-Welfare. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 348 (oct 2021), 41 pages. https://doi.org/10.1145/3476089

[81] Douglas Schuler and Aki Namioka. 1993. *Participatory design: Principles and practices*. CRC Press.

[82] Wonsun Shin and May Oo Lwin. 2017. How does "talking about the Internet with others" affect teenagers' experience of online risks? The role of active mediation by parents, peers, and school teachers. *New Media & Society* 19 (2017), 1109 – 1126.

[83] Robert Slonje, Peter K Smith, and Ann Frisén. 2017. Perceived reasons for the negative impact of cyberbullying and traditional bullying. *European journal of developmental psychology* 14, 3 (2017), 295–310.

[84] Gordon Tait. 1995. Shaping the 'at-risk youth': risk, governmentality and the Finn Report. *Discourse: studies in the cultural politics of education* 16, 1 (1995), 123–134.

[85] Elizabeth Thompson. 2021. U.S. technology company Clearview AI violated Canadian privacy law: report. https://www.cbc.ca/news/politics/technology-clearview-facial-recognition-1.5899008

[86] Ma Dolores C Tongco. 2007. Purposive sampling as a tool for informant selection. *Ethnobotany Research and applications* 5 (2007), 147–158.

[87] Michael Veale and Irina Brass. 2019. *Administration by algorithm? Public management meets public sector machine learning.* Oxford University Press, 121–151. https://doi.org/10.1093/oso/9780198838494.001.0001

[88] Judith A Vessey, Rachel L Difazio, Laura K Neil, and Anna Dorste. 2022. Is There a Relationship Between Youth Bullying and Internet Addiction? An Integrative Review. *International Journal of Mental Health and Addiction* (2022), 1–25.

[89] John C Ward, Michael G Dow, T Saunders, S Halls, Kathy F Penner, Kristina A Musante, and N Sachs-Ericcson. 2006. Children's functional assessment rating scale.

[90] Melissa Wells and Kimberly J. Mitchell. 2014. Patterns of Internet Use and Risk of Online Victimization for Youth With and Without Disabilities. *The Journal of Special Education* 48, 3 (2014), 204–213. https://doi.org/10.1177/0022466913479141 arXiv:https://doi.org/10.1177/0022466913479141

[91] Pamela Wisniewski, Arup Kumar Ghosh, Heng Xu, Mary Beth Rosson, and John M Carroll. 2017. Parental control vs. teen self-regulation: Is there a middle ground for mobile online safety?. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 51–69.

[92] Pamela Wisniewski, Haiyan Jia, Na Wang, Saijing Zheng, Heng Xu, Mary Beth Rosson, and John M Carroll. 2015. Resilience mitigates the negative effects of adolescent internet addiction and online risk exposure. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 4029–4038.

[93] Pamela J. Wisniewski, Jessica Vitak, and Heidi Hartikainen. 2022. *Privacy in Adolescence.* Springer International Publishing, Cham, 315–336. https://doi.org/10.1007/978-3-030-82786-1_14

[94] Pamela J Wisniewski, Heng Xu, Mary Beth Rosson, and John M Carroll. 2014. Adolescent online safety: the" moral" of the story. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1258–1271.

[95] Pamela J. Wisniewski, Heng Xu, Mary Beth Rosson, Daniel F. Perkins, and John Millar Carroll. 2016. Dear Diary: Teens Reflect on Their Weekly Online Risk Experiences. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016).

[96] Ben Wolford. 2018. What is GDPR, the EU's new data protection law? https://gdpr.eu/what-is-gdpr/ (Accessed on 08/15/2022).

[97] Douglas Zytko, Pamela J. Wisniewski, Shion Guha, Eric P. S. Baumer, and Min Kyung Lee. 2022. Participatory Design of AI Systems: Opportunities and Challenges Across Diverse Users, Relationships, and Application Domains. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 154, 4 pages. https://doi.org/10.1145/3491101.3516506

## A  INTERVIEW QUESTIONS

Below you will find the interview questions used for our semi-scripted interviews. The interviews were conducted as part of a project funded by the National Science Foundation (NSF) I-Corp program. For transparency, the NSF I-Corp program was introduced to participants in this study.

**Interview Warm-up**

(1) Introduce yourself
(2) Introduce the purpose of I-Corps
(3) Ask permission to use for research purposes
(4) Ask permission to record

**Interview Questions**

(1) Please tell me a little about your organization?
(2) What is your title and role?
(3) In what capacity do you work with teens?
(4) Approximately how many teens does your organization work with in a given year. If they are in a local system, as how many teens are engaged in this system within the entire U.S.
(5) In relation to how you engage with teens in your job, what are your key performance metrics for success?
(6) Which of these metrics are you most concerned about?
(7) What role do you think online risks experiences of these youth impact these goals?
(8) How many teens that you work with have experienced significant online risks?
(9) What online risks are the most concerning or considered as the biggest problem?
(10) How do you currently find out if a teen is experiencing these types of online risks?
(11) What are the limitations of this approach?

**Technology Interview**

(1) Briefly introduce the technology idea
(2) What would be some of the benefits of this approach?
(3) What do you see as some of the potential drawbacks?
(4) Who would be a key decision-maker in implementing a risk assessment tool like this?

**Discovery Questions**

(1) How do you currently fund your existing risk assessment tools for youth?
(2) If our proposed idea was a validated risk assessment tool, what do you think the biggest barriers to adoption would be? How could we possibly overcome these barriers?

**Closing Remarks**

(1) Is there any question that I didn't already ask that you think would be useful for me to add to future interviews? If so, what would it be, how would you answer it, and why do you think this question would be valuable to ask?
(2) Could you please refer to me at least three other people who I would benefit interviewing about this project? I really appreciate your time and insights and would love to keep in touch with you as we work on this project.

## B   PARTICIPANT CHARACTERISTICS

Table 4.  Participant characteristics and risks they identified

| Participant ID | Provider Type | Job Responsibilities | State | Identified Risks |
|---|---|---|---|---|
| 1 | Mental Health | Mental health Therapist | Colorado | Internet Dependency, Sexual Risk, Cyberbullying |
| 2 | Juvenile Justice | Prevention and intervention | Colorado | Sexual Risk, Cyberbullying |
| 3 | Child Welfare | Children, youth, and family services | Colorado | Internet Dependency, Cyberbullying |
| 4 | Mental Health | Mental health Therapist | Colorado | Sexual Risk |
| 5 | Education | Public health professor | Colorado | Sexual Risk |
| 6 | Education | Public health professor | Colorado | Internet Dependency |
| 7 | Child Welfare | Youth development services | Colorado | Sexual Risk |
| 8 | Education | School safety resources | Colorado | Cyberbullying |
| 9 | Education | Public safety resources | Colorado | Sexual Risk |
| 10 | Education | Public library administrator | Colorado | Sexual Risk |
| 11 | Juvenile Justice | Probation officer | Colorado | Cyberbullying, Internet Dependency |
| 12 | Juvenile Justice | Probation officer | Colorado | Internet Dependency |
| 13 | Child Welfare | Youth development services | Colorado | Internet Dependency |
| 14 | Education | Health and wellness teacher | Florida | Cyberbullying, Sexual Risk |
| 15 | Child Welfare | Child and family services | Florida | Cyberbullying, Sexual Risk |
| 16 | Juvenile Justice | Government consultant | Florida | Internet Dependency, Cyberbullying |
| 17 | Mental Health | Youth and family therapist | Florida | Sexual Risk |
| 18 | Mental Health | Youth and family therapist | Florida | Sexual Risk, Cyberbullying |
| 19 | Education | Language arts teacher | Florida | Cyberbullying |
| 20 | Mental Health | Mental health therapist | Florida | Internet Dependency |
| 21 | Mental Health | Mental health therapist | Florida | Cyberbullying |
| 22 | Mental Health | Mental health therapist | Florida | Cyberbullying, Internet Dependency |
| 23 | Child Welfare | Youth engagement services | Florida | Sexual Risk |
| 24 | Child Welfare | Youth engagement services | Florida | Sexual Risk |
| 25 | Mental Health | Mental health therapist | Florida | Cyberbullying, Sexual Risk |
| 26 | Juvenile Justice | Retired law enforcer | Florida | Cyberbullying |
| 27 | Juvenile Justice | Retired law enforcer | Kentucky | Sexual Risk |
| 28 | Juvenile Justice | Law enforcer | Florida | Sexual Risk |
| 29 | Juvenile Justice | Law enforcer | Florida | Sexual Risk |
| 30 | Juvenile Justice | Law enforcer | Florida | Sexual Risk |
| 31 | Mental Health | LGBTQ+ youth advocate | Florida | Sexual Risk |
| 32 | Child Welfare | Youth advocate | Missouri | Sexual Risk |
| 33 | Juvenile Justice | Law enforcer | Florida | Sexual Risk |
| 34 | Education | Youth advocacy services | Missouri | Sexual Risk |
| 35 | Juvenile Justice | Retired law enforcer | Kentucky | Sexual Risk, Cyberbullying |
| 36 | Education | Administrator | Kentucky | Sexual Risk |
| 37 | Education | Administrator | Kentucky | Sexual Risk |