



Towards a Safer Digital Future: Exploring Stakeholder Perspectives on Creating a Sustainable Youth Online Safety Community

Sarvech Qadir

Vanderbilt University

Nashville, Tennessee, USA

sarvech.qadir@vanderbilt.edu

Ashwaq Alsoubai

Vanderbilt University

Nashville, Tennessee, USA

ashwaq.alsoubai@vanderbilt.edu

Andy Niser

Vanderbilt University

Nashville, Tennessee, USA

andrew.i.niser@vanderbilt.edu

Jinkyung Katie Park

Vanderbilt University

Nashville, Tennessee, USA

jinkyung.park@vanderbilt.edu

Xavier Caddle

University of Central Florida

Orlando, Florida, USA

xavier.caddle@knights.ucf.edu

Pamela J. Wisniewski

Vanderbilt University

Nashville, Tennessee, USA

pam.wisniewski@vanderbilt.edu

ABSTRACT

In this study, we synthesize insights from secondary stakeholders (i.e., IT professionals, teachers, and entrepreneurs) in youth online safety regarding the use of online community platforms to raise awareness, and their effectiveness in sponsoring community engagement for developing youth online safety solutions. We created an online platform comprised of a youth online risk detection dashboard and stakeholder engagement features. We conducted semi-structured interviews with secondary stakeholders ($n=10$) in youth online safety to gain insights related to the use of 1) an online risk detection tool for youth, and 2) an online community platform. We present findings on the youth-focused risk detection dashboard, its educational integration, and effective consortium-building practices. Findings indicate that stakeholders emphasized the importance of privacy in managing youth's social media data, within online youth communities. They highlighted the potential of such communities in educational settings to boost digital literacy, advocating for enhanced transparency and data protection.

CCS CONCEPTS

- Human-centered computing → Empirical studies in HCI.

KEYWORDS

Youth online community, stakeholder perspectives, social media

ACM Reference Format:

Sarvech Qadir, Andy Niser, Xavier Caddle, Ashwaq Alsoubai, Jinkyung Katie Park, and Pamela J. Wisniewski. 2024. Towards a Safer Digital Future: Exploring Stakeholder Perspectives on Creating a Sustainable Youth Online Safety Community. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3613905.3651019>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0331-7/24/05

<https://doi.org/10.1145/3613905.3651019>

1 MOTIVATION AND BACKGROUND

Social media platforms are increasingly popular among children and, particularly, adolescents and young adults globally [36]. These platforms offer novel avenues for self-expression [19], educational growth [2], online communication [9], and the handling of privacy and personal boundaries [47, 52]. The digital platforms facilitate communication and socialization, connecting youth with peers and mentors globally, thus fostering exchange and broadening cultural understanding [11]. At the same time, the rapid evolution of digital technologies and online behaviors poses a continuous challenge in safeguarding young internet users [31, 33]. The growing integration of the internet into daily life, particularly for young individuals, brings forth a spectrum of risks, including cyberbullying, privacy breaches, and exposure to harmful content. [3, 6, 34, 38, 39, 41]. Existing youth online safety solutions have been focused on family-based and restrictive solutions such as parental control apps that allow parents to monitor and control their child's online activity [21]. The use of restriction and monitoring by parents may shield teens from online risks, but at the cost of trust between parents and teens, and positive family value as a whole [37, 43, 54]. Moreover, such parent-centered solutions could leave more vulnerable teens behind as they often do not have parents to actively engage with them to ensure their online safety [7, 8]. Therefore, there is a need for a proactive and expert-driven approach to promoting online safety [12], highlighting the broader view of multi-stakeholders in developing effective and responsive solutions [13, 55]. This calls for a community-based approach that incorporates the perspectives and experiences of a variety of stakeholders including educators, parents, policymakers, and youth themselves [4, 15, 41]. Thus, we aim to build a community-driven online platform to promote youth online safety. Central to our research is an exploration of perspectives of key stakeholders in the domain of youth online safety (e.g., educators, developers, IT professionals, etc) on the potential and effectiveness of a community-based platform as a youth-focused safety initiative.

Previous literature on open-source collaboration mostly focused on three areas: motivation and barriers to open-source participation [40, 46, 51], community norms within these environments [16, 26, 45], and risks associated with open-source collaboration

[22, 49]. Sholler et al. conducted open-source project participation revealing that intrinsic motivators like the desire to learn and contribute to the community, alongside extrinsic factors such as reputation enhancement and meeting external needs, significantly drive youth engagement [48]. In addition, a study by Steinmacher et al. conducted qualitative data analysis on student contributors to projects and observed that providing clear guidance and orientation for newcomers facilitates rapid youth engagement [50]. Liu et al. [32] conducted an in-depth analysis of open-source communities, focusing on their organizational structures, decision-making processes, and the diversity inherent within these groups. The study offers valuable insights into the dynamics that drive the functioning and evolution of open-source projects [32]. Additionally, content analysis conducted by Harris et al. and Ryoo et al. addressed challenges such as security vulnerabilities, sustainability, contributor burnout, and legal complexities, underscoring the need for vigilant management and adherence to standards in sensitive projects like youth online safety [28, 44]. Prior research provides a foundation for understanding the barriers of open-source collaboration in this context [23]. In the case of an online community, the engagement and feedback of online safety experts play a pivotal role in shaping a platform contextually relevant for youth and indicate a need for more collaborative approaches in the realm of youth online safety [29, 35, 53]. Previous studies have underscored the importance of a collaborative approach that integrates insights from end users, industry experts, healthcare practitioners, and academic researchers to address issues of online safety for young people and the development of youth-oriented digital solutions [13]. As a result, there have been significant improvements in tackling prominent issues in the field of youth online safety [42]. Moreover, the success of any online safety initiative hinges on its ability to resonate with its target audience [18]. Therefore, youth rely on engaging and user-friendly online safety tools that are often ineffective or fail to meet their unique requirements [20]. We address this need by providing a comprehensive analysis of expert suggestions for the online community platform, thereby contributing valuable insights to the ongoing discourse on effective online safety practices.

By examining the various dimensions of an online community platform, from its design to its community engagement features, the study provides a nuanced understanding of its potential and efficacy in promoting a safer online environment for young users. We answer the following research questions:

- **RQ1:** *What are online safety stakeholder expert opinions on a youth-centric risk detection dashboard for youth online safety?*
- **RQ2:** *What are the key capabilities and best practices for building an online consortium to support youth online safety?*

To address these questions, we conducted semi-structured interviews with ($n = 10$) participants to understand their perspectives regarding the youth online community and its features. Then, we conducted a grounded thematic analysis. For RQ1, we found that the platform stands out as an effective instrument in educational settings to enhance digital literacy, with a strong focus on ensuring transparency and safeguarding data. For RQ2, stakeholders identified key capabilities that platforms could incorporate to enhance community building and user engagement and made suggestions regarding features that would be valuable to integrate as part of

the community. The purpose of this study was to solicit feedback regarding an online risk detection dashboard and youth online safety community from the perspective of stakeholders who work in this space. This study contributes empirical research findings, presenting valuable data from studies and research that shed light on young users' behaviors, preferences, and the challenges they face in digital environments. Also, it introduces design innovations in user interface and interaction techniques, enhancing the online experience for youth and bolstering their safety.

2 METHODS

We conducted 10 semi-structured interviews with design probes via Zoom to interact with the participants and learn their insights on the features and mechanisms required of youth online safety communities. The semi-structured interview design was utilized to allow for flexibility in discussion and design probes as part of the platform allowed participants to interact with the prototype and share their reflections [24]. Each interview was split into two sections: 1) dashboard demonstration and 2) community prototype usage. Each section included design probes. Section 1 elicited feedback from participants regarding the potential use of the risk detection dashboard by youth, any barriers that might prevent youth from using it, and features that might be useful to youth. In Section 2, participants used the online community prototype to execute 4 tasks inclusive of creating a user profile, searching for current members they might want to reach out to, creating forum posts on the platform, and proposing a project to the community. Figure 1 provides an overview of the session flow with each participant. To ensure ethical treatment of the data entrusted to us, we received Institutional Review Board (IRB) approval from the last author's institution. All research team members were also required to complete CITI training [27] as well. Each session lasted for 60 minutes and participants were offered \$20 Amazon gift card compensation for participation. The following sections outline the participant recruitment process and the semi-structured interview in more detail.

2.1 Participant Recruitment

Participants were mainly recruited from existing contacts and were contacted via email to schedule a session based on their availability. They were recruited based on specific inclusion criteria to ensure that the study gathered information from a diverse and relevant subset of individuals. The inclusion criteria for recruiting secondary stakeholders were adapted and revised from a study by Caddle et al. [13]. These criteria required participants to be 18 years old or older, be an English speaker, and fall into at least one of the following categories: 1) Parents of teenagers, to capture the perspectives of those who are directly responsible for managing the online safety of youth. 2) Experts or stakeholders in promoting youth online safety, to draw on professional and academic expertise in the field. 3) Individuals who work or have worked in youth online safety within the last five years, including practical, field-based insights into the current challenges and solutions in online safety.

2.2 Interview Procedure

In the first section of the interviews, we asked participants to watch a recorded video demonstration of an online safety dashboard [5]. The dashboard (Figure 2) is an online risk assessment tool that allows youth ages 13-21 to upload social media data files from Instagram and Twitter to be assessed for instances of risk such as cyberbullying and exposure to sexually explicit data. This web-based dashboard allows youths to assess AI-identified risks within their online interactions, specifically in private conversations. The evaluation provided by youths will be used to improve the quality of machine learning models. In addition, the dashboard is not yet a complete system, but a prototype that is shown to participants in the study. After watching the demonstration, we asked participants questions about the dashboard regarding its utility, applicability, and educational impact.

In the second section of the interviews, video tutorials were shown to participants to assist them in performing tasks on a prototype online community platform. These tasks required participants to interact with the initial features ¹ of the prototype of an online community. The features participants explored included member profile creation, member search, forum posting, project proposals, news dissemination, and comment on all of the preceding as shown in Table 1. These situations were designed to prompt discussion about the platform and reveal deeper insights into the participants' beliefs and attitudes toward what features need to be present in community-oriented youth online safety platforms to make them useful to the youth online safety community.

- First, participants were walked through the steps to create their contributor profile (Figure 3). Questions were then asked about the profile creation process, important details they would like to capture and showcase, and information that they would not be willing to share on their profile.
- Next, participants used the platform's member search feature to search for a community member that they might be interested in contacting to work on a future project. They also sent a message to a user that interested them. We then asked the participants when they would be seeking out new contacts on a platform, what fields would be important to identify them, and what suggestions they had to upgrade the current online community platform.
- Thirdly, participants proposed a project to the community (Figure 4). They also learned how to find other projects, see their history, and engage in discussion. After that, participants were asked about their expectations about collaborating on projects, what types of details they would like to gather from users for a project, any barriers they believe would prevent users from proposing projects, and any rules they think should be added to ensure no one can steal ideas or projects from others.
- Finally, participants learned how to participate in the forum and created a post on the platform. Participants were then asked what topics they would like or not like to see on the forum and what would encourage or discourage them from engaging in community discussions.

¹initial features were identified from previously published literature but citation was removed for anonymity

To conclude the interview, participants were asked questions about their general thoughts on the platform as a whole. These focused on how we could build an online community, what modifications and additions they would like to see, and what features they valued most among other topics.

2.3 Data Analysis

All sessions were video and audio recorded, and the recordings were fully transcribed by the researchers. Our preliminary thematic qualitative analysis [10] involved reviewing the recorded sessions and grouping recurring insights to identify major themes about youth-centric risk detection dashboards, community engagement in youth online safety, and best practices for developing an online consortium for youth online safety. This method allowed us to identify key patterns within data, thus providing an intricate understanding of the various perceptions and attitudes of participants towards the community platform. The major themes we identified from the interviews included the utility and barriers of the dashboard, community initiative features, project creation, post creation, and the benefits of the community platform.

3 RESULTS

We had a total of 10 stakeholders out of which three worked with law enforcement (courts, lawyers, and police). Five participants had prior experience working with youth as a school social worker, in youth empowerment, and mentorship. Two participants worked directly with teens in educational settings. Below, we present our findings related to online safety stakeholders' expert opinions on 1) the integration and educational impact of a youth-centric risk detection dashboard, 2) the community features and accessibility options within the platform, and 3) their suggestions for best consortium building practices.

3.1 Stakeholders highlighted the significance of incorporating the online risk detection dashboard in educational and wellness programs (RQ1)

The results showed that experts ($n = 4$) expressed strong support for integrating an online risk detection dashboard into school health and wellness programs and professional development training for staff. Experts ($n = 3$) viewed dashboards as a valuable tool for educating youth about recognizing and handling bullying and facilitating safe communication between stakeholders and members. A recurring theme was the **importance of educational impact**, with the belief that an online risk detection dashboard could play a crucial role in educating teenagers about the repercussions of online interactions, particularly regarding bullying. For instance, one participant described the risk detection dashboard as “*useful to check the veracity of websites (e.g. scam detection)*” describing it as a “**critical tool for identifying risks**” in youths’ online interactions (P8). At the same time, a few participants ($n = 2$) expressed concerns about the usability of the platform. For example, a participant expressed skepticism about whether teenagers would actively use such a platform, citing,

"Would [the youth] use something like this or would this be something that they would find accessible or be interested in? And I don't know... I'm trying to envision any of them taking the time to really upload that and assess those kinds of risks." (P1, educator)

Also, half of the experts ($n = 5$) noted potential resistance from teenagers to sharing personal data, such as Instagram activity, due to **privacy concerns and security of user data**. Participants identified the problems with how personal data is collected, stored, and used by online risk detection platforms. Stakeholders also raised concerns such as privacy factors and the security of user data due to its sensitive nature. Therefore, participants ($n = 2$) expressed that incorporating website authenticity checks i.e. SSL/TLS certificates and content verification could significantly strengthen the platform's security against suspicious activities on the platform, especially in educational and wellness settings where security is a key. The participants' emphasis on SSL/TLS certificates can be understood as a call for robust security measures as a foundational step in building a secure platform that users can trust.

3.2 Stakeholders highlighted key practices for a successful online youth safety consortium, focusing on features that enhance user engagement, community interaction, and platform security (RQ2)

In building the online consortium supporting youth online safety, experts ($n = 7$) emphasized the necessity of incorporating features to facilitate **user engagement** and positively impact user retention and interaction. For example, experts ($n = 5$) suggested that forums should be well-formatted and filtered by major topic areas and educational levels to ensure that users could easily access relevant and valuable content. Moreover, some participants ($n = 4$) suggested including a real-time message monitoring feature by providing email notifications when a user receives a message from another community member on the platform. For instance, one participant mentioned,

"An e-mail notification like hey, someone messaged you click [on the link] to go [to the platform]." (P6, law enforcement)

Furthermore, some stakeholders ($n = 3$) stressed the importance of adopting Google sign-up credentials for account creation which can simplify the overall onboarding process of the platform. Finally, participants ($n = 3$) also emphasized the importance of the platform's visual appeal to maintain user engagement. In addition, stakeholders recommended integrating features to **encourage community coordination**. In discussing proposing, viewing, and commenting on projects, participants ($n = 4$) described the inclusion of a voting mechanism as beneficial to allow community members to attach priority to proposed projects on the platform. Moreover, participants ($n = 4$) applauded the feature of adding a user bio to the profile which enables community members to connect based on individual preferences in user bios.

Some participants ($n = 3$) suggested features in which community members can add tags or keywords to their profiles based on their work experiences. For instance, one participant who has previously

worked with youth, identified the significance as "*allowing [users] to put tags on [the platform] would be helpful.*" (P1, educational setting). This can further be coupled with robust search tools or filtering techniques to search for users' profiles and projects based on the individual preferences the members would like to connect. Other stakeholders ($n = 2$) suggested for project and post-creation tools on the platform. These tools include prompts for project timelines, role specifications, expected contributions to the community, and alignment with broader mission statements. Lastly, when engaging with youth, stakeholders ($n = 2$) emphasized the importance of making youth feel like experts in the field, which will ultimately lead them to think more deeply. One of the participants further clarified:

"Acknowledging the fact that they are experts in the field. Maybe like, you know, make them more brainstorm about different things about that they know." (P7, youth mentorship)

This emphasizes the importance of reminding youth of the significant contributions that will eventually make them more confident in sharing their feedback. Thirdly, the majority of participants ($n = 6$) emphasized embedding features to **enhance platform security and credibility** through strict adherence to data protection laws, clear consent mechanisms, and a well-defined code of conduct outlining the security expectations of an online community. Finally, a few participants ($n = 2$) also suggested deploying monitoring checks during the sign-up and profile creation phases to effectively minimize unwanted account activities on the platform and keep the community safe. As such, participants shared diverse ideas to improve overall user experiences on the online consortium to promote active engagement of the online safety stakeholders.

4 DISCUSSION

4.1 Educational impact of online community for youth

Our results showed that platforms focused on youth online communities can be potentially integrated into school curriculums as components of health and wellness education, positioning them as proactive instruments against cyberbullying, and promoting digital literacy. Furthermore, they could serve as invaluable resources for educators and parents, providing them with strategies and insights to better support young individuals in online safety practices. For example, integrating an administrator view for adult mentors (e.g., parents, educators, health professionals, etc) could complement current youth online safety approaches allowing for more support. The emphasis on controlled stakeholder-to-member communication within youth online safety platforms could also foster a monitored and educational dialogue space, ensuring a more secure experience. We propose the development of collaborative monitoring systems designed to improve communication between stakeholders and youth. Specifically, we suggest the development of applications that enable both parties to jointly oversee online interactions, fostering a safer and more transparent digital environment [14]. Alternatively, we also suggest designing prototypes using co-design approaches that foster mutual understanding and shared responsibility in managing youth online experiences [1].

4.2 Enhancing Privacy and Communication Channels for Youth Engagement

Beyond the primary evaluative focus of our study, additional findings indicate a significant stakeholder demand for platform transparency, especially regarding data utilization and protection practices. This is paralleled by calls for platforms to continuously evolve to keep pace with the shifting digital landscape, ensuring sustained relevance and efficacy for its user base. Such adaptability is critical, given the rapid innovation and change inherent in digital technologies and youth culture [17]. The results showed significant concerns around privacy and data sharing among adolescents. To ensure a balance between privacy and protection, particularly when adopting socio-technical solutions, we advocate for teen privacy concerns to be integrated into a broader policy agenda. The agenda needs to encompass protocols for data handling ensuring teens and their guardians understand how their data is used and providing mechanisms for teens to have a say in how their personal information is managed. In addition, the experts underscored the importance of developing clear communication channels tailored to diverse age groups to foster effective engagement with the platform's features. We propose the development of customized interfaces specifically designed for different age groups. These tailored interfaces would cater to the unique cognitive abilities, interests, and technology proficiency levels of developmentally different age brackets. These initial insights are significant in forming the roadmap for the subsequent development of guidelines for the youth online community.

4.3 Extending the Study's Impact: Implications for Youth Online Community Platforms

In the broader context, the implications of this study extend beyond the specific focus on youth online safety platforms. Prior literature suggested a “multi-stakeholder approach” to mitigate youth online risk [30]. Our results strengthen this argument by providing empirical evidence that an online consortium can be an effective way to make this happen. In addition, our results extend the literature by providing insights into ways in which we can support multi-stakeholders in the online consortium. For instance, principles such as adapting interfaces to cater to specific age groups, ensuring platform transparency in data utilization, and the importance of adapting to the digital landscape are applicable beyond our platform. Particularly, the importance of digital privacy and data protection in online communities is emphasized in prior literature [25]. Therefore, our insights can be effectively translated into a variety of applications such as educational software tailored for different student age groups, developing social media apps that require age-appropriate content filtering and privacy settings, and e-commerce platforms that modify their interface for varying users. These principles can enhance user engagement, trust, and effectiveness across various digital platforms, whether they are educational tools, social networks, or health-related resources. Moreover, the study's emphasis on stakeholder engagement and feedback can serve as a model for inclusive and user-centered design in technology development. This approach is crucial in creating digital solutions that are not only technically sound but also resonate with and meet the evolving needs of diverse users. Our findings thus

contribute valuable perspectives to the broader discourse on responsible and responsive digital platform design and management.

5 LIMITATIONS AND FUTURE WORK

The total number of interviews conducted ($n=10$) is small, hence, our results are not representative of all stakeholders. Future efforts involve more proactive efforts to engage more stakeholders from varying backgrounds using different approaches i.e. snowball sampling and flyers. Moreover, further directions for this research include implementing design modifications informed by stakeholder feedback to improve user interface and privacy measures. Additionally, efforts to engage a broader spectrum of stakeholders will be instrumental in diversifying perspectives and ensuring the platform addresses the needs of a wider user community effectively.

6 CONCLUSION

Our study highlights the critical role of expert-driven youth online safety platforms in empowering both youth. Guidelines such as capability, commitment to user-centered design, and ongoing engagement with stakeholders stand out as core tenets to adhere to towards the support of platform longevity and meeting core mission directives. These elements are essential in ensuring that the platform not only addresses current online safety concerns but is also equipped to evolve with the changing digital landscape. The continued collaboration with online safety experts is crucial in steering these platforms toward intersecting the needs of youth with necessary modes of protection.

ACKNOWLEDGMENTS

This research was supported by the U.S. National Science Foundation under grants IIP-2329976, IIS-2333207, and by the William T. Grant Foundation grant 187941. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of our sponsors.

REFERENCES

- [1] Zainab Agha, Karla Badillo-Urquiola, and Pamela J Wisniewski. 2023. "Strike at the Root": Co-designing Real-Time Social Media Interventions for Adolescent Online Risk Prevention. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–32.
- [2] June Ahn. 2011. The effect of social network sites on adolescents' social and academic development: Current theories and controversies. *Journal of the American Society for Information Science and Technology* 62, 8 (2011), 1435–1445.
- [3] Ashwaq Alsoubai. 2023. A Human-Centered Approach to Improving Adolescent Real-Time Online Risk Detection Algorithms. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI EA '23). Association for Computing Machinery, New York, NY, USA, Article 480, 5 pages. <https://doi.org/10.1145/3544549.3577045>
- [4] Ashwaq Alsoubai, Xavier V Caddle, Ryan Doherty, Alexandra Taylor Koehler, Estefania Sanchez, Munmun De Choudhury, and Pamela J Wisniewski. 2022. MOSafely, Is that Sus? A Youth-Centric Online Risk Assessment Dashboard. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing*. 197–200.
- [5] Ashwaq Alsoubai, Xavier V Caddle, Ryan Doherty, Alexandra Taylor Koehler, Estefania Sanchez, Munmun De Choudhury, and Pamela J Wisniewski. 2022. MOSafely, Is that Sus? A Youth-Centric Online Risk Assessment Dashboard. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing*. 197–200.
- [6] Ashwaq Alsoubai, Jihye Song, Afsaneh Razi, Nurun Naher, Munmun De Choudhury, and Pamela J Wisniewski. 2022. From 'Friends with Benefits' to 'Sexortion': A Nuanced Investigation of Adolescents' Online Sexual Risk Experiences. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–32.

- [7] 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.
- [8] Karla A Badillo-Urquiola, Arup Kumar Ghosh, and Pamela Wisniewski. 2017. Understanding the unique online challenges faced by teens in the foster care system. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 139–142.
- [9] Amal Mohammed Nabil Abdul Azim Badr, Tarek Ismail Mohamed, Nassereldin Abdel Qadir Osman, and Alexey Mikhaylov. 2022. A review of social media website users' interaction paths with governmental accounts during the COVID-19 pandemic. In *Informatics*, Vol. 9. MDPI, 50.
- [10] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis*. American Psychological Association.
- [11] Catherine F Brooks and Margaret J Pitts. 2016. Communication and identity management in a globally-connected classroom: An online international and intercultural learning experience. *Journal of international and intercultural Communication* 9, 1 (2016), 52–68.
- [12] Xavier Caddle, Ashwaq Alsoubai, Afsaneh Razi, Seunghyun Kim, Shiza Ali, Gianluca Stringhini, Munmun De Choudhury, and Pamela Wisniewski. 2021. Instagram data donation: A case for partnering with social media platforms to protect adolescents online. In *ACM Conference on Human Factors in Computing Systems (CHI 2021)/Social Media as a Design and Research Site in HCI: Mapping Out Opportunities and Envisioning Future Uses Workshop*.
- [13] Xavier Caddle, Jinkyung Park, and Pamela J Wisniewski. [n. d.]. A Stakeholders' Analysis of the Sociotechnical Approaches for Protecting Youth Online. ([n. d.]).
- [14] Xavier V Caddle, Nurun Naher, Zachary P Miller, Karla Badillo-Urquiola, and Pamela J Wisniewski. 2023. Duty to Respond: The Challenges Social Service Providers Face When Charged With Keeping Youth Safe Online. *Proceedings of the ACM on Human-Computer Interaction* 7, GROUP (2023), 1–35.
- [15] 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*. 315–318.
- [16] Amanda Casari, Julia Ferraioli, and Juniper Lovato. 2023. Beyond the repository: Best practices for open source ecosystems researchers. *Queue* 21, 2 (2023), 14–34.
- [17] Calvin ML Chan, Say Yen Teoh, Adrian Yeow, and Gary Pan. 2019. Agility in responding to disruptive digital innovation: Case study of an SME. *Information Systems Journal* 29, 2 (2019), 436–455.
- [18] Emily Chen and Margarita DiVall. 2018. Social media as an engagement tool for schools and colleges of pharmacy. *American journal of pharmaceutical education* 82, 4 (2018), 6562.
- [19] Tae Rang Choi and Yongjun Sung. 2018. Instagram versus Snapchat: Self-expression and privacy concern on social media. *Telematics and informatics* 35, 8 (2018), 2289–2298.
- [20] Susan Edwards, Andrea Nolan, Michael Henderson, Ana Mantilla, Lydia Plowman, and Helen Skouteris. 2018. Young children's everyday concepts of the internet: A platform for cyber-safety education in the early years. *British journal of educational technology* 49, 1 (2018), 45–55.
- [21] Lee B Erickson, Pamela Wisniewski, Heng Xu, John M Carroll, Mary Beth Rosson, and Daniel F Perkins. 2016. The boundaries between: Parental involvement in a teen's online world. *Journal of the Association for Information Science and Technology* 67, 6 (2016), 1384–1403.
- [22] Jim Euchner. 2013. The uses and risks of open innovation. *Research-Technology Management* 56, 3 (2013), 49–54.
- [23] Parisa Haim Faridian and Donald O Neubaum. 2021. Ambidexterity in the age of asset sharing: Development of dynamic capabilities in open source ecosystems. *Technovation* 99 (2021), 102125.
- [24] Bill Gaver, Tony Dunne, and Elena Pacenti. 1999. Design: cultural probes. *interactions* 6, 1 (1999), 21–29.
- [25] Elias Grünewald, Johannes M Halkenhäußer, Nicola Leschke, Johanna Washington, Cristina Paupini, and Frank Pallas. 2023. Enabling versatile privacy interfaces using machine-readable transparency information. In *Privacy Symposium: Data Protection Law International Convergence and Compliance with Innovative Technologies*. Springer, 119–137.
- [26] Kholekile L Gwebu and Jing Wang. 2011. Adoption of Open Source Software: The role of social identification. *Decision support systems* 51, 1 (2011), 220–229.
- [27] Kristie B Hadden, Latrina Prince, Laura James, Jennifer Holland, and Christopher R Trudeau. 2018. Readability of human subjects training materials for research. *Journal of Empirical Research on Human Research Ethics* 13, 1 (2018), 95–100.
- [28] Lauren E Harris and Jerry A Jacobs. 2023. Emerging Ideas. Digital parenting advice: Online guidance regarding children's use of the Internet and social media. *Family Relations* 72, 5 (2023), 2551–2568.
- [29] Brian W Head. 2011. Why not ask them? Mapping and promoting youth participation. *Children and Youth services review* 33, 4 (2011), 541–547.
- [30] Yujin Jang and Bomin Ko. 2023. Online Safety for Children and Youth under the 4Cs Framework—A Focus on Digital Policies in Australia, Canada, and the UK. *Children* 10, 8 (2023), 1415.
- [31] Anastasia Kozyreva, Stephan Lewandowsky, and Ralph Hertwig. 2020. Citizens versus the internet: Confronting digital challenges with cognitive tools. *Psychological Science in the Public Interest* 21, 3 (2020), 103–156.
- [32] Manlu Liu, Clyde Eirikur Hull, and Yu-Ting Caisy Hung. 2017. Starting open source collaborative innovation: the antecedents of network formation in community source. *Information Systems Journal* 27, 5 (2017), 643–670.
- [33] Sonia Livingstone and Peter K Smith. 2014. Annual research review: Harms experienced by child users of online and mobile technologies: The nature, prevalence and management of sexual and aggressive risks in the digital age. *Journal of child psychology and psychiatry* 55, 6 (2014), 635–654.
- [34] Sana Maqsood and Sonia Chiasson. 2021. "They think it's totally fine to talk to somebody on the internet they don't know": Teachers' perceptions and mitigation strategies of tweens' online risks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [35] Lance T McCready and Geoffrey B Soloway. 2010. Teachers' perceptions of challenging student behaviours in model inner city schools. *Emotional and Behavioural Difficulties* 15, 2 (2010), 111–123.
- [36] Emmanuel K Ngwainmbi. 2019. Social media use among the youth and working class: Conditions for remediating globalization and cultural space. *Media in the global context: Applications and interventions* (2019), 49–93.
- [37] Jinkyung Park, Mamtaj Akter, Naima Samreen Ali, Zainab Agha, Ashwaq Alsoubai, and Pamela Wisniewski. 2023. Towards Resilience and Autonomy-Based Approaches for Adolescents Online Safety. Available at SSRN 4608406 (2023).
- [38] Jinkyung Park, Joshua Gracie, Ashwaq Alsoubai, Gianluca Stringhini, Vivek Singh, and Pamela Wisniewski. 2023. Towards automated detection of risky images shared by youth on social media. In *Companion Proceedings of the ACM Web Conference 2023*. 1348–1357.
- [39] Jinkyung Park, Vivek Singh, and Pamela Wisniewski. 2023. Supporting Youth Mental and Sexual Health Information Seeking in the Era of Artificial Intelligence (AI) based Conversational Agents: Current Landscape and Future Directions. Available at SSRN 4601555 (2023).
- [40] Huiyan Sophie Qiu, Alexander Nolte, Anita Brown, Alexander Serebrenik, and Bogdan Vasilescu. 2019. Going farther together: The impact of social capital on sustained participation in open source. In *2019 ieee/acm 41st international conference on software engineering (icse)*. IEEE, 688–699.
- [41] Afsaneh Razi, Ashwaq AlSoubai, Seunghyun Kim, Shiza Ali, Gianluca Stringhini, Munmun De Choudhury, and Pamela J Wisniewski. 2023. Sliding into My DMs: Detecting Uncomfortable or Unsafe Sexual Risk Experiences within Instagram Direct Messages Grounded in the Perspective of Youth. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–29.
- [42] Michael D Richardson, Pamela A Lemoine, Walter E Stephens, and Robert E Waller. 2020. Planning for Cyber Security in Schools: The Human Factor. *Educational Planning* 27, 2 (2020), 23–39.
- [43] 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. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–23.
- [44] Jungwoo Ryoo, Bryan Malone, Phillip A Laplante, and Priya Anand. 2015. The use of security tactics in open source software projects. *IEEE Transactions on Reliability* 65, 3 (2015), 1195–1204.
- [45] Maha Shaikh and Emmanuelle Vaast. 2016. Folding and unfolding: Balancing openness and transparency in open source communities. *Information Systems Research* 27, 4 (2016), 813–833.
- [46] Bianca Shibuya and Tetsuo Tamai. 2009. Understanding the process of participating in open source communities. In *2009 ICSE Workshop on Emerging Trends in Free/Libre/Open Source Software Research and Development*. 1–6.
- [47] Dong-Hee Shin. 2010. The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with computers* 22, 5 (2010), 428–438.
- [48] Dan Sholler, Igor Steinmacher, Denae Ford, Mara Averick, Mike Hoye, and Greg Wilson. 2019. Ten simple rules for helping newcomers become contributors to open projects. *PLoS computational biology* 15, 9 (2019), e1007296.
- [49] Mario Silic and Andrea Back. 2016. The influence of risk factors in decision-making process for open source software adoption. *International Journal of Information Technology & Decision Making* 15, 01 (2016), 151–185.
- [50] Igor Steinmacher, Tayana Conte, Marco Aurélio Gerosa, and David Redmiles. 2015. Social barriers faced by newcomers placing their first contribution in open source software projects. In *Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing*. 1379–1392.
- [51] Igor Steinmacher, Tayana Uchoa Conte, Christoph Treude, and Marco Aurélio Gerosa. 2016. Overcoming open source project entry barriers with a portal for newcomers. In *Proceedings of the 38th International Conference on Software Engineering*. 273–284.
- [52] Ken Strutin. 2011. Social media and the vanishing points of ethical and constitutional boundaries. *Pace L. Rev.* 31 (2011), 228.

- [53] Jennifer White. 2016. Reimagining youth suicide prevention. *Critical suicidology: Transforming suicide research and prevention for the 21st century* (2016), 244–263.
- [54] 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.
- [55] Lucy Yardley, Bonnie J Spring, Heleen Riper, Leanne G Morrison, David H Crane, Kristina Curtis, Gina C Merchant, Felix Naughton, and Ann Blandford. 2016. Understanding and promoting effective engagement with digital behavior change interventions. *American journal of preventive medicine* 51, 5 (2016), 833–842.

A APPENDIX

A.1 Session Flow

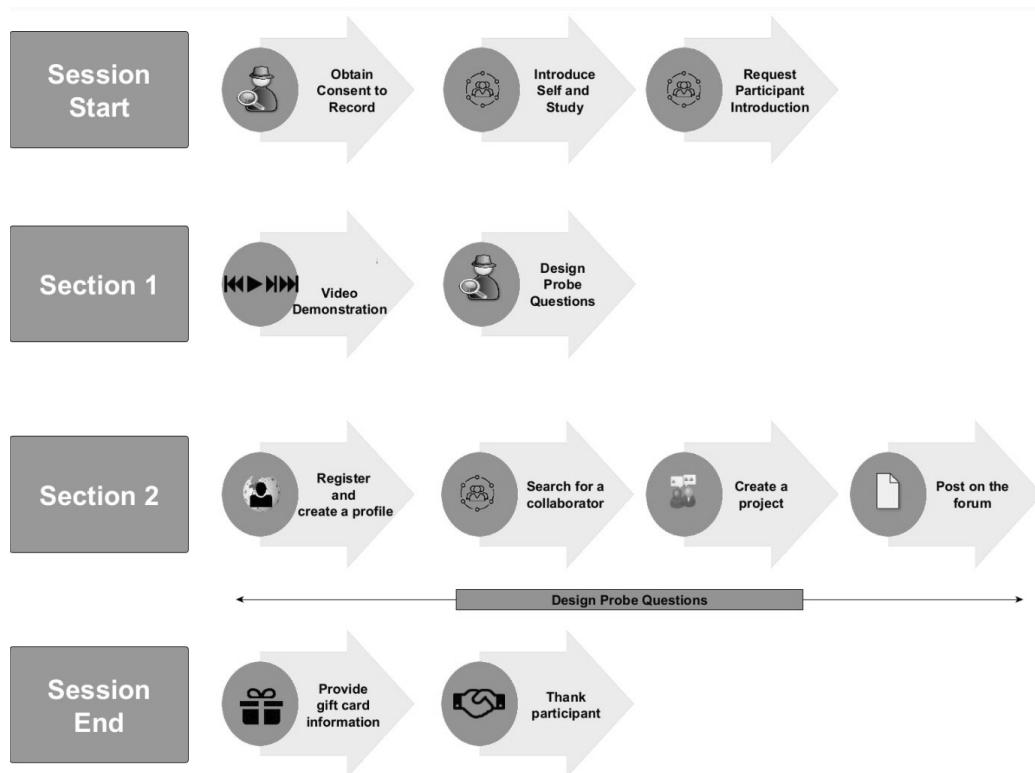


Figure 1: Session Flow

A.2 Online Risk Detection Dashboard

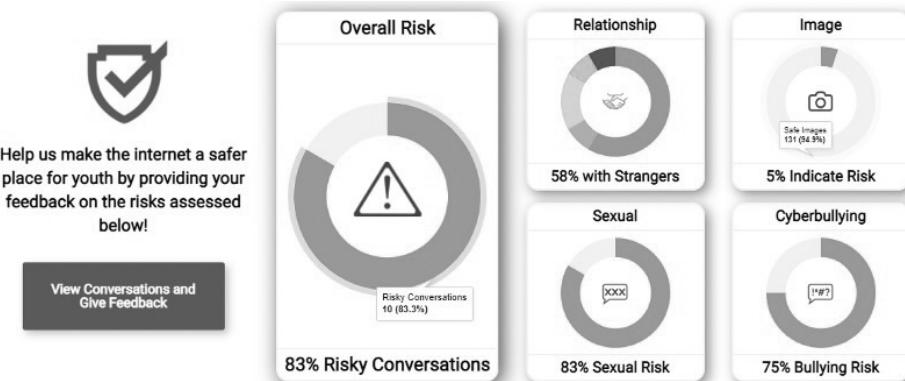


Figure 2: Online Risk Detection Dashboard

A.3 Online Community Feature List

Table 1: Online Community Features

Feature	Supports
News	Disseminating scientifically backed information and updates
Posts	User interaction and engagement
Member Search	User interaction
Projects	Informed design and development of youth online safety solutions
Workshops	Disseminating scientifically backed information

A.4 User Registration page

User Registration

By creating an account on this site you are indicating that you have read, agree to, and will abide with the terms set in the following pages:

- Terms of Use
- Privacy Policy
- Community Guidelines

Already have an account? Click here to login!

Have a Github account? Click here to register using Github!

 User Name

User Name

 Email

Email

 Password

Password

 Password Again

Password Again

 First Name

First Name

 Last Name

Last Name

 Organization/Affiliation

Organization/Affiliation

 Area(s) of Expertise

Nothing selected ▾

I am human



Register User

Figure 3: User Registration

A.5 Project creation page

Create Project

Name:

Title:

GitHub Repo Name:

Content:

A horizontal toolbar for rich text editing, featuring icons for bold (B), italic (I), underline (U), strikethrough (del), font style (-apple-system), font size (A), alignment (left, center, right, justify), lists (bullet, numbered), tables (grid), links (link), images (image), and other document controls (undo, redo, search, help).

Figure 4: Project Creation