Researching the (Potentially) Sexually Explicit Material of a Minor

Arup Kumar Ghosh, Karla A. Badillo-Urquiola, Uzair Tariq, & Pamela J. Wisniewski
University of Central Florida
Orlando, FL USA
arupkumar.ghosh@ucf.edu, kbadillo@ist.ucf.edu, uzairtq@knights.ucf.edu, pamwis@ucf.edu

ABSTRACT

Adolescent online safety has historically been framed as preventing teens from being exposed to various online risks [13], such as sexual solicitations, cyberbullying, and exposure to explicit content [9]. However, in trying to accomplish this goal, researchers must first deeply understand the types of risks teens encounter online. Some of these risks, however, involve very sensitive and possibly illegal artifacts, such as images that could be classified as child pornography, and pose serious ethical issues associated with conducting research.

Author Keywords

Adolescent online safety, sexting, privacy, ethics

ACM Classification Keywords

K.4.1 [Public Policy Issues]: Ethics, Human safety, Privacy

INTRODUCTION

In order to build technologies that can help keep teens safe from serious online risks, researchers must be able to study these risky behaviors to understand and know how to detect and mitigate them. However, a number of ethical challenges arise [10] when dealing with potentially illegal and sexually explicit materials that involve minors. We discuss some of our research thrusts that have necessitated tackling such ethical challenges, as well as note how we have attempted to manage these challenges thus far.

ADOLESCENT ONLINE SAFETY RESEARCH

Qualitative Approaches for Capturing Risk Experiences

In the last authors' prior work as a post-doctoral researcher at Pennsylvania State University, she made considerable efforts to qualitatively understand teens' and parents' perceptions about the online risks teens encountered [12,14]. She and her collaborators did this through a novel, web-based diary study approach where parents and teens separately reported weekly on four types of online risk events (information breaches, sexual solicitations, online harassment, and exposure to explicit content) that teens encountered over the course of a two-month period. However, given the political environment at Penn State after the Sandusky scandal, the university was on high-alert regarding any contact faculty and staff had with minors. For instance, in 2015, Penn State employees were required to complete mandated child abuse training each new academic year and were also given strict requirements on how to engage with minors in the event that they were on campus (e.g., Summer Camps). That being said, the IRB protocol for the web-based diary study took over *two months* to get approved. We were caught off-guard when we had to create protocol for mandated child abuse reporting in the event that imminent risks to minors were identified during the study. This is how Dr. Wisniewski began to realize the ethical challenges associated with her research trajectory.

Algorithmic Approaches for Risk Detection

As Dr. Wisniewski's Computer Science Ph.D. students, we are beginning to work on extensions to her earlier work, involving more automated approaches for online risk detection. For instance, one project involves gathering a data set of teens' public, semi-public, and private social media activities from Instagram, Facebook, and other social media sites through a third-party parental monitoring app. This will involve a user study of parents and teens using the third-party parental monitoring app as well as the data scrape of the social media content. The goal of the data collection efforts is to create and analyze a representative data set of teens' social media behaviors and to identify appropriate risk thresholds and key risk dimensions unique to this population. This type of contextual and in-depth knowledge of adolescent risk behaviors is needed before developing advanced and automated algorithms for risk detection. Our analysis will help future researchers who prefer more automated approaches to better understand the nuance in risk dimensions and thresholding when it comes to risky social media content and teens.

Computational Computer Scientists have created algorithmic solutions for automatically detecting and blocking inappropriate content, such as sexually explicit imagery and forms of malicious text (e.g., cyberbullying) from various forms of media [1,2,6,11,15,16]. While they achieve fairly high accuracy rates in terms of risk detection, they are constrained because the data sets being used are not generalizable to teens. They often resort to using convenient data sets that are scraped from publicly available social media and are not representative of teen users. Representative data sets are needed for computational algorithms to detect key attributes of data in order to create intelligent and automated systems that detect online risks [17]. Further, establishing a “ground truth” [8] for what signifies online risk in order to determine the accuracy of these algorithms has not included the perspectives of either parents or teens. Risk is a normative and highly subjective construct that, when operationalized quantifiably, has considerable impact on safety outcomes [7]. For instance, does it matter if the teen perpetrated the risk (e.g., generated the risky content) or was an innocent or a victim (e.g., unintentionally consumed content posted by a “friend”)?

These types of questions must be addressed prior to the automated detection of online risks, which is the goal of our research. However, creating such a data set, while highly valuable, is also highly sensitive as the social media content of teens may contain sexually explicit or otherwise questionable material involving minors. Distribution of sexually explicit materials of a minor is a crime; therefore, we must be very careful how we manage this data once it has been collected. For instance, National Science Foundation (NSF) grant proposals require a detailed Data Management Plan that encourages public sharing of data in order to advance science. In this case, such data *must not* be shared publicly and unique protocols for sharing the data among research scientists will need to be carefully devised.

Another research trajectory we have begun is a proof-of-concept for developing a low-powered sensor that can be mounted on a mobile phone camera lens to detect nudity *prior to* digital capture. A complication of developing risk detection solutions for mobile devices is that they either need to be lightweight enough to be handled on the client-side device or sent to a centralized server for more advanced processing. Does sending potentially illegal and/or explicit sexual materials of a minor (that had otherwise not been shared beyond that mobile device) to a centralized server constitute “distribution,” and thus, warrant a criminal charge? If this data is stored on servers maintained by a university system and is breached, what could that mean for the university? These are some of the questions we embark on as we pursue this research.

A final thought about the ethical challenges of conducting adolescent online safety research is *how* we handle the adolescent risk behaviors once they are detected? A teen sexting with his or her romantic partner or one who transmits a nude photo beyond an intimate relationship could be charged with the distribution of child pornography and listed in the National Sex Offender Public Website extending well into his or her adult life. Thus, would such risk detection techniques actually be *harmful* to teens by facilitating legal enforcement through open-source intelligence (e.g., surveillance) of such behaviors? Alternatively, is it right to use technology as a proxy for parenting or for helping parents invade their teens’ personal privacy for the sake of keeping them safe? These are only some of the ethical questions we have encountered along the way.

TACKLING THESE CHALLENGES

While we do *not* pretend to have answers for all of the challenges discussed above, here are some of the ways we have attempted to mitigate the ethical risks related to our research, as well as some of our outstanding questions.

- ***Informed Assent:*** In our IRB protocols, we have chosen to obtain both parental consent (which is required) and teen assent to participant in our studies. Although teen assent is not required, we believe that this is a way to give teens a sense of increased agency in the research.
- ***Mandated Child Abuse Reporting:*** We have created detailed protocols given our roles as mandated child abuse reporters. This includes what to do in the event that child abuse or imminent risk to a minor has been detected and disclosing this protocol to parents and teens within the informed consent documents. Also, we have had to make a clear distinction between abuse and illegal activities, as we are not necessarily mandated to report illegal activity.
- ***Collecting Potentially Illegal/Explicit Data:*** It is necessary that parents are clearly informed of the data (text and images) being collected unobtrusively through third-party apps and how it may be used. In order to store the data securely, we have proposed storing it on encrypted flash drives only to be accessed from a dedicated computer system – instead of being stored on a networked drive.
- ***Analyzing Sensitive Data:*** We are in the process of figuring out how to feasibly analyze a large social media data set that may contain imagery that may violate the moral or religious beliefs of the researchers who are charged with interpreting this data. How will this be managed?
- ***Furthering Science:*** How do we make sure we are maintaining the confidentiality, privacy, and security of our data and advance science? For example, once the social media data set has been collected, how should researchers be vetted prior to allowing them access to the data? When presenting the results of our qualitative analysis, how do we make sure the data has been properly anonymized while still providing enough contextual details to be relevant?
- ***Carefully Framing Online Safety:*** To avoid having our research become a form of policing teens, we have chosen to take a value sensitive design [4,5] and user-centered approach to online safety, which emphasizes teens as users and agents of their own online safety (instead of focusing just on parents). Our goal has been to be respectful of teens’ developmental growth and privacy as we set out to help teens protect themselves from and be more resilient to online risks [3,12,14].

CONCLUSION

While we have tackled some of the ethical challenges of adolescent online safety research, we hope that participating in the GROUP 2016 Ethics workshop will help us address some of our remaining questions.

ABOUT THE AUTHORS

Arup Kumar Ghosh

Arup Kumar Ghosh is a doctoral candidate in the Department of Computer Science at the University of Central Florida. He is a member of the Socio-Technical Interaction Research (STIR) Lab at UCF. His primary research interests include Human-Computer Interaction (HCI), Mobile Computing, and Social Media. He received M.S. in Computer Engineering from UCF.

Karla A. Badillo-Urquiola

Karla Badillo-Urquiola is a Ph.D. student in the Modeling and Simulation program at the University of Central Florida. She is a member of the Socio-Technical Interaction Research (STIR) Lab at UCF. Her current research interests include Human-Computer Interaction (HCI), Human Factors Psychology, and Instructional Design. She plans to leverage her interdisciplinary background to develop better training and intervention strategies for the online safety of teens, especially those underrepresented and at-risk.

Muhammad Uzair Tariq

Uzair Tariq is a Ph.D. student in the Department of Computer Science at the University of Central Florida. He is a member of the Socio-Technical Interaction Research (STIR) Lab at UCF. His research aims to combined Human-Computer Interaction (HCI) with machine learning and Computer Vision to build better solutions for keeping teens safe online.

Pamela J. Wisniewski

Dr. Wisniewski is an Assistant Professor in the Department of Computer Science at the University of Central Florida and Director of the Socio-Technical Interaction Research (STIR) Lab at UCF. Her research expertise lies at the intersection of Human-Computer Interaction, Social Media, and Privacy. Her goal is to help people meaningfully engage with one another online, and to do so safely. Her work has received Best Paper Awards at ACM SigCHI conferences.

REFERENCES

1. Awais Adnan and Muhammad Nawaz. 2016. RGB and Hue Color in Pornography Detection. In *Information Technology: New Generations*, Shahram Latifi (ed.). Springer International Publishing, 1041–1050. Retrieved July 21, 2016 from http://link.springer.com/chapter/10.1007/978-3-319-32467-8_90
2. Giuseppe Amato, Paolo Bolettieri, Gabriele Costa, Francesco la Torre, and Fabio Martinelli. 2009. Detection of Images with Adult Content for Parental Control on Mobile Devices? In *Proceedings of the 6th International Conference on Mobile Technology, Application & Systems* (Mobility ’09), 35:1–35:5. http://doi.org/10.1145/1710035.1710070
3. Lee B. Erickson, Pamela Wisniewski, Heng Xu, John M. Carroll, Mary Beth Rosson, and Daniel F. Perkins. 2015. The boundaries between: Parental involvement in a teen’s online world. *Journal of the Association for Information Science and Technology*: n/a-n/a. http://doi.org/10.1002/asi.23450
4. Batya Friedman, Peter H. Kahn Jr, Alan Borning, and Alina Huldtgren. 2013. Value Sensitive Design and Information Systems. In *Early engagement and new technologies: Opening up the laboratory*, Neelke Doorn, Daan Schuurbiers, Ibo van de Poel and Michael E. Gorman (eds.). Springer Netherlands, 55–95. Retrieved May 22, 2016 from http://link.springer.com/chapter/10.1007/978-94-007-7844-3_4
5. Batya Friedman, Peter H. Kahn, and Alan Borning. 2002. *Value Sensitive Design: Theory and Methods*. Retrieved from http://faculty.washington.edu/pkahn/articles/vsd-theory-methods-tr.pdf
6. Homa Hosseinmardi, Sabrina Arredondo Mattson, Rahat Ibn Rafiq, Richard Han, Qin Lv, and Shivakant Mishra. 2015. Detection of Cyberbullying Incidents on the Instagram Social Network. *arXiv:1503.03909 [cs]*. Retrieved July 20, 2016 from http://arxiv.org/abs/1503.03909
7. Ralph M. Perhac Jr. 1996. Defining risk: Normative considerations. *Human and Ecological Risk Assessment: An International Journal* 2, 2: 381–392. http://doi.org/10.1080/10807039609383615
8. Scott Krig. 2014. Ground Truth Data, Content, Metrics, and Analysis. In *Computer Vision Metrics*. Apress, 283–311. Retrieved July 21, 2016 from http://link.springer.com/chapter/10.1007/978-1-4302-5930-5_7
9. Kimberly J. Mitchell, Lisa M. Jones, David Finkelhor, and Janis Wolak. 2014. *Trends in Unwanted Online Experiences and Sexting : Final Report*. Crimes Against Children Research Center. Retrieved from http://scholars.unh.edu/cgi/viewcontent.cgi?article=1048&context=ccrc
10. Erika S. Poole and Tamara Peyton. 2013. Interaction Design Research with Adolescents: Methodological Challenges and Best Practices. In *Proceedings of the 12th International Conference on Interaction Design and Children* (IDC ’13), 211–217. http://doi.org/10.1145/2485760.2485766
11. Jafar Shayan, Shahidan M. Abdullah, and Sasan Karamizadeh. 2015. An overview of objectionable image detection. 396–400.

12. Pamela Wisniewski, Haiyan Jia, Na Wang, et al. 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* (CHI ’15), 4029–4038. http://doi.org/10.1145/2702123.2702240

13. Pamela Wisniewski, Heng Xu, Jack Carroll, and Mary Beth Rosson. 2013. Grand Challenges of Researching Adolescent Online Safety: A Family Systems Approach. Retrieved May 25, 2016 from http://aisel.aisnet.org/amcis2013/SocialTechnicalIssues/GeneralPresentations/10

14. Pamela Wisniewski, Heng Xu, Mary Beth Rosson, Daniel F. Perkins, and John M. Carroll. 2016. Dear Diary: Teens Reflect on Their Weekly Online Risk Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI ’16), 3919–3930. http://doi.org/10.1145/2858036.2858317

15. Rui Zhao, Anna Zhou, and Kezhi Mao. 2016. Automatic Detection of Cyberbullying on Social Networks Based on Bullying Features. In *Proceedings of the 17th International Conference on Distributed Computing and Networking* (ICDCN ’16), 43:1–43:6. http://doi.org/10.1145/2833312.2849567

16. Filter for Web Pornographic Contents Based on Digital Image Processing - ProQuest. Retrieved July 21, 2016 from http://search.proquest.com/openview/371dca98f94376b38a9e1979012a61b6/1?pq-origsite=gscholar&cbl=696410

17. NSF Award Search: Award#1566511 - CRII: RI: Multi-Source Domain Generalization Approaches to Visual Attribute Detection. Retrieved July 20, 2016 from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1566511