Understanding Online Harassment: Definitions, Ethics, and Prevention

Lindsay Blackwell	Zahra Ashktorab	Katherine Lo	Liz Lee
lblackw@umich.edu	parnia@umd.edu	kmlo@uci.edu	liz@onlineSOSnetwork.org

INTRODUCTION

Online harassment is a pervasive problem; 73% of adult internet users have witnessed online harassment, and 40% have personally experienced harassment [4]. Of users who have experienced harassment online, 45% have been the target of more severe threats, including physical threats, stalking, sexual harassment, and repeated threats over a sustained period of time [4]. Young women are particularly vulnerable to more severe forms of online harassment [4]: in a recent study, 25% of young women surveyed had been sexually harassed online (compared with 6% of all internet users), and 26% had been physically threatened (compared with 8% of all users).

The effects of harassment vary from person to person, ranging from anxiety, humiliation, and self-blame to anger and physical illness. Online harassment, however, can "cast a long shadow"—severe harassment can damage an individual's reputation, comfort or safety, due to the persistence and searchability of digital media [4]. Most critically, online harassment contributes to a culture of abuse that threatens to silence the voices of technology users, particularly women [4] and members of racial and sexual minorities [3,4].

It is imperative for the future and success of the Internet that researchers, developers, designers, and users work in tandem to build a better and more empathetic web. In this BoF, we will discuss the various research and design lenses through which online harassment can be better understood. The facilitators come from different disciplines: data science, human-computer interaction, social science, business, and advocacy. Ultimately, we hope to facilitate a group conversation around the many dimensions of online harassment, to help better understand and work toward the prevention of these behaviors.

AUDIENCE

In this session we hope to reach the following audiences: 1) undergraduate or graduate students in computing-related disciplines who are interested in researching and developing solutions for online harassment; 2) technology designers or industry researchers who are interested in the detection and prevention of online abuse on their platforms or more broadly; and 3) those interested in learning about online harassment through a multi-disciplinary perspective (for example, journalists or career professionals). The discussions generated in this session will appeal to participants at all levels.

PLAN OF ACTION

The goal of this BoF session is to encourage an interactive discussion of online harassment among participants. Below we have outlined the structure of our BoF, which will easily scale to accommodate a large audience.

Introduction (10 minutes): The facilitators will introduce themselves, their qualifications, and a brief definition of relevant terms. Facilitators will also provide a short overview of the session so participants know what to expect.

Next, audience members will divide into small groups to allow participants to communicate and discuss the various facets of online harassment. The groups will be given written instructions and discussion prompts, and facilitators will move around the room to answer questions and help to facilitate discussions when needed. The three groups will concurrently discuss the following themes: defining online harassment; the ethical responsibilities of those studying online harassment; and the affordances of online platforms that may enable or help prevent online harassment.

Concurrent Groups:

Group 1 (*20 minutes*): **Discussion of what is and is not online harassment.** While online harassment is widely discussed amongst technology researchers, designers, journalists, and users, there is no standard definition for what *online harassment* entails. In this discussion, we will consider various forms of online abuse (such as trolling, flaming, doxxing, and public shaming) to determine a community definition of *online harassment.* Participants are asked to write down their definition so that definitions can be shared with all participants following the session.

Group 2 (*20 minutes*): **Discussion of the ethical responsibilities of researchers who study online harassment.** Researchers and data scientists use algorithms and data mining to detect and model harassment on various platforms. What is their responsibility to the communities from which they gather data? What are social scientists' responsibilities to the perpetrators, and not just victims, of online abuse? What scientific practices are and are not ethical when researching these types of problems? Groups will be given real-world case studies to discuss, including Minority Media's recent and controversial presentation on virtual reality abuse at the Game Developers Conference (GDC).

Group 3 *(20 minutes):* **Discussion of the technical affordances that may enable or help prevent online harassment.** In order to discourage harassment in online spaces, we must first understand the different characteristics of these environments that may afford abusive behaviors. Thus, a critical step in the prevention of online harassment is to study the technological affordances that enable—and in many cases, exacerbate—harassing behaviors. We will examine online community platforms where harassment is known to occur (e.g., Yik Yak, reddit, and Twitter) and discuss the technical affordances on these platforms that may enable various forms of harassment (e.g., anonymity; persistence; moderation tools). We will further discuss the implications of these affordances in designing and developing for online platforms.

Regrouping *(20 minutes):* The facilitators will open the floor to each of the discussion groups, who will prepare a summary of their discussions to share with the rest of the participants. Group 1 will share their discussions on a definition of online harassment. Group 2 will share their discussion on the ethical responsibilities of researchers who study online harassment. Group 3 will share the results of their discussion on the technological affordances that may enable or help mitigate abusive behaviors online.

Conclusion (10 minutes): The facilitators will again take the floor, summarizing any salient points from the community discussion and sharing additional information and resources if needed. With any remaining time, participants will be encouraged to ask questions of the four facilitators or bring additional points to the group for discussion.

OUTCOMES
At the conclusion of the BoF session, we hope to: 1) generate a community definition of *online harassment*; 2) begin discussion on the ethical responsibilities of researchers who study online harassment; and 3) conclude with in-depth discussion of the technological affordances that both enable or may help mitigate online harassment. Participants will leave with a better understanding of what online harassment entails, as well as what differentiates certain types of online abuse from others. In addition, participants will leave with a framework for the ethical responsibilities of researchers across industry and academia who study online abuse, preparing our participants to be more thoughtful about their own work and in the consideration of others'. Ultimately, participants will leave with a better understanding of the technological affordances that may enable and can help prevent online abuse. This will help participants better understand the existing online communities they currently study and use, and will better equip participants to design and develop solutions for the prevention of online harassment in current and future online community platforms.

PARTICIPATION STATEMENT
All facilitators have made a commitment to attend the conference if accepted.

FACILITATORS
Lindsay Blackwell is a PhD student at the University of Michigan School of Information and a fellow with the Institute for Research on Women and Gender. Her research explores misbehavior in online communities, including trolling and online harassment. She uses traditional social science methods such as interviews, surveys, and content analysis to better understand the underlying social mechanisms and technological affordances that both enable [5] and may help mitigate [3] online abuse.

Zahra Ashktorab is a PhD student at the iSchool at the University of Maryland College Park. Her research focuses on cyberbullying detection and mitigation through mixed methods. Using a data-driven approach, she has studied Ask.fm, an anonymous question-answer site popular with teens, to understand the nature of discourse that occurs alongside the site's pervasive cyberbullying [1]. She has also led participatory design sessions with teenagers to design cyberbullying mitigation tools on social media platforms [2].

Katherine Lo is a PhD student at the University of California, Irvine. Her research examines the dominant narratives that motivate diversity initiatives in technology and how they are differently enacted in practice. In particular, she examines the affordances of online community platforms that enable diverse and marginalized people to organize safely. She is currently working with Intel on an initiative to develop guidelines for online citizenship in video game communities.

Liz Lee is the Co-founder and Executive Director of Online SOS (www.onlineSOSnetwork.org), a nonprofit providing free crisis coaching and other professional services to people experiencing online harassment. In addition to providing services, Online SOS focuses on technology and advocacy to bring attention to the cost of online harassment and better address the needs of affected individuals. Liz is using her investment and business experience at Morgan Stanley Alternative Investment Partners to broaden the conversation about online abuse to investors, while ensuring individuals experiencing harassment are able to get the help they need.

REFERENCES
[1] Ashktorab, Z., Kumar, S., De, S., and Golbeck, J. (2014) "iAnon: Leveraging Social Network Big Data to Mitigate Behavioral Symptoms of Cyberbullying."

Social Media Expo at *iConference 2014.* Berlin, Germany. March 4-7, 2014.

[2] Ashktorab, Z. and Vitak, J. (2016). "Designing Cyberbullying Mitigation and Prevention Solutions through Participatory Design With Teenagers." In *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems* (CHI '16). San Jose, CA. May 7-12, 2016.

[3] Blackwell, L., Hardy, J., Ammari, A., Veinot, T., Lampe, C., and Schoenebeck, S. (2016). "LGBT Parents and Social Media: Advocacy, Privacy, and Disclosure during Shifting Social Movements." In *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems* (CHI '16). San Jose, CA. May 7-12, 2016.

[4] Duggan, M. (2014). "Online harassment." Pew Research Center.

[5] Ellison, N., Blackwell, L., Trieu, P., Lampe, C., and Morioka, T. "Anonymous Online Interaction and its Benefits for Adolescents." Presented at the annual meeting of the International Communication Association (ICA '16). Fukuoka, Japan. June 9-13, 2016.