online sos

Into 2020:

The State of Online Harassment and Opportunities for Collaboration

How we got here and
what do we do now

OnlineSOS is a nonprofit organization that supports individuals facing online harassment through actionable resources and advocacy.

Publication Authors

Founder	**Elizabeth (Liz) Lee**
Editor, Writer	**Terri Lee Harel**
Contributing Writer	**Allison Gauss**
Advisor	**Katherine Lo**
Design	**Alex Chen**

OnlineSOS is a nonprofit organization that supports individuals facing online harassment through actionable resources and advocacy.

Support for This Publication

WALLACE GLOBAL FUND | Craig Newmark Philanthropic Fund | Vodafone Americas Foundation

The funding for this report is made possible by the visionary leadership of Wallace Global Fund, Vodafone Americas Foundation, and Craig Newmark Philanthropic Fund. This report was developed in partnership with the countless organizations and individuals—experts, advocates, researchers, journalists, partners, and friends—who contributed their personal stories, precious time, brilliant talent, sharp minds, and hearts.

You recognized online harassment and abuse as a problem before it became hot topic or obvious. You had the insight, foresight, and boldness to act. Thank you for being trailblazers, risk-takers, and innovators. Thank you for your commitment to change and willingness to take action, for your commitment to impacting the next generation of journalists and activists, women and girls, to thrive on the internet, free from online harassment and abuse.

Contact Us

The intention of the State of Online Harassment report is to catalyze and facilitate collaboration. For any questions, comments, or to get involved, please contact team[at]onlinesos[dot]org.

Table of Contents

Welcome Letter

Hi there! Welcome to the conversation!

No matter how you got to this report, we're so grateful you are joining the dialogue around online harassment and abuse.

Ten years ago, I graduated from UPenn and moved to New York City to start my first job in finance. What started as a period of excitement turned frightening when I was stalked, threatened, and extorted online over several months. The harassment started anonymously, but escalated to my apartment door being kicked in. I was confused and afraid, with nowhere to turn.

While working at Morgan Stanley, I hesitated to share my story, fearing stigma and a possible negative impact on my career. But in 2015, I realized that what happened to me was happening to others and with devastating consequences. I asked myself, "How many more people need to die before I do something?" At the peak of my professional success and confidence, I quit my job to create the solutions I wish I once had. OnlineSOS was born in 2016.

Since then, the escalation from online to real life violence and abuse has been alarming:

- The distinction between online and offline has disappeared. This is especially evident in the aftermath of the 2016 presidential election and the violence that devastated the Capital Gazette newsroom and communities in Charlottesville, Parkland, and Christchurch.
- Online harassment is used to amplify media manipulation, workplace issues, and domestic violence. Technology is abused to systematically create professional, political, and ideological intimidation and exacerbate discrimination and the sexual and employment-related harassment and abuse surfaced by the #metoo movement.

As OnlineSOS helped targeted individuals with mental health support and other resources, online harassment evolved with more ferocity than any solutions could handle. So a major question still loomed: "How could we help lay the foundation for the next decade of change?"

This report shares what we've learned from being on the front lines of support. Through individual stories and data, as well as offering trends and frameworks, this report highlights untold stories, victories, gaps, and challenges related to online harassment. How can we connect the dots to create a line or, better yet, a graph that shows the behavioral patterns in the system of online harassment? How do we use that information to better influence change?

We also aim to honor the visionaries who advocated and educated the general public about the harms of online harassment before it was a widely recognized phenomenon. Before Jeff Bezos wrote his Medium post about sextortion, there were countless individuals who spoke out on Twitter, in boardrooms, at dinner parties, and at public convenings. These individuals often came from marginalized communities, had been personally targeted, and/or used their talents to advocate for others. Their efforts must not be forgotten—as their initiatives, experiences, and lessons are key to creating effective and lasting change.

Now, in 2019, we are at an inflection point. There's a unique window to catalyze change. An unprecedented level of public consciousness about online harassment and its related issues—privacy, civility, viral extremism, and violence against women—makes it time to act.

I believe every conversation and every action makes a difference. There's an opportunity to come together at the intersection of tech, media, research, law, policy, finance, government, and civil society. What kind of future is possible by combining our talents, by collaborating across the aisle, by taking risks? In the words of Albert Einstein, "We can't solve problems by using the same kind of thinking we used when we created them." Let's invite new voices, new ideas, new collaborations to create—together—an internet where everyone thrives; the internet we're proud to leave behind.

— Liz Lee, founder of OnlineSOS

Executive Summary

- Online harassment continues to grow and evolve despite its broader recognition as a problem. More than 85 million American adults have experienced online harassment. Women, people of color, religious minorities, and other marginalized groups are more likely to experience harassment both more often and with more intensity. Certain professionals, like activists or journalists, are also more vulnerable to harassment for simply doing their jobs.
- Tactics of online harassment deployed against individuals or specific groups are similar to those used in media manipulation and disinformation. Online harassment has steadily emerged not only as an effective tool to silence, exclude, and harm individuals, but also as a way to reshape democracy in practice.
- To be effective, solutions to online harassment need to be centered on the individual's experience. Without cultivating a rich picture of the end-to-end experience, the context of harassment, and its lasting effects, proposed interventions will continue to fall short and the status quo upheld.
- To create lasting change, developing a common vocabulary for describing various tactics of online harassment is key. For an individual, a common vocabulary can affect how they seek help, report abuse, find resources, decide what to do, and recover. For the experts, groups, or researchers addressing online harassment, a common vocabulary can streamline collaboration and facilitate more effective data collection, research, policy, legal remedies, and victim support. We can't change what we can't name.
- Whether a solution is focused on serving an individual or seeking change-at-scale, it will require a combination of interventions, disciplines, and expertise. Online harassment is complex, so mutual understanding and collaboration is critical to systemic change.
- Looking towards 2020, it's more urgent than ever to address online harassment. Most important is to focus on creating better outcomes for targets of online harassment. When targets can recover and reintegrate, rather than be silenced or excluded, harassers and manipulators lose ground. Left unaddressed, online harassment (and its consequences) will only become an increasingly effective tool for manipulation and marginalization.

Here is the TL;DR version of the report in five key takeaways:

1. Online harassment is one of the most important social issues of our time.
2. People's experiences hold the key to effectively understanding and addressing online harassment.
3. Creating a common vocabulary is a critical step to scoping the problem, engaging in productive discourse, and developing solutions.
4. Grassroots groups supporting individuals play a key role in the ecosystem; it is important to provide funding to maintain and distribute their offerings.
5. Creating better outcomes for individuals is possible through multi-stakeholder efforts.

Below are a few key frameworks and graphics developed to share our understanding of online harassment

The report aims to portray a snapshot of online harassment in this moment of time. We hope the report serves as a valuable tool to catalyze meaningful dialogue and collaboration that, in turn, results in the refinement and further development of the report with additional input.

Note from the Editor

In developing this report, every day presented our team with a new thread to pull. Dozens of stories a week revealed yet more examples of online harassment in action and its consequences, both personal and societal. There are endless angles from which to understand and address online harassment. How could we encapsulate the expanse of this issue in a single document? It still keeps me up at night. Each anecdote and idea warrant deeper investigation and interdisciplinary understanding or collaboration.

With that in mind, this report represents what our team has identified as salient and urgent needs in the lead up to 2020. While, indeed, it's impossible to capture every deserving detail in one report, we hope this overview of the online harassment landscape not only inspires the historical thinking needed to learn from the past, but also sparks the difficult discussions civil society, technologists, policy makers, educators, academics, and average internet citizens need to have to create new systems of online participation and community that incentivize safety, freedom, and productive discourse.

— Terri Harel

Introduction

Scope and Methodology

The OnlineSOS Landscape Report provides a comprehensive look at the current state of online harassment in the U.S. This includes an analysis of the evolution of online harassment, the tactics (or combination of tactics) a targeted individual may experience, and how organizations, experts, technology companies, and academics address or approach the topic. In this report, you will find:

- How we got here: an analysis of online harassment over the years
- What a targeted individual may experience, including what they need when faced with harassment
- Who addresses online harassment and how
- Gaps in research, solutions, or resources and why it's urgent to fill these gaps
- Opportunities and recommendations for productive collaboration
- Directories of 1) resources and initiatives and 2) research about online harassment and adjacent topics

To create this report, the OnlineSOS team catalogued and categorized over 700 pieces of information about online harassment, including reports, resources, news articles, and research papers. Some sections of this report also include insights derived from the real life incidents OnlineSOS triaged and supported from 2016 to 2018. During these years, the organization provided direct service support to individuals facing online harassment. Finally, the OnlineSOS team connected with more than 30 stakeholders to inform this report. This group included online harassment experts, researchers, academics, journalists, tech leaders, lawyers, and others involved in online harassment work.

The American public widely recognizes the problem of online harassment—referenced in this report as an umbrella term for a large set of tactics meant to silence, intimidate, threaten, or inflict pain on a targeted individual. (For more on specific tactics, please see Behaviors and Tactics in Part 2.) Yet there are few available concrete solutions for reducing the frequency and intensity of online harassment and the reach of abusers. Despite efforts by private individuals, organizations, researchers, and private sector companies to address the problem, little progress has been made to curb abuse [1]. The questions, then, are:

1. How did online harassment become what it is today?
2. Why is progress slow or stagnant, despite wide recognition of the problem?
3. What will accelerate progress toward reducing online harassment?

In this report, we aim to address these questions. In doing so, we will also highlight the nuances of online harassment, its effects, and its relationship to broad, systemic issues in the United States. Lastly, we will shed light on efforts to address online harassment in the last few years, and identify potential opportunities to reduce it in long-lasting ways.

Focusing on Online Harassment

In this report, we'll focus exclusively on online harassment—an umbrella term for a large set of tactics meant to silence, intimidate, threaten, or inflict pain on a targeted individual. Although we discuss "trolling," which has become a common, colloquial way to refer to any behavior meant to annoy, rile up, bother, or harass someone else, we will be more specific about the tactics and behaviors used to harass individuals [2].

In addition, OnlineSOS will look at online harassment specifically as it relates to adults. "Cyberbullying," which is most often used to describe abusive online behavior between children and young adults, is not the focus of this report. Although it features some overlap in patterns and behaviors, cyberbullying requires unique interventions and considerations, including schools, parents/guardians, and legal options applicable only to minors.

Part 1

The Evolution of Online Harassment

Below, we'll provide an overview of the evolution of online harassment, with a primary focus on how the 2016 elections marked a turning point in its reach and development. We'll also look at what will be important to consider as we head into 2020.

Online Harassment Before Gamergate

Online harassment isn't new. As soon as networked computers popped up, people found ways to abuse these new forms of communication and community. An article published in the New Yorker in 1994 details its author's first "flame," and, no, it's not the romantic kind. The writer, naive and excited about the new world wide "net," has his cyber-bliss rudely interrupted by a vile email supposedly sent by a fellow journalist. Newly "flamed," the writer sets off to find out what other bad behavior might exist on the net. To his horror, he discovers "flame wars"—exchanges of insulting messages—on message boards and websites around the net. He's frightened by what he finds and his once bright outlook on the wild, wild web turns suspicious and dark [3].

This kind of flaming or "trolling" didn't always have its roots in malicious intimidation. But it soon morphed into a tactic of more harmful intent. Researcher Ben Radford argued that trolls saw themselves as digital clowns, exposing the foibles and folly of a community. Whatever noble view trolls held of themselves, however, they still acted with the intent to create an unpleasant experience for others online. Importantly, the provocation often wasn't targeted toward individuals.

An Example of Trolling

A troll joins an astronomy discussion forum, posing as a genuinely interested user. They then vehemently assert that the Earth is flat in order to provoke an emotional and verbal reaction from community members.

Level 1 — Trolling for the Lulz

Figure 1.1A

This early type of trolling was not unlike someone who comes upon a peaceful pond and throws rocks in it to make waves and disturb the ecosystem. It even seems like a victimless crime, if you don't account for all the creatures that live in the pond.

Tactics of "trolling for the lulz," as this was known, were easily translated into behaviors with malicious intent, including:

- **Sending messages that range from** rude remarks to encouraging suicide
- **Cyberstalking**, an invasion of privacy and source of intimidation that can also involve identity fraud and financial hacking
- **Doxxing**, releasing private information to the public, which puts targets at risk for stalking, violence, and physical intimidation
- **SWATing**, sending police to a target's home, which puts targets (and innocent bystanders) in harm's way
- **Coordinating mob harassment**, which can silence and discredit, especially journalists, activists, and oppressed groups
- **Spreading false information**, which damages targets' professional reputations and their ability to work

In this period, spamming, doxxing (when information considered private, like a home address or social security number is published or broadcast online), and non-consensual distribution of intimate images were used to threaten, intimidate, and silence specific targets. Examples of high-profile cases and targets include Kathy Sierra (2007), Anita Sarkeesian (2012 and 2014), Caroline Criado-Perez (2013), the spamming case at Occidental College (2013), and Jennifer Lawrence (2014).

These cases inflicted psychological trauma and distress on targeted individuals, and had a serious impact on their personal and professional lives. Blogger and software designer Kathy Sierra, fearing for her own safety and that of her family, disappeared from online spaces for an extended period of time and moved across the country. Anita Sarkeesian needed to leave her home. Caroline Criado-Perez received so many threats of rape and violence on Twitter that she was unable to function. In response, Twitter, which had been around for seven years by then, added a "Report Abuse" button [4].

> *When the hacking thing happened, it was so unbelievably violating that you can't even put it into words...I think that I'm still actually processing it.*
>
> — Jennifer Lawrence in a 2017 interview, three years after nude photos of her were hacked and distributed without her consent.

Level 2 — Trolling for Individual Attacks

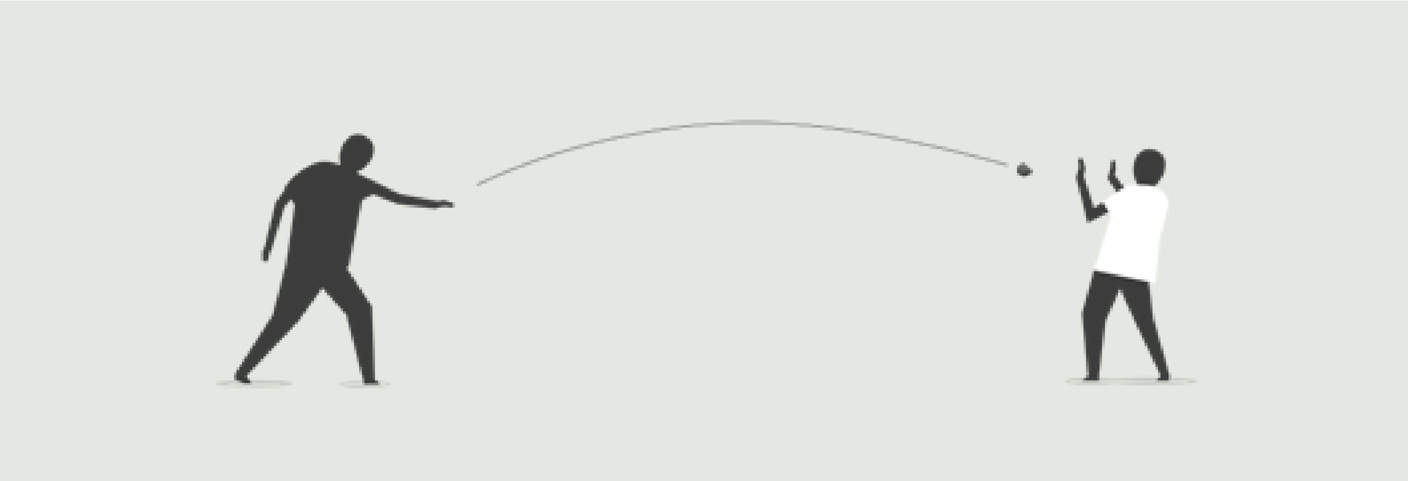


Figure 1.1B

- Unlike "trolling for lulz," tactics of trolling were used to target individuals. Instead of inciting arguments within a community for the sake of entertainment, a troll might attack an individual with whom they disagree. Old tactics such as impersonation and outrageous claims would be used to harm a target.
- This can be compared to someone throwing rocks at another individual. The tools are the same, but by turning them on an individual, the troll is able to do a great deal of harm.
- Example: A troll impersonates their ex-partner and posts offensive content to draw criticism and negative attention to the target. The goal is to cause harm specifically to the ex-partner, in the form of emotional suffering, reputational damage, and further harassment.

The Gamergate Period: 2014 – 2016

By 2014, distinctions between life on and offline had gotten blurry. iPhones and other smartphones had been around for almost a decade, and most people in America happily carried around their digital lives in their pockets. Twitter and other social media platforms were not only ubiquitous for fun, but also became more important to people's professional lives.

But this increased access to online communities and platforms was a two way street. It also gave malicious actors greater reach. They could find and target individuals 24 hours a day, 7 days a week. It became increasingly easy to ramp up the intensity of abuse too, using platforms like Reddit or 4chan as staging grounds. Pew Research's 2014 report about online harassment found that 40 percent of Americans had been harassed online by that point and that 73 percent had witnessed some form of harassment, ranging from "efforts to purposefully embarass someone" to more severe forms like sexual harassment and stalking.

Zoë Quinn, an independent game developer, was quickly harassed online after releasing a game in 2013. Misogyny had been present in online spaces—especially within the gaming community—since the advent of the internet, so this was not surprising. However, the end of her short relationship with a man named Eron Gjoni sparked an online harassment campaign that lasted years. Its effects still reverberate.

Sarah Jeong, a journalist and the author of The Internet of Garbage (2015, updated 2018), succinctly summarizes Gamergate and its impact:

> *Gamergate is complicated. It's also fairly simple: it's a harassment campaign instigated by Zoë Quinn's ex-boyfriend, Eron Gjoni. Quinn was already being harassed before Gjoni, but her ex amplified it many times over: Before Gjoni's post, she had received 16 megabytes of abuse. When she stopped saving threats last December [2014]—because she couldn't keep up with the bombardment—she had 16 gigabytes: 1,000 times more.*
>
> *... Quinn and Gjoni dated for five months. After the relationship ended, he created 'The Zoe Post,' a blog post alleging that she had been unfaithful to him during their relationship. A Boston Magazine profile of Gjoni states, 'By the time he released the post into the wild, he figured the odds of Quinn's being harassed were 80 percent.' He was right. Quinn received a barrage of threatening messages, like If I ever see you are doing a pannel [sic] at an event I am going to, I will literally kill you. You are lower than shit and deserve to be hurt, maimed, killed, and finally, graced with my piss on your rotting corpse a thousand times over.*
>
> *Quinn was doxed—her personal information, including her address and Social Security number, was published. She moved. The harassment continued— and with some patient investigation, Quinn was able to document Gjoni egging on her harassers from behind the scenes. What Gjoni was doing was both complicated and simple, old and new. He had managed to crowdsource domestic abuse.* [5]
>
> — Sarah Jeong, Internet of Garbage, page 16-17

During Gamergate, men's rights advocates (MRA groups) and other anti-feminists used platforms like 4chan to coordinate their attacks on specific individuals. This marked a turning point in coverage of online harassment, particularly mob harassment.

Even months and years after "The Zoe Post", Gamergate worked like a hurricane that sweeps up and destroys everything in its path. Anyone who critiqued sexism in gaming, voiced support for targets, or was deemed a SJW ("social justice warrior," a term created by harassers to describe anyone promoting progressive views, including feminism) could be targeted.

What's more, 4chan denizens uninvolved in the original events of Gamergate were easily recruited and egged on to become harassers. "Lot of support, and a ton of people are picking up the self-chastising when people start getting insulting. It took a few days of 4-5 of us doing it but it's taking off," wrote one user involved in organizing IRC chats, according to an Ars Technica report. In another transcript, a user wrote, "i couldnt care less about vidya [Quinn's game], i just want to see zoe receive her comeuppance." [6]

Gamergate even consolidated other campaigns in its wake, making it a Category 5 storm of ongoing mob harassment. Historically, trolls had often posed as a member of an existing community so that their content would be heard and received by targets. But this behavior was weaponized for ideological purposes, as was the case with the #endfathersday campaign—a hashtag that trended on Twitter in June 2014—which was taken up by Gamergate. The campaign intended to create discord and backlash within the feminist community. Harassers created sockpuppet accounts—an identity used to deceive and/or maintain anonymity—designed to look like politically active feminists and posted content to disrupt communities and weaken one side of a debate [7].

Gamergate sparked new models for harassment and left many victims and targets in its path. For example, it capitalized on the manufactured narrative that the controversy had nothing to do with Quinn personally and everything to do with "ethics in gaming journalism." It became a successful experiment in creating unfounded ideological strife.

> *Much of 2015 would also cover the continued harassment of Gamergate, with other victims like Brianna Wu, Anita Sarkeesian as well as others...There are many victims of harassment who are not as publicly known- Gamergate affected everyday games developers, lesser known or upcoming critics, artists, etc. Gamergate had many targets and many victims.*
>
> — Caroline Sinders, "An Incomplete, but Growing, History of Online Harassment Campaigns since 2003"

Level 3 — Trolling for Ideological Attacks

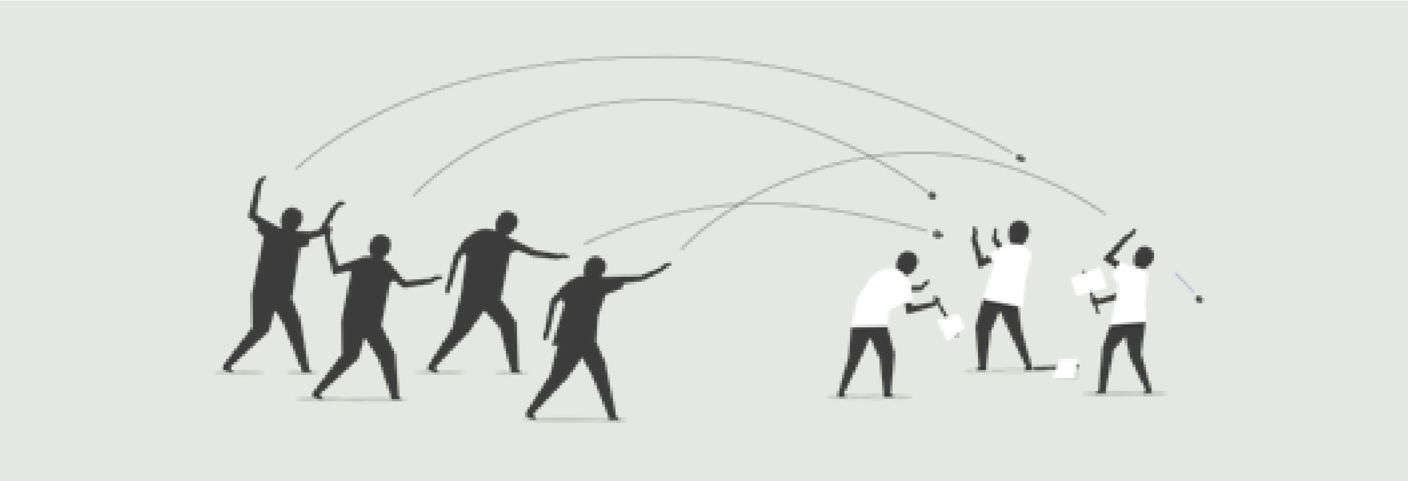


Figure 1.1C

- Whereas many earlier cases of harassment sought to harm targeted individuals, during and after Gamergate harassers began using digital tools and tactics to weaken an ideological group or community (although Gamergate isn't the first such case, it brought these tactics to the mainstream). These coordinated campaigns may also promote a specific ideology. Harassers may target individuals who are part of the opposing group or use misinformation to turn public opinion against the targeted group.
- Going back to the rocks analogy, this form of harassment can be visualized as someone throwing rocks at a group of people based on their identities or ideologies. The aggressor may be a lone troll who skillfully uses trolling tactics (rocks) or it could be a coordinated attack by like-minded trolls. The goal is to weaken or intimidate the other side of a debate.
- **Example:** A group of anti-feminists uses bots, dogpiling (overwhelming a target with questions, threats, insults and other tactics meant to discredit, silence, or shame a target), and other tactics to threaten and intimidate leaders of a women's rights group. The goal is to silence these activists and discourage others from getting involved in the cause.

@sassycrass:

#YourSlipIsShowing: Documenting a Hoax

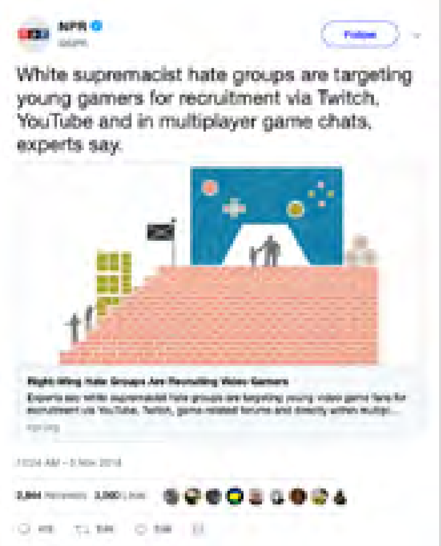
Figure 1.2A Twitter Moments preserved by @kazamacat, developed by @sassycrass

It's important to note that the concerns about such organized harassment that came into sharp relief during and after Gamergate were well-voiced, but mostly ignored by the general public and platforms. Campaigns against Black women, in particular, were present long before Gamergate but little had been done to stop it. For example, Black women on Twitter had for years raised concerns about harassment and pointed out campaigns to discredit and silence them. [8]

During the Gamergate period, Shafiqah Hudson (@sassycrass), @so_treu, and other Twitter users revealed that the #endfathersday campaign was a coordinated effort by trolls posing as feminists. Over 200 deceptive accounts were identified and documented—with the hashtag #yourslipisshowing—in a Storify in 2014 to help people identify and block such accounts. [9] (Unfortunately Storify shut down in 2018, but Tweets were preserved by Twitter user @kazamacat in Twitter Moments.) [10]

Activists have been instrumental in highlighting the danger of such abuse and many warned that Gamergate and the tactics used to organize and perpetuate it would have far-reaching consequences for campaigns of harassment, manipulation, and hate. Those who have raised the issue of online harassment early and often have been, themselves, members of targeted marginalized communities. Shireen Mitchell, a founder, speaker, author, and activist, has pointed this out extensively, as has Jamia Wilson, the director of Feminist Press, among many others. For a deeper read, refer to Mitchell's Twitter Moment on this topic or this 2016 Guardian article that references much of Wilson's important work, noted and linked in the endnotes [11] [12].

> Activists have been instrumental in highlighting the danger of such abuse and many warned that Gamergate and the tactics used to organize and perpetuate it would have far-reaching consequences for campaigns of harassment, manipulation, and hate. Those who have raised the issue of online harassment early and often have been, themselves, members of targeted marginalized communities.



NPR
@NPR

Follow

White supremacist hate groups are targeting young gamers for recruitment via Twitch, YouTube and in multiplayer game chats, experts say.

Right-Wing Hate Groups Are Recruiting Video Gamers

Experts say white supremacist hate groups are targeting young video game fans for recruitment via YouTube, Twitch, game-related forums and directly within multip...

npr.org

10:24 AM - 5 Nov 2018

[illegible] Retweets [illegible] Likes

To further illustrate this, in 2018 NPR tweeted about an investigation into how the alt-right recruited members on gaming networks. The replies, linked under Figure 1.2B, offer a deep look into how far back such efforts and tactics reach. Replies included, "Anybody who's used these platforms has been aware of this and saying it for YEARS" and "no surprise there."

Gamergate and its copycat campaigns had a profound impact on how abusers and bad actors learned to organize for future campaigns, including the ones that targeted American voters in the 2016 presidential election.

Figure 1.2B: See full thread
https://twitter.com/NPR status/1059511664951799808

Online Harassment in a Post-2016 Election World

The same tactics used in Gamergate, like sockpuppeting and mob harassment, were used by Russian trolls in the run up to the 2016 presidential election. Posing as American voters, these trolls spread polarizing and often false content on social media platforms and provoked extremism among unsuspecting citizens[13].

Such actions aren't limited to paid campaigners. Groups like the alt-right and the anti-vaccination movement are also taking their cues from Gamergate and paid trolls to harass and silence targets.

Recently, anti-vaccination activists have targeted medical professionals who support vaccination in publications and online forums. The Guardian reported that "networks of closed Facebook groups with tens of thousands of members have become staging grounds for campaigns that victims say are intended to silence and intimidate pro-vaccine voices on social media. The harassment only exacerbates an online ecosystem rife with anti-vaccine misinformation, thanks in part to Facebook's recommendation algorithms and targeted advertising."[14] These tactics, including staging mob attacks and spreading misinformation, harken back to Gamergate.

Platforms' engagement-based models enable violent extremism, white supremacy, and other hateful content. This allows a relatively small group of people to have an outsized voice in discourse.

Facebook, Twitter (which no longer sorts posts only chronologically), Reddit, YouTube, and Google Search all rely on specific indicators to determine what's shown to people. In general, algorithms pick up on trending topics, combinations of engagement indicators (likes, retweets, comments, upvotes), and personal preferences to determine what reaches the tops of users' newsfeeds. These models have encouraged misinformation to spread and also enabled hateful comments or content to become more visible.

We'll explore how platforms are addressing this issue later in the report.

> *Far-right actors frequently game Twitter's trending topics feature to amplify certain stories or messages. And YouTube gives a platform to conspiracy theorists and fringe groups who can make persuasive, engaging videos on outrageous topics.*
>
> *Often, anons will work together to get a hashtag to trend, sometimes by creating large amounts of fake accounts. In other instances, they will take an extant hashtag, like #BlackLivesMatter, and manipulate or "hijack" it—in this case, posting messages critical of BLM to diminish the ability of supporters to use the hashtag to find each other.*[15]
>
> — Caroline Sinders, "An Incomplete, but Growing, History of Online Harassment Campaigns since 2003"

Online Harassment and Disinformation

At first glance, online harassment may not seem part and parcel of the disinformation campaigns we now associate with Russian "trolls." However, when the strategy to sway public sentiment, mold an electorate, and play into distinct political and geopolitical goals includes coordinated, sophisticated action meant to incite anger and fear, we're talking about classic trolling.

Moreover, when the goal is to sway public opinion, online harassment is an incredibly potent tool. It's also hard to control once unleashed. The exact tactics and tools used by early trolls and later adapted to vicious online harassment campaigns are tailor-made to silence opposing viewpoints. In turn, they make it much more difficult for a regular person to discern fellow citizens from paid agents, or to discern individual opinions from facts or campaigns of propaganda.

Level 4 — Trolling for Large-Scale Manipulation

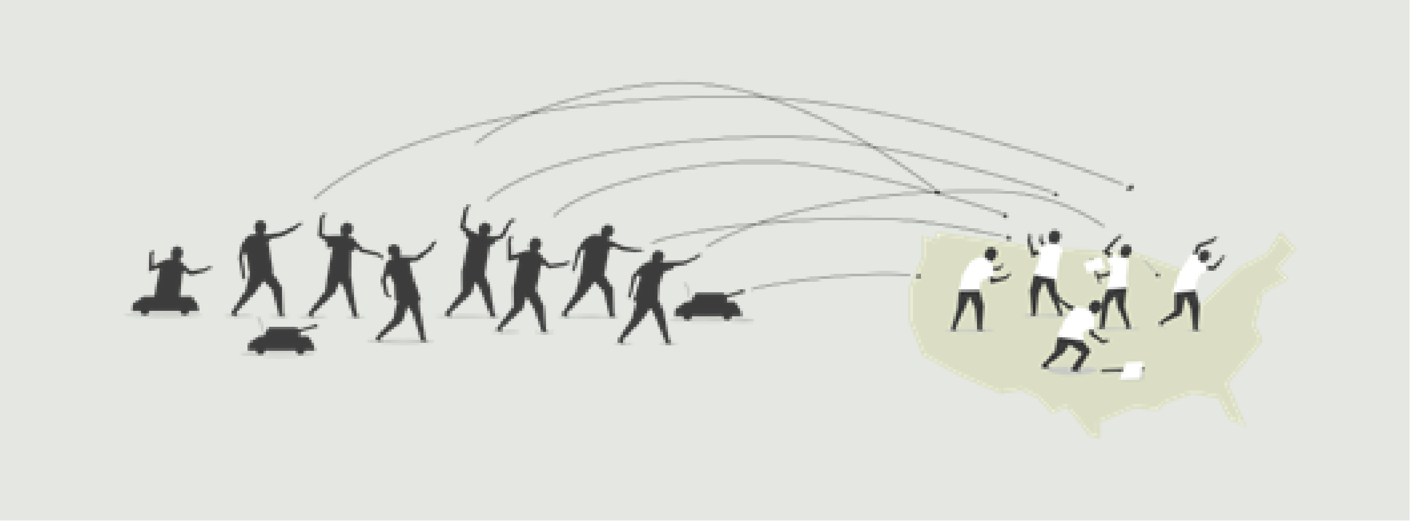


Figure 1.1D

- Today, harassment tactics have been further weaponized to manipulate the wider population. While the perpetrators' intent remains waging and winning an ideological battle, these campaigns encroach on politics, society, and democracy when absorbed by larger audiences. In addition, while harassment for ideological reasons may consist of direct attacks from like-minded individuals, this newer, more broad manipulation may be financed by a leader or organization. Tactics may be amplified by paid agents or bots to make an issue or viewpoint seem much greater than it is. Coordinated trolling may be used to manipulate people who may not have any direct connection to the issue or political mission at hand.
- This could look like a powerful individual or organization mobilizing rock-throwing mercenaries. The rocks are not aimed at individuals but are more numerous and cause fear and mistrust in the population being bombarded. The powerful individual might then claim they are the only leader who can protect the population from rocks.
- Example: A think tank organizes a campaign to publicize any crimes committed by undocumented immigrants. Their paid trolls and bots spread stories far beyond their usual audience and share rumors and false reports. The goal is to convince the population that immigrants are a threat to the country.

In 2015, Russia's Internet Research Agency (IRA) started a campaign with the hashtag #ColumbianChemicals. The fabricated backstory was that a chemical plant in Louisiana had exploded. Sockpuppet accounts of "local concerned citizens" and "eyewitnesses" started to document the horror by using the hashtag, leading to a deluge of misleading reports.

According to an investigative report by the New York Times, a user named @EricTraPPP tweeted to a New Orleans-based reporter named Heather Nolan, "Heather, I'm sure that the explosion at the #ColumbianChemicals is really dangerous. Louisiana is really screwed now."

In another example, sockpuppet accounts, also managed by Russian agents, used Instagram as the battleground to target African-American voters. A report, as summarized on ABC News, said that the IRA "created an expansive cross-platform media mirage targeting the Black community, which shared and cross-promoted authentic Black media to create an immersive influence ecosystem." [16] This network then exploited existing racial tension in America to sow discord among American voters.

A Guardian article highlighted this behavior in an interview with Theodore Johnson, a Brennan Center for Justice fellow. In the interview, Johnson said, 'Equally important was putting black activist language out on social media in order to scare white citizens into thinking their nation was changing,' pointing to posts that falsely showed Black Lives Matter activists with guns and claimed they planned to exercise their second-amendment rights." [17]

These were not grassroots efforts by independent actors, but hired agents of a political power. Within the current functionality of platforms like Twitter and Facebook, a mass-produced bot or puppet account can be indistinguishable from a legitimate user. Paid labor can now produce a chorus of voices tuned to the buyer's message.

With these tactics in mind, the 2016 election season brought online harassment tactics and behavior out of "niche" circles like tech, gaming, and activism and to the forefront of American politics, culture, and discourse.

Online Harassment and the Press

Perhaps the most prominent and easy-to-identify example of online harassment in the United States comes from the current administration's attacks on the press. The president has repeatedly referred to the media as the "enemy of the people," and made personal attacks on specific journalists—both live and on Twitter.

Attacks on journalists are nothing new, given the public nature of the profession and the regular arrest and harassment of journalists by governments worldwide. Yet it is an increasingly recognized and rapidly growing issue in the United States. In 2018, the U.S. was included in Reporters Without Borders' annual list of the most dangerous countries for press alongside the war-torn nations of Syria, Yemen, and Afghanistan. This is the first time the United States was included on the list since Reporters Without Borders began publishing it in 1995.

Five staffers were killed in a shooting at the Capital Gazette in Maryland in June 2018. The shooter had harassed the paper on Twitter and on its comment sections for years. [18]

President Trump's assault on the profession and his declarations that the media is "the enemy of the people" — a phrase he's used increasingly since May 2018 — has only served to legitimize such abuse directed at journalists.

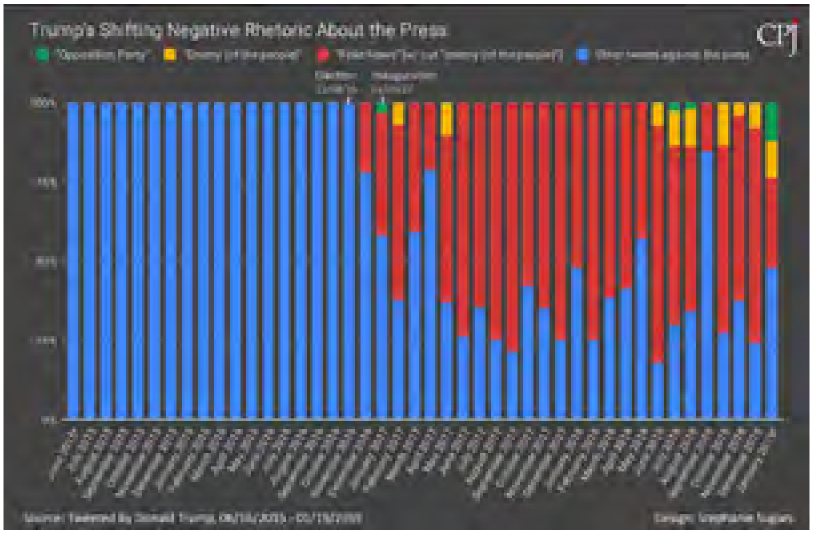


Figure 1.3, Committee to Protect Journalists, 2019.

Trump insulted individual journalists via Twitter 280 times as a candidate. CPJ has documented cases of several journalists who said after being targeted by him on Twitter they were harassed or doxxed.

A study by the Committee to Protect Journalists (CPJ) shows that the president's comments aren't without consequence (although he's claimed they're harmless, most recently in an exclusive interview with the New York Times' publisher A.G. Sulzberger.) [19] The United Nations has also taken notice of the potential for the president's action to increase the threat of violent attacks on journalists. [20]

> *Each time the President calls the media 'the enemy of the people' or fails to allow questions from reporters from disfavoured outlets he suggests nefarious motivations or animus.*
>
> — David Kaye, UN Special Rapporteur on freedom of expression and Edison Lanza, Inter-American Commission on Human Rights

That CPJ report analyzed the president's attacks on the news media and called out that, "Trump insulted individual journalists via Twitter 280 times as a candidate. CPJ has documented cases of several journalists who said after being targeted by him on Twitter they were harassed or doxxed." [21] This is just another demonstration that what happens online can lead to offline consequences.

The impacts of harassment on journalists are chilling:

- Deterioration of mental and physical health
- Reclusion from professional and personal pursuits
- Self-censoring
- Completely leaving online spaces, which may be critical to their work
- Avoiding writing about or covering certain topics, particularly ones that might be considered controversial (for example, immigration)
- Leaving home or seeking physical protection

All of this amounts to a dangerous decline in the diversity of voices in the public discourse, as well as a disturbing trend that online harassment can effectively meet abusers' goals: to scare, silence, and expel.

Despite this unnerving declaration by the president, journalists have always been targets for abuse. Journalists have regularly drawn ire for their work and most are open to discourse and criticism—they are, after all, public figures airing their words and reports for public consumption and assessment.

However, the move to digital media by publishers has increased pressure and demands on journalists, as well as avenues for their harassment. Many journalists are now compelled to engage with readers in online comments sections or to moderate these sections entirely, while also maintaining a public presence on social media.

Unfortunately, these arenas are often prime territory for people looking to harass a journalist rather than respectfully disagree with them. Reporters Without Borders came up with a helpful framework to understand the abuse and harassment that journalists face online[22]. The framework includes three stages of harassment of journalists: disinformation, amplification, and intimidation. Note that this relates to harassment outside of personal attacks that may be waged simply because of the topics a journalist covers, their physical appearance, or other traits.

As mentioned before, disinformation campaigns waged by the new wave of sophisticated and state-sponsored "trolls" employ classic online harassment strategies and tactics to intimidate, silence, and coerce individuals.

Challenges or threats to press freedom shut down important voices, viewpoints, and topics that should be part of the national discourse in healthy democracies. Journalists can be shut down and silenced because of online harassment tactics, from sustained vitriolic messages to death and rape threats.

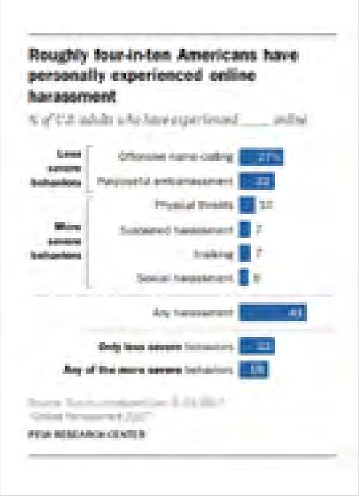


Figure 1.4, Pew Research, 2017.

Online Harassment by the Numbers

In 2017, Pew Research released another report about online harassment. Compared to the 2014 results, the 2017 numbers appeared similar on the surface. Sixty-six percent of Americans reported that they have witnessed online harassment (compared to 73 percent in 2014); 62 percent said online harassment is a major problem (not reported in 2014); and 40 percent have personally experienced online harassment (similar to 2014). [23]

What may need to be considered here, however, is that new norms in social communication can impact how online harassment is perceived and/or addressed (or not addressed). Harassing behavior can either be seen as normal, expected, or even encouraged and there is no indication that normalization was taken into account in the survey.

The true scope and depth of online harassment remains unclear. While large surveys conducted by Pew Research help, they don't account for underreported or unreported cases of online harassment.

Additionally, because online harassment happens through so many mediums and to so many people, many of whom do not report harassment, the scale at which harassment occurs may be grossly underestimated. Individuals might decide not to report harassment for a number of reasons including: fear of retribution or escalated harassment; feelings of guilt, shame, or distress; concern for physical safety, career, and reputation; or uncertainty about what can actually be done to stop the harassment.[24]

Aside from these large, nationally representative polls, other studies look at subsections of the population (for example, by demographic or profession). Such studies are more useful and in-depth than high-level studies, but make it difficult to discern the scale and impact of online harassment today.

Because toxicity online has become an accepted and almost expected part of online life, it has become even more difficult to measure the true scope of online harassment. What some people might see as simply an annoying or aggravated comment might be viewed by others as harassing behavior—particularly if a targeted individual has been receiving such comments for weeks, months, or even years.

The 2017 Pew Research report found that 27 percent of those surveyed weren't sure if they considered their most recent incident "online harassment." The report doesn't record why respondents aren't sure, but this may come from normalization, comparing one's experience to another's "worse" experience, and a lack of ways to describe the harassing behavior (see Part 3, Section 1 for more on this).

> *Most of these situations happened in game rooms or on forums where smack talk is common and no one takes it too seriously, but personal details are usually left out.*
>
> — anonymous respondent from the 2017 Pew Research survey

> *We all also got tons of rape threats constantly... The bad thing about that (aside from the obvious) is that it all ends up sort of bleeding together. You get so used to it that things stop standing out.* [25]
>
> — Kara Brown, podcaster and writer, interviewed for Mashable

Even more recent numbers, released by The Anti-Defamation League (ADL) in early 2019, show that 37% of Americans have experienced severe online harassment, including "sexual harassment, stalking, physical threats, and sustained harassment." As the ADL report's executive summary highlights, that "figure is substantially higher than the 18% reported to a comparable question," in the 2017 Pew Research poll. [26]

Those numbers are certainly concerning. But an even more disturbing takeaway is the disproportionate impact of online harassment on particular individuals or communities, including women, people of color, people who identify as LGBTQ+ and religious minorities, as well as individuals specifically targeted for their political views, activism, or profession. As writer and activist Bailey Poland points out in her book, *Haters: Harassment, Abuse, and Violence Online*, frequent and disproportionate attacks on underrepresented individuals are not adequately addressed in these studies. [27] As we'll see later in the report, more research is needed to develop an accurate understanding of the problem scope.

Heading into 2020

Advancing Technology

While abusive online behavior and content clearly isn't new, both technology and the expansiveness of online spaces has increased the ways an abuser can harass targets. Virtual reality, computer vision, deepfakes (the realistic doctoring of videos using image synthesis and AI), and quick meme creation and distribution make it easy for an abuser to create abusive content. Technology will get better and more accessible and harassers are bound to adapt.

> *Danny (not his real name) has stalked and harassed me, online and off, for almost 15 years — more than half my life at this point. He has used a variety of methods to do so — phone, text, email, Facebook and other social media — updating his tactics with every advance in technology.* [28]
>
> *— Roni Jacobson, writer*

Democracy and Plurality

It is clear that the same tactics used for harassing individuals are now being weaponized in large-scale, sophisticated attacks that undermine our democracy. As we learn more about the social and cultural power of online engagement, we also understand its use as a political and social organizing tool. This is creating rapid changes in accepted norms of online behavior that appear as—or are recognized as—harmful, but to what extent remains unknown.

For example, is it acceptable to dox someone violating the norms of an online community or of society? What about doxxing an avowed racist, or someone who makes racially-charged statements? That's what some vigilantes have been doing, as documented thoroughly in a report by ProPublica and, more recently, in The New Republic.

Tactics used for harassment of oppressed groups are now being used to punish those who violate a community's values and norms. Most people would agree that a racist or sexist aggressor using doxxing or dogpiling to silence a target based on their identity is wrong. Now the same tactics are used to punish racist or sexist behavior potentially fueled by the historic lack of consequences for such behavior.

As mentioned in The New Republic's investigation, Danielle Citron, the scholar and author of *Hate Crimes in Cyberspace*, told ProPublica in 2017, "This is a very bad strategy leading to a downward spiral of depravity. It provides a permission structure to go outside the law and punish each other." [29]

While these issues won't be resolved by 2020, they indicate a concerning normalization of harassment as an effective tool for ideological warfare. More strategies, tactics, and avenues exist today than ever to weaponize the internet, its current infrastructure, and the people who use it.

Popular Culture and Society

While online harassment and its relation to pop culture warrants a much deeper exploration by culture and media critics, journalists, and researchers than we can offer here, it's worth noting that new norms in social communication can impact how our society perceives online harassment and how we address it (if at all). Harassing behavior can either be seen as normal, expected, or even encouraged.

One such example is communication norms in online gaming (where players interact with one another during live play) and communication norms on social media (both on incumbent platforms and in quickly growing upstarts like TikTok and Snapchat). In the former example, female players have posed as men to avoid harassment. And, in a strange twist, men have posed as females to harass other gamers. [30]

In the latter, online harassment was used as a marketing tactic that helped at least one (now infamous) rap star gain fame and notoriety. His preferred behaviors were trolling and commenting with threats of violence on people's Instagram posts. [31] In an article for Rolling Stone, journalist Stephen Witt notes, "He'd seen the response to his social media provocations, and he sensed that this could be leveraged to build a larger audience…He'd find someone on social media, leave nasty comments, and dare them to fight him…he wasn't threatening in real life." This echoes classic rebuttals by abusers that they were "only joking" or "didn't really mean it," which relieves them of responsibility for bad behavior.

Another cultural phenomenon to consider moving forward is the shift in how people seek out and find intimate relationships. People using dating apps and websites, like Tinder, Grindr, and Hinge, can easily be harassed, taken advantage of, or approached by impersonators. In 2018, one Grindr user filed a lawsuit against the company, *Herrick vs Grindr LLC*, after a former partner created and used fake profiles to harass him for months. [32] Made famous by Gamergate (although by no means the first example of this), cases of intimate partner abuse today almost always include digital facets.

In some cases, harassment is used as entertainment among people who may not recognize the behavior as inappropriate. At the end of 2017, two gamers arguing over a $1.50 bet escalated to a SWAT threat (swatting or SWATing refers to sending police to a target's home). A third gamer, Tyler Barris, known for his SWATing prowess, was commissioned to SWAT one of the gamers involved in the bet. The SWAT targeted one of the gamers' old addresses, whose current resident, an innocent bystander named Andrew Finch, was shot and killed by officers in Wichita, Kansas. Barris had been to jail for a wave of bomb threats in 2015 but was released two years later on good behavior, and also built a business on his stalking and SWATing expertise. [33] Finch had no connection to Barris or any other gamers involved.

When being interviewed for a 2018 WIRED magazine article, Barris said, "[SWATing] was like a kind of online power. Knowing that you're breaking the law, and knowing that they won't be able to find you, and knowing you just sent the SWAT team or bomb squad somewhere, and knowing you could do that over and over again."

While the Wichita SWATting case was extreme, SWATting is not uncommon, particularly in certain circles. Yet, it has the potential for deadly consequences.

Even when news sites report the facts about incidents of doxxing, SWATting, non-consensual pornography, and other forms of online harassment, they may be unintentionally promoting the behaviors. It has been found that media coverage of suicides [34] and school shootings [35] can lead to short-term increases in similar incidents. The same could be possible for extreme forms of online harassment.

Online harassment is also normalized in popular culture in other ways: While most media outlets stop short of explicitly celebrating harassment, websites such as Buzzfeed [36] and Thought Catalog [37] have published lists of the "best" online insults and "burns." Although these can be construed as comments simply made in bad taste, they may still desensitize or encourage other harassing behaviors.

Summary

- More than 85 million Americans have experienced online harassment.
- Anyone is subject to harassment. It has affected private citizens and public figures; everyday Americans and CEOs of Fortune 500 Companies.
- The same tactics used for harassing individuals are being used in large-scale, sophisticated, and state-sponsored attacks and disinformation campaigns.
- The problem of online harassment is growing in scope and intensity, rather than receding.
- Learning from previous real-life examples of online harassment can help identify patterns of behavior and create better outcomes in the future.
- Harassing behavior may become more normalized and taken less seriously. We should consider this as we seek effective, proactive, and preventative solutions to online harassment.

The Four Levels of "Trolling"

Level 1 —
Trolling for the Lulz

A rock is thrown into a pond to see the ripples and disturb an ecosystem.

Level 2 —
Trolling for Individual Attacks

A rock is thrown directly at an individual, intend to harm or inflict pain on the individual.

Level 3 —
Trolling for Ideological Attacks

Rocks are thrown at individuals based on their identity or ideologies.

Level 4 —
Trolling for Large-Scale Manipulation

A powerful individual or group mobilizes mercenaries to throw rocks with the aim of causing fear and mistrust by their bombardments.

Figure 1.5, The 4 Levels of "Trolling", OnlineSOS, 2019.

For a deep read on the history of online harassment and the various cases, movements, and literature that have defined it up until this point, please reference:

- An Incomplete—But Growing—History of Online Harassment since 2003
  by Caroline Sinders
  (last updated November 2018)
- Haters: Harassment, Abuse, and Violence Online by Bailey Poland
  (published 2016)
- The Internet of Garbage by Sarah Jeong
  (published 2015, updated 2018)
- Hate Crimes in Cyberspace by Danielle Citron
  (published in 2014)

Part 2

The Individual Experience and Responses

In Part 2, we'll provide an overview of the current ecosystem of resources available to individuals facing online harassment through the lens of how people experience online harassment. This includes:

- What the individual experience of online harassment is like
- What needs and options individuals have when facing online harassment
- Current online harassment research trends and challenges
- Current legislation or policy trends and challenges
- A 5-Point Solutions Framework to help guide the development of future individual-centered online harassment interventions or responses

Experiencing Harassment

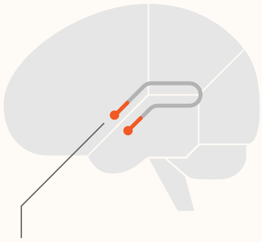
When an individual experiences online harassment, the experience can typically cause a stress response colloquially known as going into "fight or flight mode." A person may experience symptoms not unlike post-traumatic stress disorder (PTSD).[38] For example, their ability to react, think, or clearly communicate may be impaired. This is normal when a person is under emotional distress because of structural and chemical changes in the brain as outlined in the exhibit below. Experiencing harassment can feel overwhelming, crushing, or inescapable. These feelings may be compounded for individuals who have been assaulted or who have experienced other trauma in their past.

Trauma and the Brain

Psychological Trauma

Unique, individual experience of an event or enduring conditions, in which:

- The individual's ability to integrate his/her emotional experience is **overwhelmed**, or
- The **individual experiences** (subjectively) a threat to life, bodily integrity,or sanity.



Structural changes to the brain that impact ability to respond

▼ Decrease hippocampus size (memories) creates difficulty remembering or recalling details

▼ Decrease in VPC volume (emotional regulation) creates difficulty regulating emotions

▲ Increase in activity in amygdala (emotions) create hyperarousal, possibly leading to symptoms of anxiety or panic

Source: Pearlman & Saakvitne, 1995

Figure 2.1A, OnlineSOS, 2017.

> *Each time I reported a post and had it deleted, another one appeared. I began to fear for my safety every time I went to the grocery store, refueled my car at the gas station, or went hiking. I prayed I wouldn't be assaulted every time I left my house.*
>
> *— Rebecca Sheffler (pseudonym)*[39]

The Role of Emotional Distress

A major barrier to receiving adequate support: Emotional distress can hinder an individual's ability to communicate effectively

What We've Observed

In a state of emotional escalation or trauma, individuals have challenges in communication:

1. Explaining concisely
2. Describing a coherent timeline with relevant details
3. Supplying proper documentation

Evidence

This observation is consistent with:

1. Psychological research
2. Feedback from platforms (re: user reports missing information or context)
3. Feedback from service providers, like lawyers and law enforcement

Mitigants

1. Receive help with documentation from supports
2. Seek structured forms or ways to document
3. Create document with incident details that can be easily shared (upfront work for more effective outcomes)

Figure 2.1B, OnlineSOS, 2017.

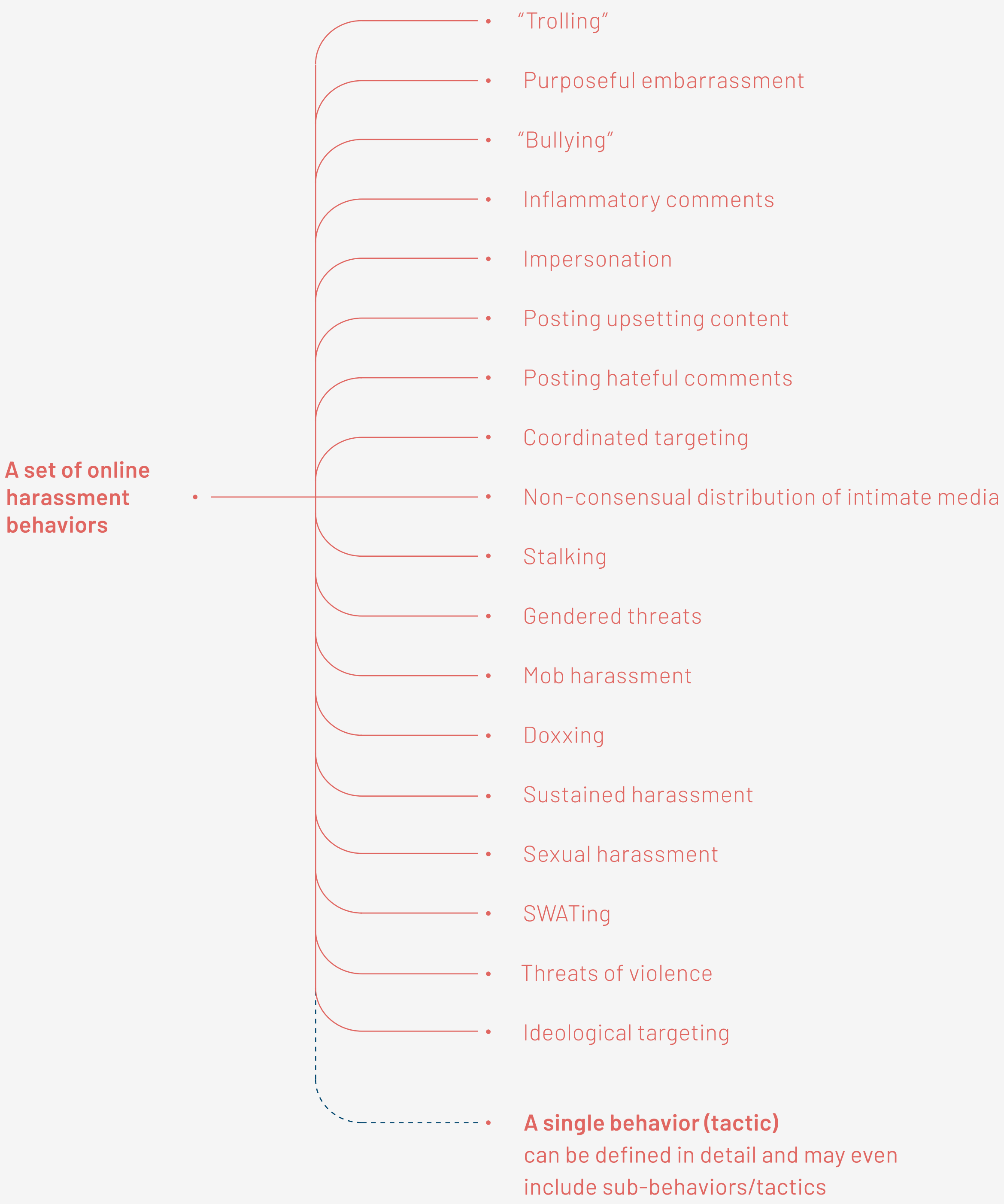
> The potential experience of psychological trauma and emotional distress can also directly impact one's ability to effectively seek and obtain the help they need.

Behaviors and Tactics

Part 1

Online harassment is an umbrella term that refers to a set of specific, damaging behaviors and tactics. People may experience more than one tactic at any given time.

Online Harassment is a Set of Harmful Behaviors



A set of online harassment behaviors

- "Trolling"
- Purposeful embarrassment
- "Bullying"
- Inflammatory comments
- Impersonation
- Posting upsetting content
- Posting hateful comments
- Coordinated targeting
- Non-consensual distribution of intimate media
- Stalking
- Gendered threats
- Mob harassment
- Doxxing
- Sustained harassment
- Sexual harassment
- SWATing
- Threats of violence
- Ideological targeting

A single behavior (tactic) can be defined in detail and may even include sub-behaviors/tactics

Part 2

Tactics can be grouped into subsets or even broken down into more specific "sub-tactics." The following table, Figure 2.3, is one example of how tactics might be organized. In this example, we grouped tactics by how the harassing content is often presented:

1. In text (e.g. described in writing);
2. Through multimedia (e.g. presented in images or video);
3. By operation (e.g. a specific action taken that may trigger subsequent behaviors).

Note that the table is not an exhaustive list of tactics, but rather one for demonstrative purposes.

Tactics by Medium: Text, Multimedia, and Operation

Text	Multimedia	Operation
Behaviors that mostly involve written communication to a target or posted about the target.	Behaviors that mostly involve images or videos delivered to the target or posted about the target.	Behaviors that may be succeeded by other harassing behaviors.
• "Trolling" • "Bullying" • Inflammatory comments • Hate speech • Threats of violence (wishing or describing) • Violence explicitly directed towards target • Doxxing (location, identity, or other personal information) • Implicit or subtle threats • Defamation • Sharing falsehoods • Extortion	• Non-consensual images or videos • Superimposing photos, often with threatening imagery • Doctored images or videos • Filming of sexual assault • Deep fakes • Defamation, threats spoken in video	• Impersonation • Sock puppeting (creation of fake accounts with intent to deceive) • Multiple account creation for coordinated mob harassment • Astroturfing (deceptive presentation of views as grassroots or organic) or staging for coordinating mob harassment • DDoS • SWATing • Cyberstalking • Falsely claiming copyright infringement on text, images, or video

Figure 2.3, Tactics by Medium: Text, Multimedia, and Operation, OnlineSOS, 2019.

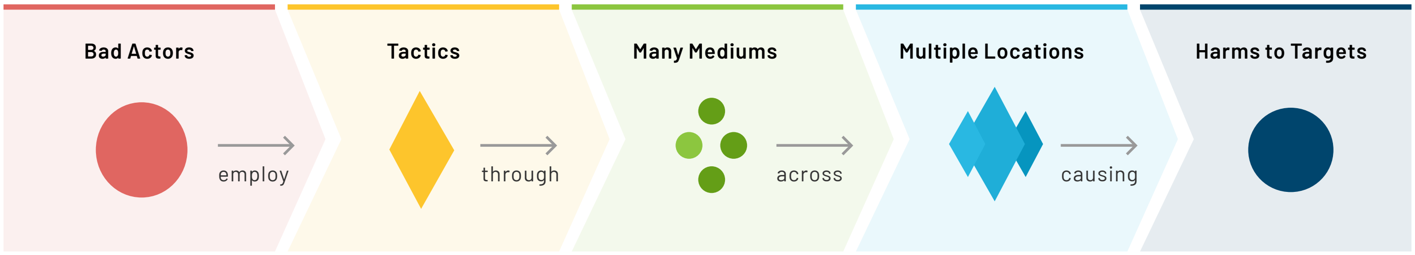
What's important to note is that targets of online harassment often deal with an onslaught of behaviors—rather than, say, one comment at a time—which naturally complicates the experience and, therefore, the response.

The online harassment experience usually exhibits one or more of the following characteristics:

1. More than one tactic at a time
2. Tactics employed across multiple channels
3. Harassment goes deeper than what is in public view or publicly shared
4. Escalation of harassment that further blurs the lines between online and offline life

Figure 2.4 demonstrates where, how, by whom, and the harm done to targets. Note that this too is just a demonstrative, rather than exhaustive list. Given the four characteristics listed above, the number of harassing interactions a targeted individual may experience is staggering. Additionally, these combinations grow exponentially for someone targeted for coordinated mob harassment.

The Components of an Online Harassment Experience



Bad Actors

- Anonymous person
- Colleague
- Ex-colleague
- Ex-friend
- Ex-partner
- Partner
- Estranged family
- Fan
- Friend or family
- Group - mob
- Ideologue
- Paid "troll"
- Pseudo-anonymous
- Reader
- Stalker

Tactics

- Bullying
- Coordinated targeting
- Doxxing
- Hateful comments
- Gendered threats
- Ideological targeting
- Impersonation
- Inflammatory comments
- Mob harassment
- Non-consensual porn
- Sexual harassment
- Stalking
- Sustained harassment
- SWATing
- Threats of violence
- Trolling
- Purposeful embarrassment
- Upsetting content

Many Mediums

- Blog post
- Comment
- Direct message
- Email
- Image
- Meme
- Social media post
- Text message
- Venmo payment
- Video

Multiple Locations

Social Media
- Facebook
- Twitter
- Instagram
- Snapchat

Message Boards
- Reddit
- 4chan
- 8chan
- WeChat
- Discord

Personal Media
- Email
- WhatsApp
- SMS/Texting
- Direct Messages

Public-facing Media
- Comment sections
- YouTube
- Google
- Wikimedia/pedia

Niche Platforms
- Twitch
- Craigslist
- Tinder
- Grindr

Harms to Targets

- Anxiety
- Capitulation
- Damaged reputation
- Death
- Discord
- Exclusion
- Fear
- Isolation
- Physical violence
- Silencing
- Trauma

Figure 2.4, The Components of an Online Harassment Experience, OnlineSOS, 2019.

Part 3

Because harassment often takes place over a period of time, the experience is made up of many combinations and permutations (or cycles) of bad actors, tactics, mediums, locations, and harms.

Therefore, it's critical that an individual is presented with options that reflect their unique situation. Online harassment is personal and, while there are some best practices (documentation of the harassment, for example), there's no single, correct way to respond.

> Because harassment often takes place over a period of time, the experience is made up of many combinations and permutations (or cycles) of bad actors, tactics, mediums, locations, and harms.

Needs and Options

In addition to tactical options that address their specific incident or concerns, people experiencing online harassment have three key fundamental needs:

1. **Physical safety** — guaranteeing the safety of oneself and one's family
2. **Emotional and psychological well-being** — managing the emotional impact of online harassment, including uncertainty and anxiety
3. **Digital security** — securing or managing online accounts to minimize risk of further exposure or harm

To address these needs, an individual commonly takes these three steps (to varying degrees):

1. **Conducts a threat assessment:** Whether an individual does this consciously (e.g. follows a guide, like the one we provide on OnlineSOS) or not, a person evaluates their risk and threat(s) to create a plan of action.
2. **Documents the harassment:** This includes saving messages, posts, comments, and other harassing content, typically to report to platforms or in case of legal action.
3. **Communications with others:** This can include written and verbal communication with social media platforms, software providers, law enforcement, friends and family, employers, or other support organizations.

There are several common decisions an individual also has to make, including:

- **Respond** to the abuser?
- **Delete** or remove the content?
- **Report** the content / behavior and, if yes, to whom?

The table below looks at some specific target needs, resources or groups that may be able to support them in one way or another and, finally, potential action steps they can take. The prioritization and sequencing of these actions can change depending on the individual's unique situation or past experiences.

In the Moment: Individual Needs, Options, and Resources

Needs	Options
🔵 Secure physical safety	• Contact law enforcement • Get a restraining order
🔵 Assess threat(s)	• Go through a threat assessment (either DIY with the support of a peer or through a private investigator, law enforcement, or domestic violence expert)
🟠 Reduce or eliminate the harassment or source of harassment ASAP	• Block, mute, or filter out the abuser • Submit a takedown request • Report the abuse to the platform(s) where the abuse is taking place • Secure accounts / digital presence to minimize risk of further exposure or harm
🟠 Answers to questions about how to deal with the situation they're facing	• Get support from a trained peer, crisis counselor or expert who has responded to other online harassment cases • Get support or answers from someone who has experienced online harassment
🔴 Document the harassment	• Follow a DIY guide for best practices or get the support of a peer, friend, or other individual
🟠 Manage the harassment to reduce trauma or re-trauma	• Ask a peer, friend, or other individual to act as a barrier between the harasser and the target
🟠 Validate and describe their experience	• Talk to an objective third party • Access taxonomy to accurately describe the harassing behavior experienced
🔴 Assess if they have a legal case to pursue action	• Contact a lawyer, legal counsel, or expert with experience in online harassment

Resources or Groups That Can Help

Law enforcement

Legal counsel / aid

Digital security assistance

Physical security assistance

Employers / HR

Social media and other internet platforms

Friends / family / community supports

Clinicians / therapists / psychologists

Crisis lines / health institutions

PR / communications

Third party resources and tools

🔵 **Physical safety**

🟠 **Emotional and psychological well-being**

🔴 **Digital security**

Figure 2.5, In the Moment: Individual Needs, Resources, and Options, OnlineSOS, 2019.

In informational interviews with journalists and activists that the OnlineSOS team conducted in 2018, respondents indicated that they often did not know how to proceed when they were being harassed. In some cases, actions they thought would be effective ended up causing more harm than good. One respondent mentioned that it would have been helpful to understand that there are a range of choices someone has when facing harassment, including whether or not to confront their harasser based on the particular situation. This respondent indicated that it would have been especially helpful to understand that confrontation is not always the best option, saying, "I thought I could engage by defending myself, but that only made it worse."

People need clear, understandable resources at their fingertips to help them through the process of dealing with online harassment. We also need to contend with the reality that completely disengaging from online spaces is not an option. This is especially the case for individuals whose job relies on being online.

> *The advice out there isn't great. Some people say that you should just delete all your accounts and get offline as if this is your own fault. But like I said, our livelihoods are so online dependent now that that's just not a solution. It's almost impossible to report this stuff to social network admins because they are dealing with a huge amount of reports of abuse everyday. They minimize and trivialize your issue.*
>
> *— Eliza Romero, pop culture writer*[40]

Below we'll look at **four key options** available to targets of online harassment:

1. **DIY Resources**
2. **Support from Family and Friends**
3. **Civil Society Initiatives**
4. **Social Media Platform Mechanisms**

1. DIY Resources

In light of Gamergate, resources for targets of online harassment flourished. Many organizations have developed excellent DIY resources and information for targets of online harassment. These include in-depth or step-by-step guides with varying angles or demographic focuses. Here is a sampling of organizations:

DIY Resources by Audience

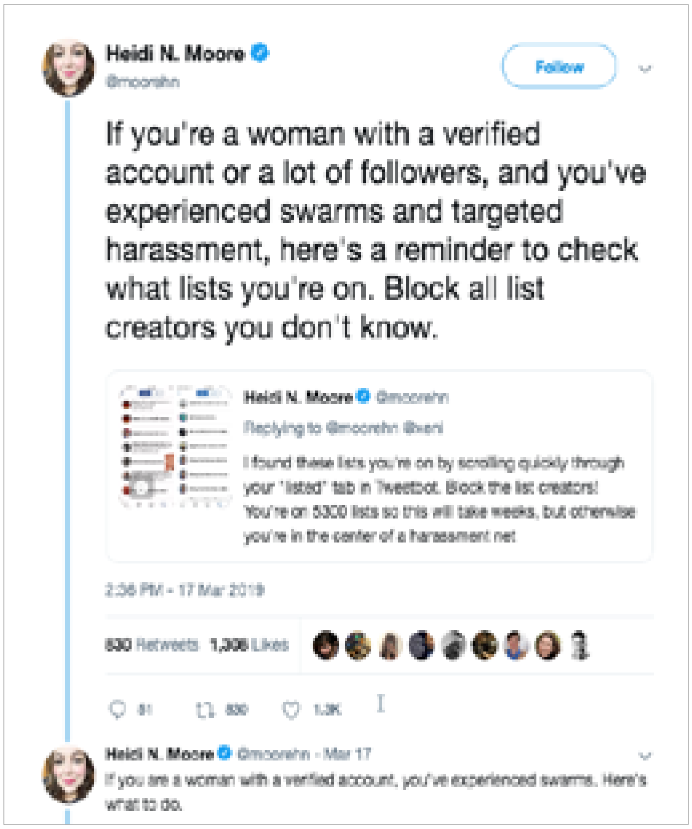
General public	Women	Targets of non-consensual intimate media	Journalists
• Crash Override • Civilination • Equality Labs • Heartmob • National Network to End Domestic Violence • OnlineSOS	• Feminist Frequency • Hack Blossom • Stop Online Violence Against Women	• Cyber Civil Rights Initiative • Without My Consent	• Committee to Protect Journalists • PEN America • Trollbusters

Figure 2.6, OnlineSOS, 2019.

These are just a few of the organizations that have put out guides or resources about online harassment. (It's also notable that in light of the growing number of individuals, groups, or communities targeted for online harassment, many organizations create online harassment resources specific to their cause sector and constituency.) Especially under distress, options and clear next steps are key. However, in many cases, people don't know where to turn to for help and are unaware that DIY resources, information, or services even exist. In addition, even if these sources are presented and easily accessible, the amount of information can be overwhelming to navigate.

It's important to note that while many resources do exist, there is no guarantee that they can be adequately maintained, updated, or expanded. It's common that DIY resources / grassroots groups become inactive because of scarce resources or founder and staff burnout.

Because of this, targeted individuals, especially among communities that are heavily targeted, may rely on backchannel networks or informal peer structures for best practices.



Heidi N. Moore @moorehn

If you're a woman with a verified account or a lot of followers, and you've experienced swarms and targeted harassment, here's a reminder to check what lists you're on. Block all list creators you don't know.

Heidi N. Moore @moorehn

Replying to @moorehn @xeni

I found these lists you're on by scrolling quickly through your "listed" tab in Tweetbot. Block the list creators! You're on 5300 lists so this will take weeks, but otherwise you're in the center of a harassment net

2:36 PM - 17 Mar 2019

830 Retweets 1,306 Likes

Heidi N. Moore @moorehn · Mar 17

If you are a woman with a verified account, you've experienced swarms. Here's what to do.

Figure 2.7, Twitter, 2019.

Full thread: https://twitter.com/moorehn/status/1107395392386277377

Although these resources and tips can be helpful, access remains a challenge. Not everyone has a peer or colleague to turn to when faced with online harassment. (In media, for example, staff writers have decreased significantly in the last decade.) In particular, this leaves freelance writers and journalists (with few resources and no support mechanisms from colleagues or an employer) acutely vulnerable.

> Promoting the availability or visibility of existing resources, whether through awareness campaigns or strong referral networks between entities, could help connect people to the resources they need when they need it most.

2. Support from Family and Friends

In our 2018 informational interviews, respondents often sought support from family and friends when faced with online harassment. The Pew Research survey from 2017 also reported that, of those respondents who were harassed online, 29 percent turned to friends and family.[41] Friends and family can provide emotional support or even help a target of online harassment filter out or moderate incoming messages to help minimize exposure to more harassment.

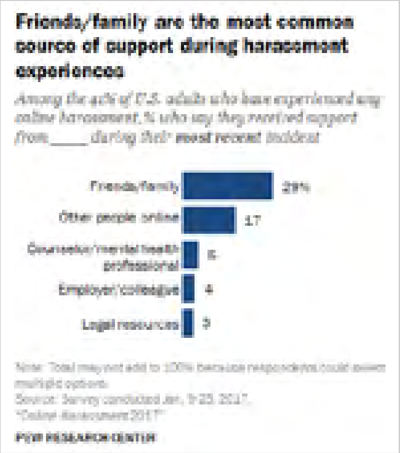


Figure 2.8, Pew Research, 2017.

Allyship has also become the basis for new proposed methods of intervention. PEN America's Online Harassment Field Manual includes a section on how to be a good ally or witness, including tips on how to offer support to someone experiencing abuse. [42]

> *When you witness another person being targeted by online hate or harassment, it can be intimidating to intervene. What if you inadvertently make the harassment worse? What if you become the target of such harassment yourself? What if the harassment you've witnessed is traumatizing to you in some way, making it difficult for you to remain in the same online space where the incident occurred?*
>
> — PEN America Online Harassment Field Manual

A research group from MIT productized the "friends-as-moderators" approach and developed "Squadbox," an inbox that let's people "put a squad of trusted friends, volunteers, or paid moderators between the world and your inbox." The trusted "squad" helps targets of online harassment feel less overwhelmed by it and filter out unwanted contact. [43]

Although people need expert support, from self-care strategies to mental health services to legal help, it may be difficult to access those resources.

It's important to note that some groups also have more access to resources than others. For example, if someone does not have health insurance, they may not be able to get the professional mental health services they need. Or, if someone cannot afford to get a hotel room for the night, move, or stay with friends or family when their physical safety has been threatened, then they are forced to risk their safety solely based on the resources available to them.

3. Civil Society Initiatives

There are two primary types of initiatives that are dedicated and designed to address online harassment:

1. Standalone upstart groups
2. Initiatives within existing nonprofit or for-profit organizations.

Standalone upstart groups can be categorized primarily as grassroots organizations, serving or dedicated to specific communities and heavily founder-driven. Most of these groups were started by (female) founders who have personally experienced online harassment.

Their organizations were born out of a desire to create the resources they wish they'd had when facing online harassment. These include:

- Civilination (founded by Andrea Weckerle)
- Crash Override (founded by Zoe Quinn)
- Cyber Civil Rights Initiative (founded by Holly Jacobs)
- OnlineSOS (co-founded by Liz Lee and Samantha Silverberg)
- Stop Online Violence Against Women (founded by Shireen Mitchell)
- Trollbusters (founded by Michelle Ferrier)
- Without My Consent (co-founded by Erica Johnstone and Colette Vogele)

Examples of dedicated initiatives within existing organizations include:

- Women Media Center's Online Safety Speech Project
- PEN America's Online Harassment Field Manual
- Hollaback's Heartmob
- Feminist Frequency's Online Safety Guide
- Anti-Defamation League's Online Harassment Survey (2019)
- Amnesty International's Toxic Twitter Report (2018)
- K&L Gates' Cyber Civil Rights Legal Initiative

Efforts within existing organizations tend to be campaign and awareness driven. Such groups are likely to conduct landscape or ecosystem research like this one, often interviewing targets or peers, writing literature reviews, and holding strategy sessions to identify how their organization is best positioned to contribute.

▶▶▶ *For a more complete list of organizations, what they offer, and their current status, refer to Appendix A.*

Direct or Active Services

While many organizations offer do-it-yourself resources and information, direct services for targets of online harassment are harder to find. Several organizations have pathways for requesting immediate or near-immediate help.

The Cyber Civil Rights Initiative has a crisis helpline for targets of non-consensual pornography. Online harassment is just one of the issues addressed by Equality Labs, but they offer a Rapid Response Request Form through which individuals and organizations can seek help with hacking, doxxing, or digital defense training. With a focus on serving female journalists, Trollbusters provides an incident report form, where targets or bystanders can report harassment. Although other organizations advertise a chatbot or helpline, some of these features were malfunctioning or unavailable.

Another approach to direct services for targets of online harassment is to create and nurture supportive communities. The organization BADASS advises targets to contact them to be added to a private Facebook group of people who have faced harassment via non-consensual pornography. HeartMob created its own online platform through which targets can ask for specific kinds of help and support. Heartmobbers, the volunteers who respond and offer help, are trained by the organization.

In addition to services specific to online harassment, other sources may provide direct services to address certain dimensions of the issue.

While these resources were not specifically designed to support individuals experiencing online harassment, they can offer meaningful support. For example, crisis lines (like Crisis Text Line) can provide emotional support, while legal aid can help someone understand the options for filing a restraining order.

The degree to which these groups understand online harassment or are trained to respond to the specific behaviors and effects of it, however, will vary. This can also depend on the agency, party, or official that happens to be available.

These groups include:

- Mental health or psychological services, including crisis management
- Digital / cybersecurity services
- Legal support
- PR / crisis communications
- Employers and HR departments
- Peer support
- Law enforcement

For example, in some cases, law enforcement might be sympathetic to targets of harassment. At other times, however, an officer may not understand what Twitter is or how it works. There may also be gaps in understanding context, and therefore an inability or unwillingness to determine severity and urgency. Law enforcement, in particular, may not do anything unless a threat is both explicit and specific (e.g. on this day and at this time and at this place). Some targets may be victim-blamed as a result, adding to their distress or confusion.

> *The first police officer I spoke with...while concerned and well-intentioned, told me he was not familiar with these types of cases and advised that I "just ignore it" and "keep reporting it to Twitter". He briefly left the room to speak with CMPD cyber crimes and when he returned, he said that cyber crimes would not investigate the case for online impersonation because no money had been stolen from me. Also, because I did not know the exact identities of my stalker(s), he discouraged me from filing a full police report and only issued me an incident number.*
>
> — Jaclyn Brzezinski, writing about her experience reporting online harassment and cyberstalking to her local police department [44]

In addition to general purpose sources of support, on the opposite end of the spectrum there are also direct services that serve a specific demographic group and with a specific remedy. Access Now has a digital security email helpline offering direct technical assistance to civil society groups and activists, media organizations, journalists and bloggers, and human rights defenders. While Access Now's focus is on security risks and needs, there are tactics of online harassment that may be in scope for which Access Now can provide support.[45]

In Focus: Solutions for Journalists

As mentioned earlier in the report, the online harassment of journalists is not new. Journalists have been threatened and targeted online in various ways before the rise of social media and digital media. Social media and comment sections simply provided harassers more access to journalists.

As part of their job, journalists are expected to share their work online and engage with readers on social media or in the comment sections of their articles. Communication from readers used to come in the mail, but it now comes directly to your email inbox, a notification on your phone's home screen, or is posted online for everyone to see. This puts journalists at higher risk of exposure to not only hateful content, but also to threats and other tactics of online harassment.

Online harassment disrupts the personal and professional lives of journalists, but also impacts the stories that get covered and the ability for a journalist to cultivate important sources for a story. This changes the composition of newsrooms and the media landscape, including the people and stories that get covered, which in turn affects what reaches public consciousness.

Michelle Ferrier of TrollBusters, as well as the International Women's Media Foundation, and the Committee to Protect Journalists have led the charge to raise awareness, conduct research, and offer resources to journalists.

Since 2015, other organizations like the Anti-Defamation League, Access Now, Amnesty International, Reporters Without Borders, the Freedom Defense Fund, First Draft, the Knight Foundation, the International Journalists' Network and more have become increasingly proactive in addressing online harassment. This is a testament to the increasing threat online harassment poses to journalists both in the United States and abroad. While this growing activity is welcomed and important, it's also critical to come together and form a stronger, louder voice against the harassment of journalists.

There are two critical points to note about journalism in relation to online harassment:

1. Journalists have little incentive to share to their newsroom what's going on and newsrooms are not well-equipped to understand or support their journalists in the face of harassment.
2. As engaging online becomes an implicit and required part of a journalist's profession, what responsibility and/or potential liability does an employer or media organization have in providing support as part of employment? This is an important area of focus for the law.

The most vulnerable population of journalists is freelancers who may not have the organizational support of news organizations, as well as journalists from marginalized groups (including women, people of color, LGBTQ+). This includes younger journalists or new journalists entering the profession who may not be used to scrutiny or exposure to the public.

This is an increasing concern for journalism schools as well as newsroom management and HR departments. As waves of layoffs have indicated in the last couple decades—along with low pay, long hours, and sometimes dangerous working conditions—journalism is a challenging profession. Online harassment exacerbates this and is already changing whether people decide to join or stay in the profession.

As funders who support the journalism ecosystem look to address misinformation, disinformation, and media manipulation, online harassment is a related (but distinct) cause area that needs attention and funding.

Based on our experience developing a program to support journalists in the U.S., it's important to design specifically for their particular needs. Although they're not exactly a public figure in the way that, say, a politician is, they face much of the same scrutiny, exposure, and danger and therefore have unique considerations in security, their safety, or responding to harassment.

4. Social Media and Social Networking Platforms

Although average daily usage statistics can vary by study, Americans spend more time on social media than ever.[46] Globally, hours of average digital media consumption has more than doubled to 5.9 hours since 2008, with mobile share of that time growing as well.[47] [48]

According to a 2018 Pew Research survey, the majority of Americans use social media platforms regularly. It's also worth noting that certain platforms have become more prominent in recent years (for example, Instagram usage has jumped 7 percentage points since 2016) and some platforms are more popular with certain demographics (78% of 18 to 24 year-olds use Snapchat).

Growth has been explosive in apps specific to certain niche audiences too. For example, Twitch, a live streaming web app popular with gamers, nearly doubled its total minutes watched from 290 billion in 2016 to 560 billion in 2018, while its users send billions of chat messages a year.[49] Not surprisingly, the growth has coincided with increased scrutiny of online harassment on these platforms.[50]

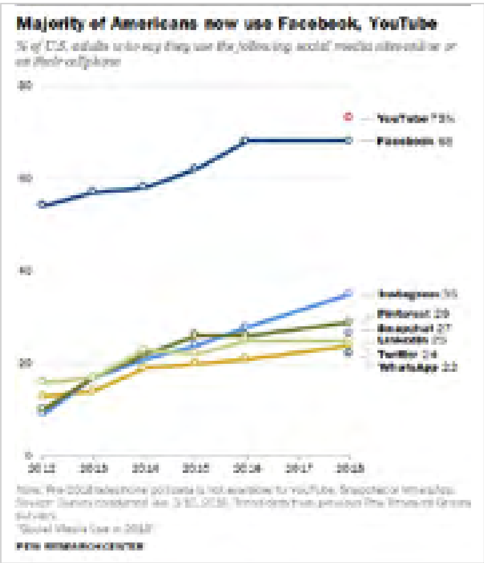


Figure 2.9, Pew Research, 2018.

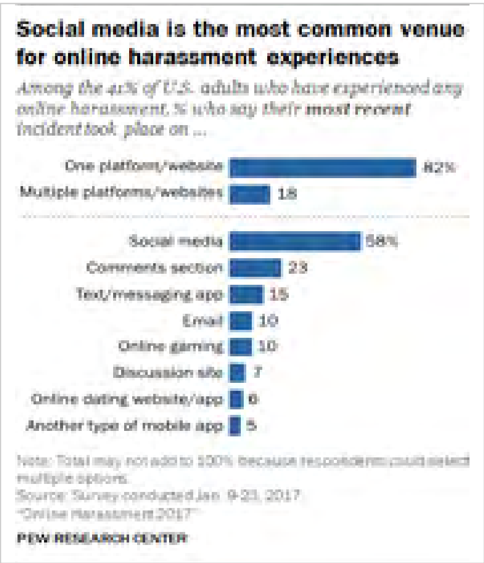


Figure 2.10, Pew Research, 2017.

Given the growth of social networking, it's also not surprising that, according to Pew, online harassment experiences commonly take place there.[51]

Blocking, Reporting, and Privacy Settings

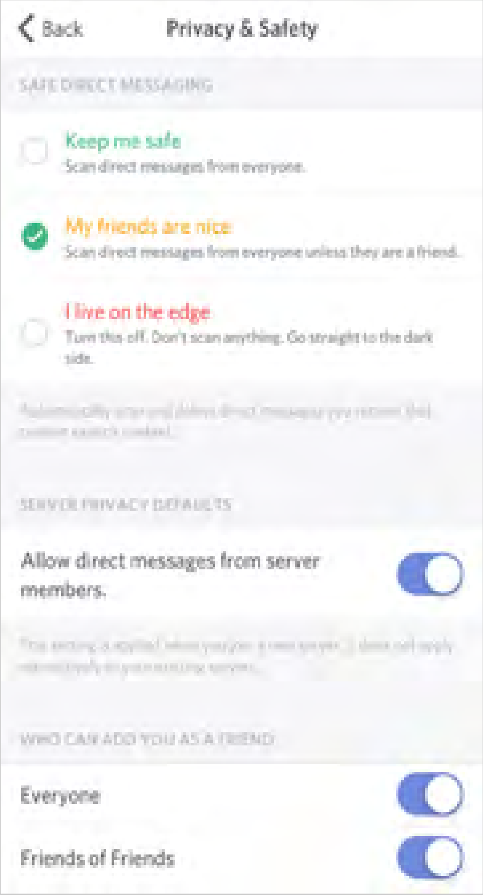
Social media companies have responded to online harassment by rolling out reactive product features like blocking, muting, filtering, reporting, or the ability to adjust privacy settings. The usability and efficacy of these features vary widely by platform. But targets of online harassment have generally expressed frustration at a) lack of responsiveness or transparency from platforms; b) inconsistent enforcement of policies or community standards; and c) the ineffectiveness of the provided tools to reduce / eliminate harassment.

> *Twitter, I find, don't remove anything. I think I've maybe managed, out of reporting probably over 100 posts to Twitter, I think they've removed two — one was a threat and the other had a pornographic image. There are 50 or 60 posts where I have specifically explained the tweets were malicious and the person that I believe is sending them has already been criminally convicted [for harassment] but they won't remove any of the tweets.*[52]
>
> — Respondent in an Amnesty International study about Twitter and the online abuse of women

Blocking, muting, and filtering are also not fool-proof and each comes with concerns of their own. When Twitter, in 2013, introduced the Mute button—a feature that allows people to hide certain users from their timelines without blocking them—one user expressed concern in the comment section of The Verge, which reported on the story:

> *This is very bad from a privacy/stalking perspective. I have a public account but there still may be people I don't want following me (and yes I know of the easy workarounds).* [53]

Muting a user, unlike blocking, still allows the muted user to interact with you (including through direct messages), but doesn't let them know they were muted. However, blocking does let a user know they've been blocked. This created (and remains) a catch-22 for people who don't want a harasser or stalker to know that they have been blocked—which could result in sock puppet accounts and escalate harassment—but also want to make sure they do not receive unwanted contact from a harasser. [54] In other words, neither is a positive outcome for a target of online harassment.



SAFE DIRECT MESSAGING

○ Keep me safe
Scan direct messages from everyone.

● My friends are nice
Scan direct messages from everyone unless they are a friend.

○ I live on the edge
Turn this off. Don't scan anything. Go straight to the dark side.

Automatically scan and delete direct messages you receive that contain explicit content.

SERVER PRIVACY DEFAULTS

Allow direct messages from server members. ●

This setting is applied when you join a new server. It does not apply retroactively to your existing servers.

WHO CAN ADD YOU AS A FRIEND

Everyone ●

Friends of Friends ●

Figure 2.11, Discord, 2019.

As late as February 2019, people still find filtering and reporting tools insufficient. The ADL's 2019 online harassment report states, "Americans also want to see private technology companies take action to counter or mitigate online hate and harassment, with 84% saying that platforms should do more. They want platforms to make it easier for users to filter (81%) and report (76%) hateful and harassing content." [55]

Privacy settings, although not a direct solution to online harassment, can help users protect themselves from future threats. Digital security is one of the few preventative measures currently available to individuals. Google's various products, like Gmail, regularly prompt users at sign in to conduct a privacy checkup, including enabling two-factor authentication or double checking what apps have access to sensitive information. [56] Facebook and Discord, a free text and voice app for gamers, have also started to introduce similar proactive privacy checkup prompts.

This can help people stay up-to-date with cybersecurity best practices that can at least minimize a harasser's ability to reach them or access their private information.

Content and Toxicity

Another trend at social media platforms has been to identify toxic language or content, especially with the help of artificial intelligence. The broad goal is to improve the health of platform-based conversation.

Twitter put out a call for help in fighting online abuse in 2018. In a public request for proposals, the company wrote:

> *We're committing to helping increase the collective health, openness, and civility of public conversation around the world, and to hold ourselves publicly accountable toward progress.* [57]

Many solutions employ artificial intelligence and content moderators to reduce toxicity. The New York Times, for example, uses Jigsaw's Perspective tool to help moderate its comment sections (Jigsaw is a project of Google).

In another example, Twitch created AutoMod, an automated chat moderation system, that is meant to help human moderators reduce or manage toxic comments and harassment on their channels and live streams. [58] (Since the bot was introduced in 2016, Twitch has also introduced tighter policies to address harassment on its platform.)

However, as we will explore later in this report, there is a limit to what AI and other technical solutions can effectively detect. This is especially true when considering the person targeted for harassment or exposed to upsetting content. Context is extremely important in understanding the severity and urgency of a report. It's unclear how this is taken into account across various platforms' moderation policies, reporting processes, and takedown or removal decisions. Targets of online harassment may not have the opportunity to present enough information that would provide richer context to a platform's trust and safety team.

Writer and journalist Sady Doyle, writing for Elle magazine in 2017, said of her experience reporting to Facebook, "According to Facebook's rubric, there was no box to check or form to fill out that would adequately explain the situation. I couldn't provide the documentation that might show a moderator why this person being able to contact me through their platform was a problem or why blocking him was not a solution. I was stuck." [59]

> *According to Facebook's rubric, there was no box to check or form to fill out that would adequately explain the situation. I couldn't provide the documentation that might show a moderator why this person being able to contact me through their platform was a problem or why blocking him was not a solution. I was stuck.*

Russian interference in the election brought platforms under fierce scrutiny for the spread of disinformation and extremist content on their networks. In response, Facebook has been scaling its safety, privacy, and policy teams. For example, the percentage of "Safety" jobs at Facebook grew "as a percentage of all job openings," according to a data analysis by Thinknum. [60]

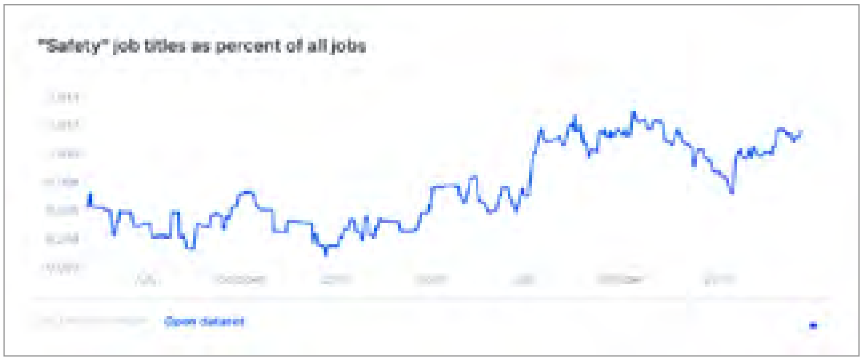


Figure 2.12, Data visualization by Thinknum, 2018

In 2017, Facebook announced that it would add 3,000 content moderators. [61] However, moderators perform emotionally dangerous work, mostly for low pay, and make decisions about questionable content in seconds. [62] While the work helps remove content that falls outside the platform's guidelines, it's unclear whether it's a viable solution to addressing harassing content on the platform.

Twitter and YouTube have also grown their Trust and Safety teams, including hiring for senior, director-level positions.[63]

Initiatives for Users

Many platforms are focused on toxicity and content. However, there is a lack of rhetoric and conversation around individual and group harassment. That means targets of harassment and other vulnerable users are still left out of the conversation about what their experiences are and what they need. Platforms fail to provide appropriate resources or recourse to those on their platforms who are most vulnerable to harassment. While they may be addressing these challenges internally, the efforts are a) not known to targets and b) not sufficiently addressing targets' concerns, given what individuals may need both proactively (before the harassment happens) and reactively (during and after an incident of harassment).

One new, recent effort was launched by Facebook in March 2019. The new initiative combines new AI technology to identify non-consensual distribution of intimate images, including doctored images, and expert-backed resources for targets. The initiative, called "Not Without My Consent," will be followed by a victim support toolkit created in partnership with Revenge Porn Helpline, Cyber Civil Rights Initiative, Digital Rights Foundation, SaferNet, and Professor Lee Ji-yeon (Hankuk University of Foreign Studies, South Korea). [64] The results of the project are yet to be seen, but initiatives that bring together researchers and experts in online harassment (including people who have experienced harassment) are promising.

Other Solutions or Initiatives

Below we'll look at the research about online harassment and adjacent fields, as well as examples of online harassment policy and legislation efforts. These efforts directly impact the individual's experience, options, and outcomes—both in the short and long-term. [65]

Academic Research

Academic research related to online harassment encompasses a wide range of focuses and disciplines, spanning from linguistics, sociology, psychology, computational linguistics, political science, economics, legal scholarship, and more. Because online harassment is a complex issue, academics are exploring the experience of online harassment, its effects on targets, and its broader implications on society from many different vantage points. In this section, we'll briefly touch on recent trends in the research and related topics, as well as the challenges of conducting such research.

This sampling of research was drawn from our conversations among both academics and practitioners. We also looked at the number of times a resource has been cited, bibliographies of relevant papers or reports, or compilations of research, such as the one led by Dr. J. Nate Matias (last updated: December 8, 2018).[66]

One of the most promising trends in academia is an effort to build on the growing body of work about online harassment, even across disciplines. Such collaborative efforts are key to developing richer understandings of online harassment, its consequences, and potential interventions. Reflecting this outlook, there are now even academic conference tracks dedicated to online harassment, proactive convenings, open-source projects, and platform-funded requests for research proposals. [67]

A Snapshot of Research Examples

Research in online harassment has increasingly focused on reflecting experiences of targeted individuals, potential coping strategies, and effects on particular demographics (e.g., women, people of color) and professions (e.g. journalists, academics), as well as their intersections.

Research Samples

Identifying Women's Experiences With and Strategies for Mitigating Negative Effects of Online Harassment. Jessica Vitak, Kalyani Chadha, Linda Steiner, and Zahra Ashktorab at the University of Maryland, College Park, MD, USA (2017)

'You really have to have a thick skin': A cross-cultural perspective on how online harassment influences female journalists. Journalism. Chen, G. M., Pain, P., Chen, V. Y., Mekelburg, M., Springer, N., & Troger, F. University of Texas at Austin. (2018)

Computational methods for characterizing and detecting online harassment and toxicity have been a persistent research focus of computer science, linguistics, cognitive science, and sociology in the past decade. These methods have seen institutional and financial support and social media platforms have invested in funding research in these areas. (Social networking sites, like Facebook and Twitter, and large tech companies, like Google, regularly publish requests for proposals across various research topics.) [68]

Research Samples

Aggression Identification Using Deep Learning and Data Augmentation. Risch, Julian & Krestel, Ralf. From the First Workshop on Trolling, Aggression and Cyberbullying at the 27th International Conference of Computational Linguistics (COLING 2018)

Community Interaction and Conflict on the Web. Srijan Kumar, William L. Hamilton, Jure Leskovec, Dan JurafskyDr. Srijan Kumar working on at Stanford University. (2018) [69]

Along with identifying models of harassment and profiles of harassers, as well as online cultural environments that correlate with increased toxicity, researchers have also focused on understanding the broader implications of social media, disinformation, toxicity, and online harassment on society.

Attempts have also been made to assess and understand successful models of prevention, intervention, and community-based solutions to online harassment in existing online communities.

Research Samples

When Online Harassment is Perceived as Justified. Lindsay Blackwell, Tianying Chen, Sarita Schoenebeck, Cliff Lampe University of Michigan School of Information. (2018)

CivilServant.io, a nonprofit founded by J. Nate Matias, a postdoctoral student at MIT, that organizes citizen behavioral scientists and works "directly with online communities to test ideas in moderation and evaluate the impact of the tech industry in our social lives."

There is also a growing body of interdisciplinary work on governance and ethics, media studies, legal studies (freedom of expression, legal consequences of online harassment, copyright law), policy, civil rights, and economics.

Snapshot of Trends in Findings

- Negative effects on personal and professional life
- Silencing of individuals (chilling effect) demonstrating people do not report or that they do not trust platforms
- Negative impact of online harassment on free press and freedom of expression
- Women and underrepresented individuals more likely to face harassment, both in frequency and intensity
- Online harassment as a public health issue and mental health crisis
- The ethics of research about online harassment (e.g. balancing user privacy and research opportunities) is challenging
- Domestic violence cases almost always have a digital component to them

Research Challenges

Challenges that have emerged, preventing or slowing researchers from making more progress:

Data collection — training data: Understanding harassment tactics and effects on individuals requires researchers to collect qualitative and quantitative data on a target's experience. Any algorithmic solution is only as good as the size and quality of its training data set. However, in the absence of high-quality, real-life training data, the algorithm's effectiveness will be limited. As we'll see in more detail below, data collection is challenging based on 1) data availability (including access to platforms); 2) an individual's willingness to contribute information; and 3) the high cost of the human-driven process for data annotation that is required for labeling data. The cost grows as more nuances are added—for example annotating new or more complex contextual dimensions.

Data collection — lack of access to data on platforms: Making platforms and the information they host available for research is necessary. However, it's proven difficult to execute while preserving the privacy of people using platforms.[70] Platforms may be hesitant to allow researchers to access information, especially given cases like the now infamous Cambridge Analytica scandal, where a researcher sold private participant data to a third party.[71]

> *You don't wanna create a situation where just because it's a researcher it's okay. You can't create that kind of situation...you have to protect Facebook's users' privacy. You have to take measures to make sure that the information you collect isn't going to be inadvertently disclosed. You can't transfer it to a third party. You can't transfer it, for example, to a data aggregator, or to any other commercial enterprise. You can use it only to inform the public about matters of public concern.*
>
> *...There's going to be disagreement about the meaning of some of these terms, and Facebook's going to have to flesh it out over time. Facebook would have to decide over time which projects it was willing to allow and which ones it was going to shut down. But in our view, that's a better situation than we're in right now, where Facebook has effectively categorically prohibited all of this journalism and research from taking place."* [72]
>
> — Jameel Jaffer, First Amendment lawyer and executive director
> of Columbia University's Knight First Amendment Institute

Ethics of research: Another ethical concern for researchers: how to systematically talk to and document the experiences of people who have been traumatized. Asking people to recount traumatic events is difficult and can re-traumatize individuals. People may also not agree to make their traumatic experiences public. This makes recruiting targets of online harassment for studies challenging. An individual may put themselves at risk simply by participating.

Accurate representation: Affecting academic and industry researchers alike, surveys, reporting data, interviews, and other log data exclude information about non-participation and underreporting. Academic researchers have the additional barrier of limited access to data outside of sampling that they are able to strategically access through APIs, scraping (including terms of service complications), and traditional survey/interview methods.

Harassment of academics: Researchers who investigate harassment and adjacent fields can, themselves, become targets of harassment.

Please refer to Appendix C for recent academic papers about online harassment spanning a variety of focuses and fields.

Legislation and Policy

One noteworthy collaboration was a collective effort that former California Attorney General Kamala Harris put together to address non-consensual distribution of intimate media, termed "cyber exploitation" by Harris. In 2015, Harris convened a task force that included leading scholars, nonprofits, victims, victim advocates, and social media platforms to address this particular tactic of online harassment. Among other results that included two California bills that enacted tougher policies on cyber-exploitation and more recourse for targets of non-consensual pornography (the bills went into effect on January 1, 2016), Google de-indexed search results when victims requested it and Twitter banned non-consensual intimate photos and videos. [73]

However, in a February 2019 profile of Harris, Politico Magazine reported that she's taken a step back from internet-related issues, writing:

> *Harris has burst quickly onto the national stage...Her legal fight against online harassment is one of her most innovative, and least-understood, contributions to public policy. But it's also potentially toxic with important Democratic constituencies.*

Since 2016, talks of legislation to address harassment at a federal level have stalled as platform data scandals and Russian election interference gripped national attention and the legislative agenda. However, 38 states and Washington D.C. have laws criminalizing the non-consensual distribution of intimate images. In addition, Congresswoman Katherine Clark (D-Mass) introduced the Online Safety Modernization Act to the House in 2017, but it was not enacted. The bill proposed an amendment to the federal penal code, directed the Department of Justice (DOJ) and the Federal Bureau of Investigation (FBI) to define and publish statistics on cybercrimes against individuals, and expanded funding for the FBI and DOJ to hire and train law enforcement officers in the investigation of cybercrimes against individuals.[74]

It's notable that non-consensual distribution of intimate images, as well as youth-focused topics like cyberbullying, have gained wide recognition and attention from policy makers. These efforts, which include both top-down and bottom-up efforts to gain traction, could serve as a playbook for addressing other tactics of online harassment.

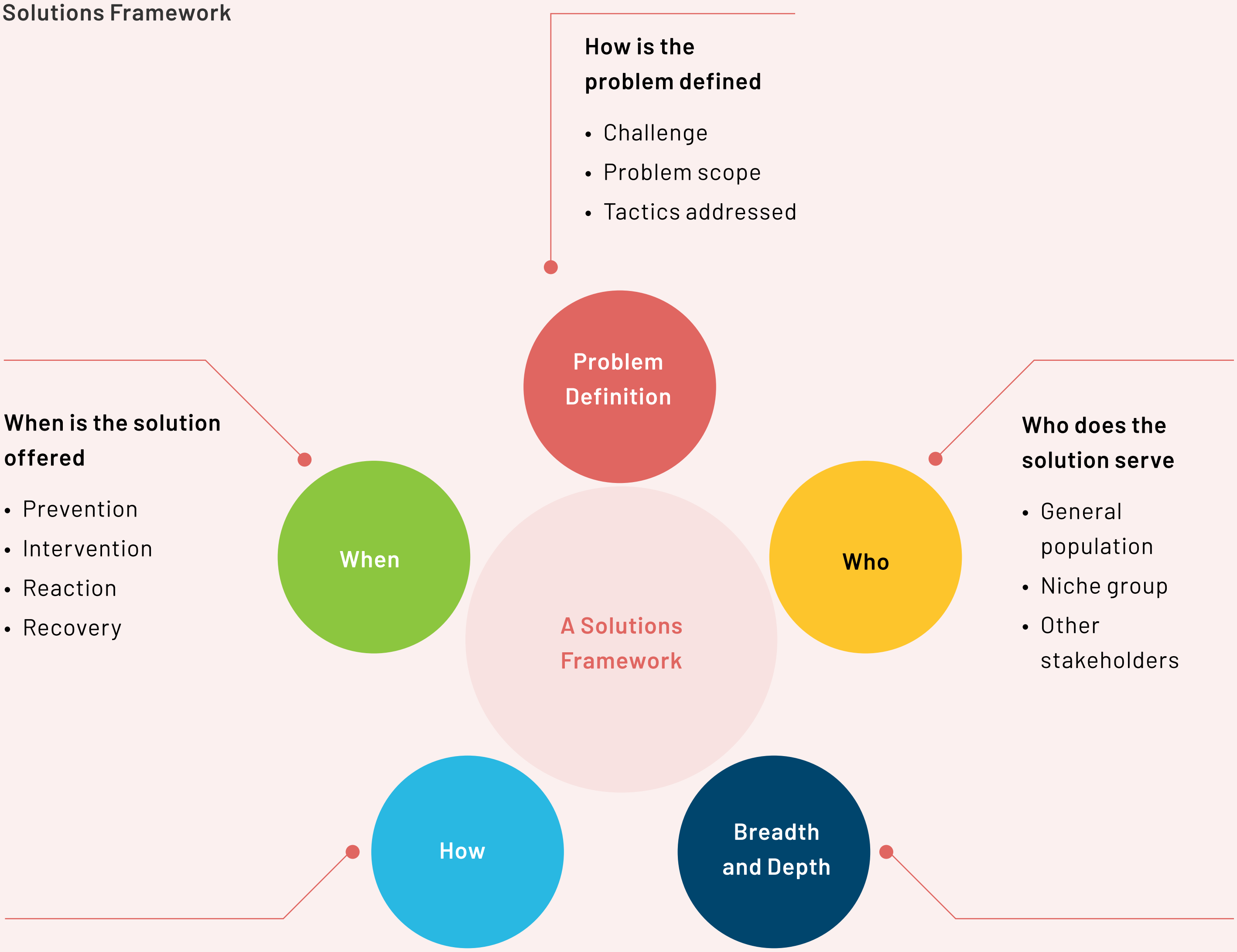
Part of this playbook is fruitful interaction between victim advocates, legal scholars, and coordination with legislators. In the example of non-consensual pornography, Holly Jacobs founded the nonprofit Cyber Civil Rights Initiative (CCRI) based on her personal experience with being a target of the distribution of intimate images without her consent. Jacobs worked closely with Mary Ann Franks, a professor of law at the University of Miami School of Law to garner the attention and contribution of legal scholars, law practitioners, advocates, legislators, the media, and social media platforms. Jacobs and Franks' efforts laid the foundation for non-consenual pornography laws (developing both model state and model federal statutes) across the country.,[75] [76] Another important part of this effort was advocating for the use of the term non-consensual pornography, instead of the more common term "revenge porn," as a more accurate way to describe this tactic of harassment. [77]

A Solutions Framework

To provide targets of harassment, solutions funders, and other stakeholders an easy way to identify the most relevant solutions to their needs or interests, we offer a five-point solutions framework of categorization.

This framework can be used to categorize existing and proposed solutions to address online harassment.

The 5-Point Solutions Framework



How is the problem defined

- Challenge
- Problem scope
- Tactics addressed

When is the solution offered

- Prevention
- Intervention
- Reaction
- Recovery

Who does the solution serve

- General population
- Niche group
- Other stakeholders

What remedy,

- philosophy,
- support,
- framework,
- or plan of action is offered?

Does the solution focus on depth or breadth of impact

- Individual care
- Community-specific
- Systemic change or change at scale

Figure 2.13, The 5-Point Solutions Framework, OnlineSOS, 2019.

5 Point Framework:
Deeper Questions Table

Problem Definition	How do you define problem scope? Are you focused on a specific tactic or behavior? What part of the challenge does your solution address?
Who	Who is your solution for? What group does your solution serve or benefit? Individuals, a group, or for other stakeholders? Are you focused on a specific demographic (e.g. journalists, women and girls)?
When	When does your solution come into play? We've identified four key stages in the online harassment cycle: • Prevention (e.g. training) • Intervention (e.g. counterspeech) • Response (e.g. crisis helpline) • Recovery (e.g. ongoing mental health services)
How	What remedy, philosophy, support, framework, product, or plan of action is offered? Do you focus on a specific field and/or medium (e.g. legal)? Does your solution rely on a community (e.g. moderators or peer-to-peer enabled)?
Breadth and Depth	Why do you support this specific solution? What are your intended outcomes? What are the driving motivations? To what degree are you focused on breadth or depth? What's at stake? Are there any other incentives, priorities, or milestones you are optimizing for?

Figure 2.14, OnlineSOS, 2019.

Summary

- An individual can experience a wide range of online harassment tactics and under varying circumstances. Therefore, people need options and a plan of action unique to their situation.

- Targeted individuals may experience trauma including emotional distress and anxiety.

- People experiencing online harassment have three key needs: 1) physical safety; 2) emotional and psychological well-being; and 3) digital security.

- Resources or solutions that currently exist include: 1) DIY guides and resources; 2) support from family and friends; 3) grassroots organizations, with limited direct or active services; 4) some legal services, with limited options for recourse; 5) blocking, reporting, and privacy tools from social media platforms, although severely limited in effectiveness and are often insufficient for targets' needs; 5) academic research furthering understanding of online harassment and its consequences; and 6) some legislation providing recourse for targets, in particular for victims of non-consensual distribution of intimate images.

- While some resources do exist, there is no guarantee that a target has adequate access to the resources or that the resources are up-to-date.

- More interdisciplinary work—from experts to technologists; activists to policy-makers—is needed to ensure targets of online harassment receive the resources, recourse, and recovery options they need.

- In developing new remedies, we offer a five point framework for categorizing how a solution best fits into the online harassment experience.

Part 3

How to Create Better Outcomes for Targets of Online Harassment

Establishing Key Priorities

Academics, technologists, and organizations are investing time and effort into understanding online harassment and developing potential interventions, as well as ways to limit abuse. However, there are still few concrete tactics for effectively reducing online harassment cases.

The reasons for this may include:

- **Lack of collaboration**: Researchers, organizations, platforms, businesses, and other sectors typically operate in silos.
- **Business models**: Social media platforms have business models that do not incentivize the time, effort, and resources needed to address the root of problems on their platform.[78]
- **Legal and policy lag**: Mechanisms to deter or penalize abusive online behavior are lacking and lagging.
- **Lack of useful taxonomy**: There is no existing shared definition of online harassment and related terms, neither within society, nor among those working to address the issue.
- **Impact ambiguity:** The impact of online harassment on society, business, and democracy remains ambiguous and could be more explicitly discussed.

Effectively tackling online harassment requires the coordination and cooperation of technology platforms, advocacy organizations, crisis management professionals, law enforcement, legal and policy experts, legislators, researchers, and educators.

OnlineSOS sees three key priorities moving forward:

1. Develop a common vocabulary to describe and understand online harassment (both tactics and harms).
2. Amplify and address the needs of targets and make it easier for them to report, assess, document, and resolve their online harassment case(s).
    a. Define and support the development of preventative measures in tech/product design and proactive security measures.
    b. Continue funding experts who can focus on:
        i. Human-centric needs
        ii. Investing in marginalized community leaders and grassroots organizers
3. Build coalitions, especially among uncommon allies, to support work, research, and initiatives that grant more groups 1) participation in building solutions and; 2) access to resources and tools needed to address online harassment within their own niche, community, or industry.

1. Develop a Common Vocabulary to Describe and Understand Online Harassment

By describing online abuse more accurately, journalists, organizations, platforms, and authorities can more effectively respond to online harassment. Naming tactics of online harassment demonstrates that it is neither normal nor inconsequential.

Of the 41 percent of respondents who identified with having experienced one of the six harassing behaviors presented in the 2017 Pew Research survey (offensive name calling, purposeful embarrassment, physical threats, stalking, sexual harassment, and sustained harassment), 27 percent were not sure if they would consider their most recent incident as harassment.

Targets who can identify what's happening to them as harassment—and a specific kind of harassment—could help surface the resources they need, identify the best next steps to take, and increase the likelihood of obtaining a satisfactory response from platforms, law enforcement, or community members (like moderators and administrators).

Defining tactics of harassment, together with a better understanding of context, has far-reaching implications for not only individuals, but also for effective policy making, platform governance, and research.

As pointed out by Pater, Kim, Mynatt and Fiesler in a 2016 research paper, "Characterizations of Online Harassment: Comparing Policies Across Social Media Platforms," social media platforms often do not define or spell out what "harassment" is in their terms of service or other formal documents.

> *"Lack of specificity and consistency for what constitutes harassment makes future work on understanding harassment policies at a deeper, more granular level challenging."*

They conclude that specific definitions of online harassment are critical to governing platforms and addressing online harassment more effectively.

> *"We believe it is imperative to understand platform-specific characterizations of harassment before examining effectiveness of policy or the ethically appropriate ways of regulating harassing behaviors online. Once we have this foundation, we can begin to assess the embedded social norms surrounding harassment within online platforms and the roles that written policies play in mitigating and reducing harassment within these communities."*

Opportunity: Improve transparency of platforms

While it's understandable that platforms cannot provide explanations for all decisions, processes, and policies made at their company, more explicitly defining what harassment means to them provides better guidelines for how their communities are expected to behave. This informs community members on how platforms remove content, ban users, or otherwise address harassing behavior.

Currently, the decision-making process and policies are often opaque, creating confusing and inconsistent methods of dealing with online harassment. This has resulted in disturbing content or abusive behavior persisting on platforms.[79]

For example, Facebook's Community Standards are ambiguous as to what constitutes harassment on the platform and what consequences a user would face as a result. They say:

> *The consequences for violating our Community Standards vary depending on the severity of the violation and a person's history on the platform. For instance, we may warn someone for a first violation, but if they continue to violate our policies, we may restrict their ability to post on Facebook or disable their profile. We also may notify law enforcement when we believe there is a genuine risk of physical harm or a direct threat to public safety.*[80]

Such ambiguity has also frustrated platform users and have led to calls for more clear, transparent guidelines around what is and is not acceptable. For instance, in early 2018, Twitch published a blog post about recent updates to its community guidelines. Quoting the update, one user commented:

> *'First, conduct we deem to be hateful will result in an immediate indefinite suspension. Hate simply has no place in the Twitch community.' Can you please give us an EXPLICIT LIST of what you consider to be hateful?* [81]

A better understanding of these processes and policies—for researchers, experts, and targets of harassment—can create a more open dialogue about what proactive and reactive measures would be most effective for addressing online harassment and its parallel behaviors, like disinformation.

One promising project is Wikimedia's Community Health Initiative, which provides a transparent look at how it's addressing harassment on its platforms. [82] Although Wikipedia and other Wikimedia projects have contributors and editors already adhering to a greater set of guidelines than the average messaging platform—and is not beholden to shareholders—this initiative could serve as a potential model for platforms like Facebook and Twitter.

Another promising initiative is a collaborative study to diminish abuse on Twitter led by researchers Susan Benesch and J. Nathan Matias launched in April 2018. Benesch announced the study on Medium:

"Today Twitter will begin testing such an idea: that showing an internet platform's rules to users will improve behavior on that platform. Social norms, which are people's beliefs about what institutions and other people consider acceptable behavior, powerfully influence what people do and don't do. Research has shown that when institutions publish rules clearly, people are more likely to follow them. We also have early evidence from Nathan's research with reddit communities that making policies visible can improve online behavior. In an experiment starting today, Twitter is publicizing its rules, to test whether this improves civility." [83]

Opportunity: Prioritize better processes and experiences for targets of online harassment

A 2017 paper about women's experience with online harassment notes, "Social media platforms have taken a first step in acknowledging online harassment as a problem, and one that disproportionately affects women. Going forward, they need to take significant and visible steps to show users that they not only care about women's experiences and the harm that negative comments cause, but that they are prioritizing efforts to reduce the quantity and severity of such content." [84]

This highlights that defining harassment is also about aligning the community's expectations with the response of platforms. Failing to explicitly define harassment creates poor outcomes for targets of online harassment who believe—often rightfully so—that their complaint is not taken seriously.

> *I have had tweets saying 'We are going to tie you up, we are going to make you drink bleach, you will be sorry when you are in a burkha'. I have also had someone incite others to rape me, with the words: 'Who wouldn't rape Sophie Walker?' When I reported these to Twitter, the response came back so quickly it was almost like an 'out of office' reply. It was so fast that it felt automated, rather than considered. They said: 'We've investigated and there is nothing to see here.' I've tried escalating the reports, but the reply comes back just as fast and just the same. In some respects, Twitter's response to abuse is more hurtful than the abuse. It feels like they are saying: 'You're on your own if you participate in this forum.*[85]
>
> — Respondent in an Amnesty International study
> about Twitter and the online abuse of women

Specificity can exact better outcomes for targets of online harassment. In the case of laws governing the non-consensual distribution of intimate images, for example, overly vague statutes give perpetrators leniency and fail targets of abuse. [86]

The words we use can also distort the public's understanding of online harassment and have a detrimental impact on targeted individuals. For example, "revenge porn" is a common term for the non-consensual distribution of intimate images (images which can also be doctored). The term undercuts the seriousness of such crimes and resorts to victim blaming. Writing for the Women's Media Center—and as described in her book, HATERS, in more depth—Bailey Poland notes,

"Calling this act 'revenge porn' implies that the woman in the picture has done something that necessitates or excuses an act of vengeance…'Revenge' has nothing to do with many instances of sharing nude images of women without their permission—the goal is, instead, to humiliate women and control access to women's bodies and lives without their input." [87]

The actor Jennifer Lawrence called the distribution of her intimate images a "**sex crime**."[88] In her work against the non-consensual distribution of intimate images, Kamala Harris referred to this behavior as "**cyber exploitation." Both clarify that the perpetrator is responsible for the behavior, rather than blame the target.**

Taxonomy work has been repeatedly identified as a key potential driver for reducing online harassment and developing better solutions or resources for targets. A 2016 MIT convening identified this as a "high impact" opportunity for online harassment research and action. [89] In another example, a 2017 study by Lindsay Blackwell, a researcher at the University of Michigan, and HeartMob, concluded that labeling harassing behaviors validated targets' experiences, helped bystanders better understand the problem, and helped set norms for appropriate and expected user behavior. [90]

Another way to consider taxonomy and categorization of online harassment more broadly within the online safety landscape is by looking at risk.

Microsoft uses four risk categories to group 21 risks an individual can face online. This is shown in the figure below, taken from Microsoft's 2018 Digital Civility Global Report (slide 58, linked below). This list includes online harassment and other behaviors (e.g. doxxing) that are often considered and colloquially described by individuals as online harassment.

Risk Categories Used by Microsoft

Intrusive Risks

Unwanted Contact
Being personally contacted (by phone or in person) by someone who obtained your information online but without inviting them to contact you.

Hoaxes, Scams & Frauds
The spreading of false rumors (e.g., chain letters), criminal attempts to obtain personal information often for monetary gain (e.g., phishing scams), malicious emails disguised as someone you know (e.g. virus).]

Hate Speech
Speech that attacks a person or group based on gender, ethnic origin, religion, race, disability, or sexual orientation.

Discrimination
A person who is discriminated against or excluded based on gender, ethnic origin, religion, race, disability, or sexual orientation

Misogyny
An expression or demonstration of dislike, contempt for, or ingrained prejudice against women

Terrorism Recruiting
An attempt by a terrorist or terrorist organization to recruit a person for the purposes of causing harm.

Behavioral Risks

Treated Mean
Words or messages sent to another person online that are unkind, unfair or malicious.

Trolling
A deliberate act to make someone mad or angry using online or social media comments in a clever, but deceitful manner.

Microaggression
Casual insults made towards any marginalized group in society (e.g., religious or ethnic minorities, women, LGBT, people with disabilities, etc...).

Cyberbullying
When the Internet, phones or other devices are used to send or post text, images, or video intended to hurt, embarrass or intimidate another person.

Swatting
The act of deceiving emergency services (e.g., police, fire, medical) into sending an emergency response based on the false report of an ongoing critical incident or crime.

Online Harassment
Threats or other offensive behavior (not sexual solicitation) sent online or posted online for others to see.

Sexual Risks

Unwanted Sexting Received
Received unwanted sexually explicit messages and imagery.

Sexual Solicitation
A person who requests to engage in sexual activities or sexual talk or to give personal sexual information that is unwanted.

Unwanted Sexting Sent
I sent unwanted sexually explicit messages and imagery.

Sextortion
When someone threatens to distribute your private and sensitive material if you don't provide them images of a sexual nature, sexual favors, or money. The perpetrator may also threaten to harm your friends or relatives by using information they have obtained from your electronic devices unless you comply with their demands.

Unwanted Sexual Attention
Unwelcomed sexually oriented teasing, joking or flirting online or via electronic means

Revenge Pornography
A sexually explicit portrayal of one or more people distributed without their consent.

Reputational Risks

Doxing
The process of collecting and distributing or posting information about a person (e.g., name, age, email, address, phone number, photographs, etc.) without their permission.

Damage to Personal Reputation
Damage or destruction to the image created of you through PERSONAL information you or others shared online in blogs, postings, pictures, tweets, videos, etc.

Damage to Professional/ Work Reputation
Damage or destruction to the image created of you through work information you or others shared online in blogs, postings, pictures, tweets, videos, etc.

Figure 3.1, Online risk definitions as presented in Microsoft's 2018 Digital Civility Global Report. Content by Microsoft, redesigned by OnlineSOS, 2019.

For a sampling of existing lexicon work, you can refer to these varying examples:

- Dangerous Speech Project: Guide and Framework
- Microsoft: Digital Civility Index Resource Guide and 2018 Digital Civility Global Report
- Online Harassment and Content Moderation: The Case of Blocklists (p.15)
- Online Harassment, Defamation, and Hateful Speech: A Primer of the Legal Landscape
- PEN America: Glossary of Terms
- Tactical Tech: The Atlas of Online Harassment
- Women's Media Center: Online Abuse 101

2. Amplify and Address Targets' Needs

Context-first awareness of online harassment

Understanding context is key to developing tools, solutions, and policies that actually address people's needs. Context changes perceptions of severity and urgency, measures of which can help elicit appropriate responses from law enforcement, crisis management teams, employers, or platforms.

Online harassment may seem complex, but as an experience it is, in fact, simple. The complexity arises when we start to consider how to prevent, address, or remedy abuse.

Online Harassment Theory

Online harassment may seem complex,
but as an experience it is, in fact, simple.

Harassment Experience

=

Content × (1 + Context)

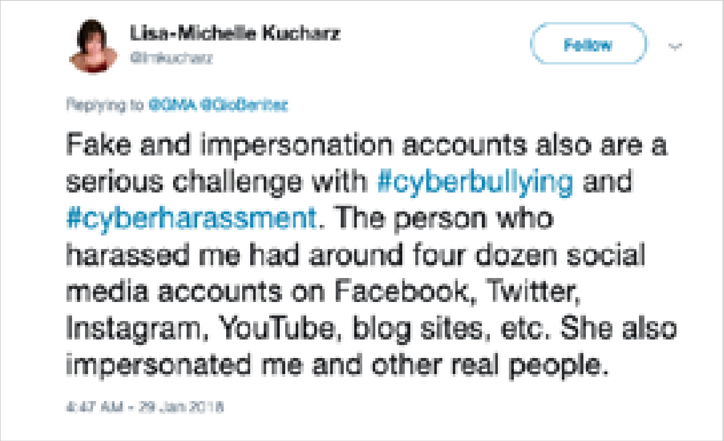
Figure 3.2 An Online Harassment Theory: How Content and Context Come Together.

Harassment involves content and context. Harassment is highly contextual, designed to cause emotional distress and fear of harm to the targeted individual. Context amplifies the experience, severity, and effectiveness/harm of online harassment.

Context amplifies the experience, severity, and effectiveness/harm of online harassment.

Online harassment involves two parts: content and context. Harassing content is easy to recognize: a vile comment, photo, or meme. Context includes:

- Who an abuser is (and what their relation is to the person being harassed).
  This can say a lot about the motivations of an abuser and their decisions. This harasser could be anonymous, known, or someone appearing to be anonymous but is known to the target.
- When a person is being harassed.
  Timing can describe why someone is being targeted and for what end goal. For example, harassers may target an investigative journalist before a major article drops to retaliate against them or intimidate and silence them.
- Where a person is being harassed.
  The communication channel(s) an abuser selects is often a function of where they can maximize the intended impact on the target. In addition, harassment often happens across multiple channels, with the intention to overwhelm a target. Again, this is meant to maximize the impact on the target. The abuser may use multiple platforms and "locations" to make their target feel helpless, even in front of bystanders.



Lisa-Michelle Kucharz
@lmkucharz

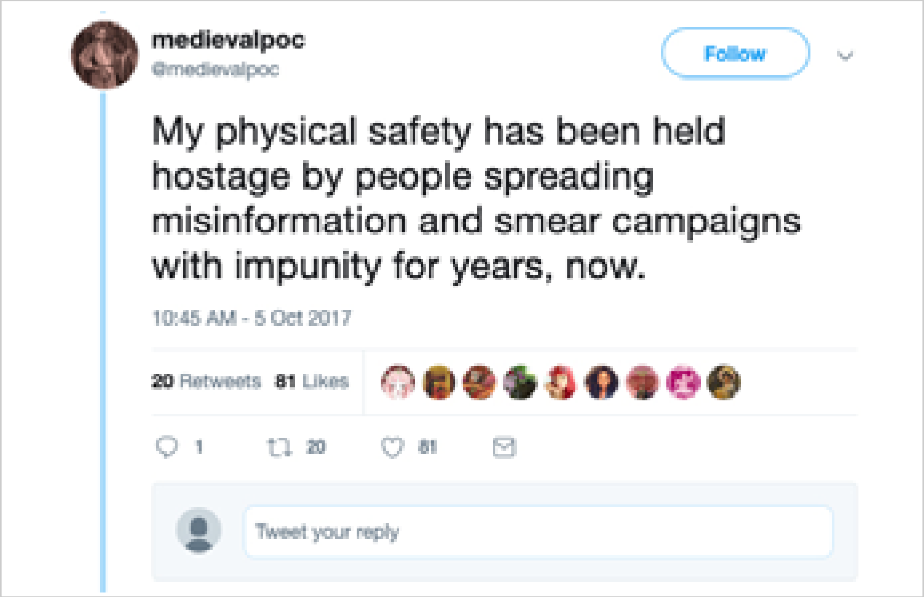
Replying to @GMA @GioBenitez

Fake and impersonation accounts also are a serious challenge with #cyberbullying and #cyberharassment. The person who harassed me had around four dozen social media accounts on Facebook, Twitter, Instagram, YouTube, blog sites, etc. She also impersonated me and other real people.

4:47 AM - 29 Jan 2018

Figure 3.3, Twitter, 2019 at https://twitter.com/lmkucharz/status/957958272337743872

- How long a person has experienced harassment.
  Even if only one comment is visible to the public, the picture shifts if we learn that an abuser is also sending a target consistent, sustained direct messages over a period of days or months.



medievalpoc
@medievalpoc

My physical safety has been held hostage by people spreading misinformation and smear campaigns with impunity for years, now.

10:45 AM - 5 Oct 2017

20 Retweets 81 Likes

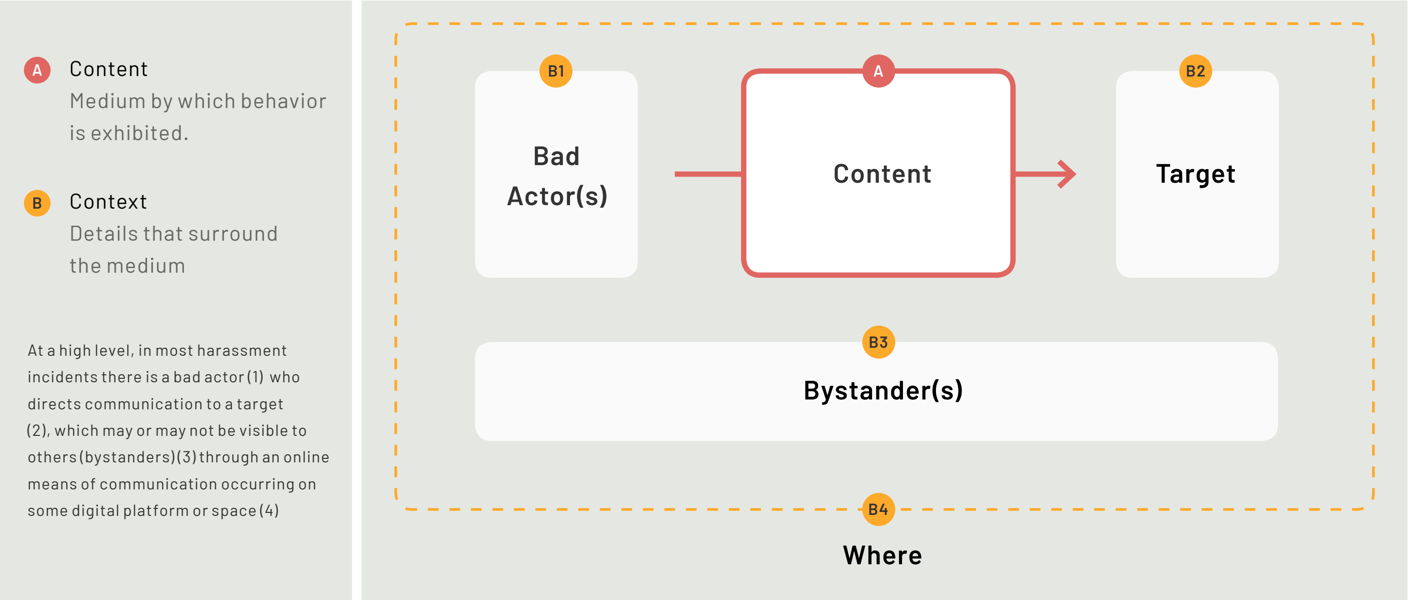
Figure 3.4, Twitter, 2019 at https://twitter.com/medievalpoc/status/915996361983758342

- How an abuser is harassing another person.
  Being able to identify what is happening leaves less room for ambiguous interpretations about seriousness or urgency. This can affect the response a target receives from reporting to, say, a social media platform. If "trolling" can mean both receiving an annoying comment and receiving repeated threats of violence, how can a target convey the seriousness and urgency of their particular situation?

Context amplifies the experience, severity, and effectiveness of the abuse. Therefore, knowing the answers to these questions creates a much richer picture about a target's experience. Understanding context can help moderators prioritize reports, law enforcement assess a threat's viability, or targets make more informed decisions about best next steps.

Described below in Figure 3.5: At a high level, most harassment incidents involve a bad actor (B1) who directs content (A) at a target (B2). The communication may or may not be visible to bystanders (B3), but the communication takes place through some online means, like email, a comment thread, a social media platform, and so on (B4).

How Online Harassment Works — The Online Harassment Interaction


A Content
Medium by which behavior is exhibited.
B Context
Details that surround the medium
B1
Bad Actor(s)
A
Content
B2
Target
B3
Bystander(s)
B4
Where

At a high level, in most harassment incidents there is a bad actor (1) who directs communication to a target (2), which may or may not be visible to others (bystanders) (3) through an online means of communication occurring on some digital platform or space (4)

Figure 3.5, How Online Harassment Works: The Online Harassment Interaction, OnlineSOS, 2018.

Combined with the other pieces of context, a clearer picture emerges about what a target is experiencing, how it's affecting them, and what threat the abuse poses to them (from psychological to physical, personal to professional).

Let's take two examples, which come from OnlineSOS's direct service triage:

1. Someone says, "You're such a stupid c*nt," to a man. The comment is in poor taste, but it's difficult to know its meaning because there is a lack of context. In this case, however, the comment was directed at a woman. It was conceived as a gender-based aggression that the woman found startling, distressing, and threatening.
2. Someone says, "That's a nice yellow shirt you're wearing." This comment might be nothing more than a compliment from a typical person, like a friend. However, the comment was made by an ex-partner stalking the target. This detail completely changes the comment's meaning. The comment now indicates that the individual is potentially being surveilled and may be in danger.

Platforms have acknowledged that behavior and context are important to addressing online harassment and threats. For example, in 2018 Facebook released a statement about fake accounts that read, "This is why it's so important we look at these actors' behavior – such as whether they're using fake accounts or repeatedly posting spam – rather than their content when deciding which of these accounts, Pages or Groups to remove." [91]

In practice, however, platforms vary widely in how much they consider (or are able to consider) behavioral patterns and other rich context when responding to cases of online harassment. Furthermore, how platforms manage reports of harassment or abuse are opaque. More transparency is needed to understand what is currently being done to take context into account. And, more research and investment is needed to assess how context can be taken into consideration effectively at scale.

Artificial Intelligence (AI) and Context

Artificial intelligence and machine learning have been discussed as potential ways to address online harassment by making predictions about what text is considered toxic or abusive. However, these proposals only cover part of the problem. **Online harassment is highly contextual and algorithmic solutions have limitations and challenges in detecting context.** It's much more challenging to detect harassment, threats, and determine severity because you need to teach the machine to identify such patterns of behavior.

So, to say that AI and tech is the solution for curbing or addressing online harassment is a misconception. While AI has the power to impact change at scale, more recognition and understanding of context and concerning behavior is necessary.

> *Machines may also be able to help intervene and stamp out bullying too. But until AI perfects a way of detecting the most subtle and cunning bullying tactics, the responsibility will still lie with us.*
>
> *— Sarah Griffiths, reporter for the BBC* [92]

Rethink product design, engineering, and business models

One challenge to addressing online harassment is entrenched product development and design processes.

Firstly, product development is centered around current business models (for example, reducing friction for users to engage with the product, keeping users on a platform longer, or bringing users back to a platform or app more often). This is at odds with development and design that could put the safest experience for users first and foremost.

> "Abusability-testing" is gaining some recognition, as more designers, product managers, and developers come to terms with what can happen when well-intentioned products are co-opted by bad actors. [93]

But so long as advertising remains the key driver of growth and revenue—which in turn relies on user growth and engagement—there's little indication that platforms will be able to prioritize how they approach product development.

Take live streaming, for example, which is an extremely effective engagement feature on platforms like Facebook and YouTube. Sarah T. Roberts, a professor of information studies at the University of California, Los Angeles recently told the New Yorker, "By and large, live streaming came online with none of that stuff sorted out… There was no real plan for, 'What are we going to do when people start using it in the worst possible way?'" [94]

There's a distinct reason for this. Former Google Product Counsel, Nicole Wong, described the early days of Google Search in a 2018 interview with Kara Swisher:

"In the mid-2000s, when social networks and behavioral advertising came into play, there was this change in the principles that...we just weren't as concerned about search anymore, instead we were focusing on this other part of the platform...Personalization, engagement...what keeps you here."

She added that it's possible to alter this—if the will exists, "So what if we change the pillars again? What if now everything that we've learned in the last two years, we say, 'That's not the internet we want to live with'? So this is just personal for me, like, what if the pillars were accuracy, authenticity and context." [95]

Furthermore, activist investor Roger McNamee, an early Facebook investor, recently wrote in a Time op-ed that "[Facebook] respond[s] to nearly every problem with the same approach that created the problem in the first place: more AI, more code, more short-term fixes. They do not do this because they are bad people. They do this because success has warped their perception of reality. They cannot imagine that the recent problems could be in any way linked to their designs or business decisions."[96]

Even after a tumultuous year at Facebook that included data privacy scandals and disinformation campaigns (among others), revenue from advertising increased more than 30% between Q4 2018 and Q4 2017. The company generated $16 billion from advertising alone between October 2018 and the end of the year. Facebook made an average of $34.86 per user in Q4 '18 compared to only $3.20 in the same period in 2011 and up from $19.81 in Q4 '16. [97]

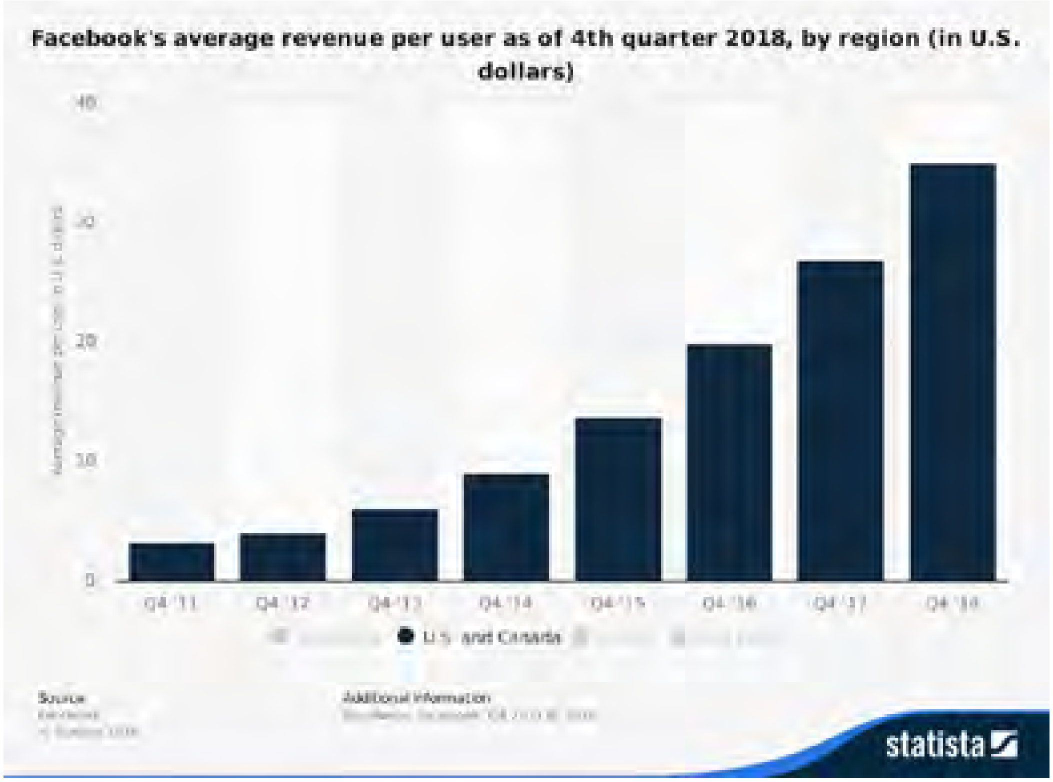


Figure 3.6, Data from Facebook, Visualization by Statista, 2019.

Secondly, platforms—and the internet more generally—have largely been built, designed, and led by white males. Yet online harassment disproportionately affects women, people of color, and LGBTQ+ people. This emphasizes the imbalance between those who build online platforms and those who cultivate and maintain the communities within them.

> *One of the main takeaways from our session at this year's Internet Governance Forum is that the better, more resilient and diverse communities we have both online and offline, the better chance we have of supporting those who are harassed in a meaningful way.*
>
> — Andreas Reventlow, an International Media Support advisor, writing about the Internet Governance Forum, 2016 [98]

This includes those who design and build the online communities themselves. The composition of founding executives and product leadership teams are not currently reflective enough of the diversity of their user bases, but platforms are exploring ways to bridge the gap through HR efforts as well as engagement of external stakeholders and users. Examples of mechanisms currently under experimentation by platforms include requests for feedback, targeted research, partnerships, community outreach, community grants, workshops/summits, and boards comprised of external experts. [99]

3. Build Coalitions

Through collaboration and a multidimensional approach to online harassment, we can develop more useful, actionable, and flexible resources for targets of online harassment.

There's strong evidence that collaboration can facilitate more productive outcomes for both a) reducing online harassment and b) improving an individual's online experience.

> *What led the reforms that Riot Games instituted was a 'player behavior team' of people with 'PhDs in psychology, cognitive science, and neuroscience to study the issue of harassment by building and analyzing behavioral profiles for tens of millions of users.' Riot Games assembled a panel of experts to design bespoke solutions for their product; their experts delivered.* [100]
>
> — Sarah Jeong, Internet of Garbage, page 77

Among the groups addressing online harassment, collaborative efforts have declined since 2017. Past efforts to build interdisciplinary coalitions are either winding down or inactive. Below are examples of productive convenings and structures for collaboration that are no longer active or meaningfully less active in 2019:

- Convening hosted by MIT Media Lab
- Convening hosted by the California Attorney General's Office
- Heartmob + Hollaback listserv
- Safety Group hosted by Soraya Chemaly, Women's Media Center director, in partnership with the National Network to End Domestic Violence Safety Net Project
- Tactical Tech Resource Directory
- #HackHarassment by Vox Media, Born This Way Foundation, and Intel

What's most troubling is that this vacuum is not being filled by new organically-formed, grassroots organizations or the top minds and leaders in online harassment research, activism, and advocacy. Instead, special interest groups and social media platforms are convening to discuss online harassment and other related topics, slowly replacing the outflow of grassroots organizations addressing online harassment. As a result, profits and other special interests could set the agenda for the conversation about online harassment, rather than ensure targets' personal experiences of online harassment remain a focus. This could increase the risk of group censorship and put the health and integrity of our democracy at risk.

> As a result, profits and other special interests could set the agenda for the conversation about online harassment, rather than ensure targets' personal experiences of online harassment remain a focus.

These groups' primary focus is on content moderation, rather than the impact and experience of individual users:

- COMO: Content Moderation and Removal at Scale http://comoatscale.com/
- FOSI: Family Online Safety Institute https://www.fosi.org/

In addition, there are sector or vertical-specific conferences or convenings—e.g. legal or research focused—that are often siloed and are specialized for discussion within a specific field.

Because of concerns about special interest groups and their potential influence, the following items need to be prioritized:

- **Promote visibility of and access to DIY resources and information:** It's clear that targets need DIY resources and information to address immediate needs or concerns when they've experienced online harassment.
- **Fund experts, communities, and awareness and advocacy initiatives:** In particular, grassroots organizations that can provide resources, services, or training to their communities are critical to developing more personalized courses of action for individuals targeted by online harassment. This includes prioritizing funding for vulnerable communities, underrepresented groups, or professions / activists / demographics that are most likely to be targeted.
- **Provide more pathways to direct services:** Providing targets of online harassment more ways to access the direct services they need, including mental health, crisis management, or personalized digital security services can help reduce "chilling effects" of online harassment and keep people engaged online. For example, an organization can conduct formal training about how online harassment works and its impacts to groups like law enforcement and legal aid; or, in another example, form partnerships between experts/researchers, grassroots organizations, and platforms that direct people to personalized services during the reporting process.

Less Formal Avenues of Collaboration

While it's important that collaboration is formalized, informal discussions, collaborations, and partnerships also drive awareness about online harassment and can initiate change. These include artistic and creative efforts around online harassment, related topics, and the individual's experience, as well as growing coverage of online harassment topics at conferences, convenings, and popular media (e.g. TED Talks). These efforts inherently create collaboration—between individuals, organizations, audiences, or more informal community groups, like a women's group—because they bring different people together for the purpose of discussion, creation, or action.

Art and Grassroots Efforts

Formal pathways of discussion are not the only way to catalyze change and, therefore, artistic, organic, and other grassroots efforts—even on-the-fly Twitter threads about the harassment experience—help educate the public about online harassment and its harms. [101]

One recent example is the film Netizens, a documentary that tells the stories of women and their experiences of cyber harassment. Screenings around the world have been accompanied by Q&A sessions and discussions about harassment among key audiences such as women and girls. students and educators, and the criminal justice community. [102]

Adjacent Gatherings

Discussions and panels about online harassment have also taken place at Online News Association conferences, the Internet Freedom Festival, RightsCon, Committee to Protect Journalists events, and UN Women events and campaigns. TED Talks about online harassment, privacy, and cyber abuse have also sparked ongoing conversations on the topic. [103]

Part 4

Looking Forward

Online harassment continues to evolve rapidly. Even between the time that the OnlineSOS team concluded the research for this report and publication:

- Facebook announced a "ban on praise, support and representation of white nationalism and white separatism on Facebook and Instagram." — **March 2019**
- Michigan enacted a new cyberbullying law. The law includes specific definitions for cyberbullying and harassing behaviors. — **March 2019**
- Courts ruled in favor of Grindr in *Herrick vs Grindr LLC* mentioned in Part 1 of this report, relieving Grindr of liability for harassment that happens on its app. — **March 2019**
- The SWATer in a case mentioned in Part 1 of this report, in which an innocent bystander Andrew Finch was murdered, was sentenced to 20 years in prison. — **March 2019**
- Dictionary.com added two new harassment-related words, "crybully" and "cybermob," to its website. — **March 2019**
- The New York Times released an investigation into harassment endured by volunteer Wikipedia editors. — **April 2019**
- The UK Government published an Online Harms White Paper as part of its plan to develop a package of measures to keep people safe online and open the conversation for public feedback. — **April 2019**

No doubt, by the time you read this report, there will be new stories, studies, and reports about online harassment published. However, as the sampling of news shows, we still are not left with many net positive changes. For every step forward there is another taken back by more reports of widespread harassment.

Therefore, what can we take from these developments and how can we build on any momentum? How can we catalyze lasting change now? After 30 years online, we have a body of history from which to learn—playbooks to study, cases to review, and successes and failures to analyze. In moving forward, we need to not only deconstruct online harassment, but also reconstruct the system within which it exists.

It's impossible to encapsulate all of the nuances and complexities of online harassment in one place, but we hope that some of these insights catalyze new conversations about online harassment and engender unconventional allyships to address it. Online harassment is one of the most important social issues of our time and yet too many people who are targeted feel hopeless and helpless in the face of harassment. It's not an option to simply retreat into the "offline" world. That wouldn't be a just option—not only for the individuals targeted—but also for the health of democracy and a more equitable society.

Into 2020:
The State of Online
Harassment
and Opportunities for
Collaboration

Appendices

Appendix A — Grassroots Organization and Initiatives Directory

How to Read Each Profile

Example Organization Name

Website — example.com
About — A brief description of this organization

Founders — Who is/are the founder(s)? Are they still active in addressing online harassment?

Products, Solutions or Services - A description of the organization's campaigns, projects, initiatives, products, or services. Who the solution serves is **bolded.**

Latest Initiatives or Public Activity: What are they working on now? Most recent update? A snapshot of the organization's most recent activity.

Amnesty International — Online Violence Against Women

Website
https://www.amnestyusa.org/online-violence-against-women/

About Organization "All people have fundamental human rights. But those rights are abused or denied every single day. When that happens, Amnesty International finds the facts, exposes what's happening, and rallies people together to force governments and others to respect everyone's human rights... Amnesty International's uniquely effective approach for protecting human rights uses a three-pronged approach: research, mobilization, advocacy."

Founded in 1961, online harassment and hate speech are just two of the many issues the Amnesty International addresses.

Products, Solutions, or Services
Stop Online Violence Against Women Toolkit
A PDF resource that explains how women are threatened and silenced in online spaces, with a particular focus on Twitter. The document also provides **the general public/allies** several ways to take action against the continued abuse of women on Twitter, along with information about how to organize on college campuses and engage with the media. **TW Note:** This resource includes graphic examples of online verbal threats.

Troll Patrol Project
A completed research project in partnership with Element AI, a global artificial intelligence software product company. More than 6,500 volunteers (called Decoders) analyzed and categorized 288,000 unique tweets mentioning 778 women politicians and journalists from the UK and US. The data was then used to draw findings about the prevalence and nature of these abusive messages.

Campus and Community Flyers

A Google Document that **campus and community organizers** can print and distribute to educate people on the online abuse of women. **TW Note:** This resource includes graphic examples of online verbal threats.

#ToxicTwitter — Violence and Abuse Against Women Online

A multimedia report on the harassment of women on Twitter, incorporating quantitative and qualitative research as well as interviews with targets of online harassment. The report discusses the intersectional nature of the problem, nonconsensual pornography, and how female activists and politicians face greater intimidation, threats, and abuse online. It also examines Twitter's reporting process and the inconsistent enforcement of its standards.

Latest Initiatives or Public Activity

- March 2019: Promoting the Toxic Twitter Report on Twitter
- Actively updates website and social media profiles

Anti-Defamation League (ADL)

Website
https://www.adl.org/

About "Since 1985, ADL has been a pioneer in confronting cyberhate. We have partnered with industry leaders and legal experts to identify and remove online hate speech. In 2016, we formed the ADL Center on Technology & Society to double down on our strategic investments to advance these efforts. Working closely with industry leaders such as Facebook, Twitter, Instagram and YouTube, in 2014 we introduced Best Practices to ensure threats and offensive content violating their community guidelines are taken offline. We have also been instrumental in game-changing breakthroughs like the formation of the Twitter Trust and Safety Council."

Founded in 1913, online harassment and hate speech are just two of the many issues that the ADL addresses.

Products, Solutions, or Services

ADL Cyber — Safety Action Guide

A web-based list of online providers and platforms. People can click on the relevant platform to see its hate speech, cyberbullying, or harassment policy. Each entry also includes the ways that a **user** can report harmful content to the platform or provider. Platforms and providers covered include:

Amazon

AT&T

Bumble

Cafe Press

Xfinity

Craigslist

eBay

Facebook

GoDaddy.com

Google

Instagram

LinkedIn

PayPal

Pinterest

Reddit

Topix

Twitter

Vimeo

WordPress

Yahoo!

YouTube

Bullying and Cyberbullying Prevention Strategies and Resources

A list of resources addressing both offline bullying and cyberbullying, particularly useful for **targets of online harassment, schools, parents,** and the **general public.** The list is categorized according to the audience (students, teachers and administrators, families and caregivers), along with links to statistics and statutes. Topics include:

- Popular mobile apps
- How students can be allies against bullies/cyberbullies
- Cyberbullying warning signs
- Guides to identifying and responding to bullying/cyberbullying

Bullying and Cyberbullying Workshops

Workshops and presentations **for students, families, and administrators.** The programs are designed to help schools identify, address, and prevent both offline bullying and cyberbullying.

Resource Library — Cyberhate Section

Within ADL's extensive resource library (which includes more than 500 resources), **people** can find articles about online harassment in the Cyberhate Section. Topics include:

- Internet Guidelines for Families
- How terrorists are using social media
- Digital security
- Online hate speech and harassment

Latest Initiatives or Public Activity

- March 2019: ADL announced its partnership with Network Contagion Research Institute to study the spread of hate and extremism on social media.
    - The first report from the partnership studies how the platform Gab is being used by far-right extremists in response to Twitter's attempts at moderation
- March 2019: Computational Propaganda and the Democratic Process
- February 2019: Online Hate & Harassment Survey and Report

BADASS

Website

https://badassarmy.org/

About "BADASS is a nonprofit organization dedicated to providing support to victims of revenge porn/image abuse, and eradicating the practice through education, advocacy, and legislation." The organization also provides resources and tools to empower targets.

Founder Katelyn Bowden — Active

Products, Solutions, or Services

List of State Nonconsensual Pornography Laws

A list of laws governing non consensual pornography in 43 states and the District of Columbia, particularly useful for **targets of non-consensual pornography** and **legal professionals.**

Facebook Group for Survivors

A private Facebook group for **targets of nonconsensual pornography.** Users can get advice and support along with education about digital security, legal recourse, and how to get images removed from sites.

Speaking Program for Schools and Workplaces
Founder and CEO Katelyn Bowden speaks to **organizations and schools** about nonconsensual pornography and online abuse.

Latest Initiatives or Public Activity

- February 2019: How Catherine Bosley Inspired BADASS
- January 2019: MTV Partnership - BADASS Takes MTV
- January 2019: BADASS Year in Review

  "We aided in dozens of arrests for NCP, telecommunications harassment, stalking, and child pornography, and have several very large cases awaiting completion. We continue to aid law enforcement in the collection of evidence and prosecution of internet sex crimes, and are working to give the police and investigators the tools and knowledge they need to ensure justice for those experiencing NCP. In 2018, we began partnering with social media platforms and individual sites to help them keep their platforms free of image abuse. We've shared our findings, our experiences, and our ideas to several large social media platforms, and helped them reinforce and refine the protections they have in place to prevent NCP."
- Active on social media

Civilination

Website
https://civilination.org/

About "CiviliNation provides both strategic direction as well as tactical tools to foster an online culture in which individuals can fully engage and contribute without fear or threat of being the target of unwarranted abuse, harassment, or lies... CiviliNation offers a number of guest lectures and training programs for organizations, universities, and online communities looking to improve people's engagement online."

Founder Andrea Weckerle — Active

Products, Solutions, or Services
CiviliNation Academy for Online Conflict Management
A library of short educational videos on topics pertaining to online harassment, geared for the **general public, targets of online harassment, and online communities.** Topics include:

- Dating and Sex on the Web
- Online Reputation
- Legal Issues

- Protecting Yourself Online
- Online Communities and Culture
- Types of Online Attacks

Attorney List
A list of attorneys with knowledge and experience with online harassment.

List of Resources and Tips
A list of tools, tips, and resources to help **targets of online harassment** and **the general public** protect themselves from or deal with online harassment. Includes related nonprofit organizations, secure communication tools, and guides to online safety and digital security.

CiviliNation Harassment Barometer — last update was March 2016
A research project in partnership with the Advertising Benchmark Index, intended to monitor the scope and state of online harassment.

Weekly Newsletter
A weekly email newsletter about the online space, including news articles and stories.

Training at Organizations and Schools
Guest lectures and training programs for schools, universities, and organizations to teach **employers, organizations, schools, and universities (employees, staff, students)** how to handle online conflict and create healthy online spaces.

Civility in the Digital Age: How Companies and People Can Triumph over Haters, Trolls, Bullies and Other Jerks
A book by the organization's founder that discusses online conflict, reputation, privacy, and legal issues.

Latest Initiatives or Public Activity

- July 2018: Latest news post on the organization's website
- Currently active on social media

Crash Override

Website
http://www.crashoverridenetwork.com/

About Crash Override is a crisis helpline [now inactive], advocacy group and resource center for people who are experiencing online abuse. "We are a network of experts and survivors who work directly with victims, tech companies, lawmakers, media, security experts, and law enforcement to educate and provide direct assistance working to eliminate the causes of online abuse."

Founders Zoe Quinn and Alex Lifschitz — Both inactive

Products, Solutions, or Services

COACH
"Crash Override's Automated Cybersecurity Helper" walks users through the steps of securing their online presence as a digital self-defense strategy. It's designed to offer bite-sized, easy-to-follow best practices to the **general public and targets of online harassment.**

DIY Security Guides
Web-based guides that walk the general public and targets of online harassment through the important elements of digital security and aspects of online harassment. Topics include:

- Preventing Doxxing
- So You've Been Doxxed: A Guide On What to Do Next
- Account Security 101: Passwords, Multifactor, Social Engineering, and You:
- Talking to Family and Police

Info Sheet for Employers
A one-page overview that teaches **employers and organizations** about online harassment. Topics include:

- What is online abuse?
- Why does it happen?
- What should employers expect?
- What employers can do

Informal Study on the Relationship Between Identity and Online Abuse

A Google Forms survey that allows **targets of online harassment** to submit information about their online harassment experience. This includes:

- Consent for use and distribution of information along with further contact from researchers
- Aspects of identity such as race, disability status, religion, sexuality, gender, and size
- Types of harassment experienced
- Identity of harasser
- Scale/frequency of harassment
- Platforms where abuse took place
- Whether user reported abuse to platform and what response they received
- Whether user reported abuse to law enforcement and what response they received
- Whether user reported abuse to law enforcement and what response they received
- Offline consequences of online harassment
- Culturally competent responses and help
- Personal experiences and case studies

Latest Initiatives or Public Activity

2016 — Although their DIY resources live on, as of December 2016, Crash Override no longer provides a hotline. From their website, "Due to overwhelming need for assistance with online abuse outpacing our current resources, the hotline is temporarily suspended." Read the full statement

Cyber Civil Rights Initiative (CCRI)

Website
https://www.cybercivilrights.org/

About CCRI fights nonconsensual pornography and other forms of online abuse through:

- Support and referral services for targets
- Advocacy and policy-making legislation
- Collaboration with the tech industry to develop design-based solutions
- Education of legal professionals, law enforcement, policymakers, and the general public about the nature and prevalence of online abuse

Founder Dr. Holly Jacobs — Active

Products, Solutions, or Services

CCRI Helpline
A crisis helpline for **targets of nonconsensual pornography** seeking support and guidance. 24/7 support is available in the U.S. only. Translation available in most languages.

Online Removal Guide
A web-based guide to reporting and requesting removal of nonconsensual pornography from most major online platforms. Includes links to platform forms and help articles. Platforms addressed:

- Google
- Microsoft
- Yahoo
- Facebook
- Instagram
- Twitter
- Snapchat
- Reddit
- Tumblr

List of “Revenge Porn” Laws
A list of laws governing nonconsensual pornography in 43 states and the District of Columbia. This is particularly useful **for legal professionals and targets of non-consensual pornography**.

List of Attorneys
A list of attorneys who have volunteered to **help targets of nonconsensual pornography** on a pro- or low bono basis. List is organized by state or territory and includes entries for the United Kingdom.

2017 Nationwide Online Study of Non-consensual Porn Victimization and Perpetration
A report summarizing the findings of a nationwide study by CCRI about the prevalence of online harassment and its health effects, motives, and prevention methods. Particularly useful for **academics/researchers or journalists/media organizations.**

2017 Research Report Infographic
An infographic for **journalists/media organizations** and the **general public** that summarizes the key findings from the 2017 study (above).

Latest Initiatives or Public Activity

- November 2018: Founder Dr. Holly Jacobs - 2018 L’ORÉAL PARIS Woman of Worth
- Active on social media

Dangerous Speech Project

Website
https://dangerousspeech.org/

About “The Dangerous Speech Project (DSP) was created to test a simple, original idea: that a particular type of public speech tends to catalyze intergroup violence, and that this knowledge might be used to prevent such violence.” DSP works in collaboration with academics, practitioners, other organizations, and platforms to diminish dangerous speech and prevent violence.

Founder Susan Benesch — Active

Products, Solutions, or Services
Dangerous Speech: A Practical Guide
A web-based text that defines and explains dangerous speech, including a framework for identifying it, common tactics, and historical examples. This guide can be useful for **advocacy organizations and agencies, journalists/media organizations, academics/researchers, or the general public.**

- "This guide, a revised version of an earlier text (Benesch, 2013) defines Dangerous Speech, explains how to determine which messages are indeed dangerous, and illustrates why the concept is useful for preventing violence. We also discuss how digital and social media allow Dangerous Speech to spread and threaten peace, and describe some promising methods for reducing Dangerous Speech – or its harmful effects on people."

What is Dangerous Speech? [video]

In this short video, Susan Benesch explains what dangerous speech is, how it differs from hate speech, and what five factors determine whether something is dangerous speech.

List of External Resources

A set of articles, guides, reports, and other resources offering information, insights, and practical steps for identifying, preventing, and responding to dangerous speech and hate speech. Topics include:

- Online harassment
- Inauthentic online content (deep fakes)
- The political and democratic consequences of online misogyny
- Counterspeech
- Algorithms
- Global examples of dangerous speech
- Hate speech on social media
- Propaganda

Counterspeech Center

A section of dangerousspeech.com dedicated to educating **users** on what counterspeech is, why it matters, and how they can participate in it. Along with describing counterspeech, it includes a collection of articles, graphics, and research studies. Topics include:

- Leaders and counterspeech
- Examples of successful counterspeech
- Counterspeech on Twitter
- Anti-muslim online bigotry
- Counterspeech as a method of preventing genocide

Latest Initiatives or Public Activity

- February 2019: Insights into elections or conflicts that could be affected by dangerous speech
- November 2018: Hosted the First International Counterspeakers' Workshop
    - "A principal goal of the workshop was to establish a list of best practices for counterspeech – a framework for understanding strategy, tactics, security measures, and guiding principles."
- April 2018: Launched a study in partnership with Twitter to understand how visibly displaying platform rules affects abuse

Equality Labs

Website

https://www.equalitylabs.org/

About Equality Labs is a South Asian organization that uses community research, socially engaged art, and technology to end the oppression of caste apartheid, Islamophobia, white supremacy, and religious intolerance. Equality Labs trains people in digital security, helping people and partner organizations understand their rights and keep themselves safe online.

Founder Thenmozhi Soundararajan - Active

Products, Solutions, or Services

Rapid Response Request Form

Targets of online harassment (both individuals and organizations) can request help with a hacking incident, doxxing, digital security training, or audit. Equality Labs will respond within 24 hours.

Digital Self-Defense Curriculum

A collection of one sheeters with tips and information about various aspects of digital security made for **targets of online harassment, activists,** or **organizations.** Topics include:

- Mobile device security
- Computer security
- Network security
- Secure digital communication tools
- Identity theft and impersonation

Digital Security Videos

Videos under 5 minutes with tips and information about **digital security for targets of online harassment, activists,** or **organizations.** Topics include:

- Mobile device security
- Computer security
- Network security

Anti-Doxing Guide for Activists Facing Attacks from the Alt-Right

A guide to securing digital presence from doxxing attempts, for **targets of online harassment and activists.** Topics and tactics discussed include:

- Self-Care
- Incident Tracking and Documentation
- Passwords
- 2 Factor Authentication
- Data Brokers
- VPNs and Tor Browser
- Secure Communications Tools
- Encrypted Email
- Social Media
- Hardware Security

Latest Initiatives or Public Activity

- August 2018: Participation in Netroots Nation 2018

Feminist Frequency

Website
https://feministfrequency.com/

About “Feminist Frequency is a not-for-profit educational organization that analyzes modern media’s relationship to societal issues such as gender, race, and sexuality... We strongly advocate for the just treatment of all people online and believe that media is an essential tool for eradicating injustice. Through consciousness-raising around issues like online harassment, we hope to cultivate new media literacies that will make us all more responsible media users in a just and more equitable virtual world.”

Founder Anita Sarkeesian — Active

Products, Solutions, or Services
Speak Up & Stay Safe(r): A Guide to Protecting Yourself From Online Harassment
A web-based guide to self defense from harassment for **the general public** and **targets of online harassment.** Primarily focused on digital, but also addresses offline mail. Topics include:

- Doxxing
- Passwords
- Online gaming
- Offline mail
- Social media
- Device security

Latest Initiatives or Public Activity

- January 2019: Released a 2018 Annual Report
- July 2018: Online Safety Guide (above) updated
- Active on social, but mostly non-harassment related content

Heartmob

Website
https://iheartmob.org/

About HeartMob is a project of Hollaback!, a grassroots nonprofit organization powered by a global network of local activists to end harassment in public spaces. HeartMob provides direct support to targets of online harassment, with the goal to reduce trauma and, ultimately, "transform the hearts and minds of those perpetuating online harassment."

Founders Co-founded by Hollaback! leaders Jae Cameron, Jill Dimond, Emily May, Debjani Roy, and Courtney Young — not all active (Dimond and Roy no longer with the org)

Products, Solutions, or Services
Bystander training — inactive
An online training module for bystanders witnessing online harassment. Particularly useful for **online communities and their members** and the **general public.**

HeartMob Platform for Targets and Bystanders;
A platform for **targets of online harassment** to share their experiences and ask for help or support. **Bystanders** can also use the platform to offer support.

The Heartbot — last activity was in July 2018
A Twitter bot that, when tagged by a **target of online harassment** or **bystander/witness** in reply to an abusive tweet, will donate $1 to HeartMob and notify the harasser.

Online Harassment Resource Guides
Web-based guides about digital security and dealing with online harassment; for **targets of online harassment; employers, organizations or companies; the general public.** Topics include:

- Technical/Digital Safety and Security
- Legal Recourse
- Self Care
- Supportive Organizations
- Organizational Infrastructure
- Social Media

Study on Effects of Classification and Heartmob Platform

An academic paper about the effects of the HeartMob support platform on users. Topics explored:

- Classification and Labeling of Online Harassment
- Intersectional Feminist Theory
- Community Norms
- Effects on Targets
- Effects on Bystanders

Latest Initiatives or Public Activity

- May 2018: Lindsay Blackwell presents her research on HeartMob
- Social media accounts seem to have been deactivated

National Network to End Domestic Violence—

Tech Safety Net Project

Website

https://www.techsafety.org/ and https://nnedv.org/content/technology-safety/

About "NNEDV's Safety Net project focuses on the intersection of technology and domestic and sexual violence and works to address how it impacts the safety, privacy, accessibility, and civil rights of victims by:

- Working with communities, agencies, and technology companies to address how current and emerging technology impacts the safety, privacy, and accessibility rights of victims.
- Educating victim advocates and the general public on ways to use technology strategically to increase and maintain safety and privacy.
- Training law enforcement and justice systems, social services, coordinated community response teams and others on tactics of technology misuse and offender accountability.
- Advocating for strong local, state, national and international policies that ensure the safety, privacy and civil rights of all victims and survivors."

Founder Cindy Southworth — Active

Products, Solutions, or Services

NNEDV's Annual Technology Summit

A training summit for **legal professionals, cause advocates and organizations, and law enforcement** on the intersections of technology and domestic violence, sexual assault, stalking, and trafficking.

Online Toolkits

A collection of web-based toolkits for targets of **online harassment, advocation orgs or agencies, and legal professionals** to help groups address online harassment and domestic violence. Includes guides, articles, and other resources. Collection includes:

- Survivors Toolkit
- Agency Toolkit
- Confidentiality Toolkit
- Legal Systems Toolkit

Online Dating: Survivor Privacy Risks & Strategies Guide

A web-based guide (also available as a downloadable PDF) that offers strategies to make online dating safer **for survivors of domestic violence / harassment** and **the general public**. It examines the privacy risks, best practices for in-person meetings, and how users can report inappropriate interactions.

Online Gaming: Survivor Privacy Risks & Strategies Guide

A web-based guide (also available as a downloadable PDF) that provides an introduction to the risks and benefits of online gaming and tips f**or survivors of domestic violence / harassment** and **the general public.**

Documentation Tips for Survivors of Technology Abuse & Stalking

A list of tips and best practices for documenting technological abuse and harassment. Topics include:

- Email
- Text messages
- Phone calls
- Social media

Images, Consent & Abuse Guide

A web-based guide (also available as a downloadable PDF) on how intimate images are being used to threaten, harass, and abuse targets. The guide discusses popular terminology like "revenge porn" and explains what **targets of online harassment** can do about it, including opportunities for legal action.

Tech Safety

An educational app that teaches **targets of online harassment** how to strengthen digital security and privacy.

Latest Initiatives or Public Activity

- July 2019: Attendance at 2019 Technology Summit
    - "We are looking forward to expanding our work to help enhance safety protections for programs and survivors and what #TechSafety means for all."
- January 2019: Data Privacy Day 2019: Location Data & Survivor Safety
- December 2018: Safety Net 2018: Looking Back and Moving Forward

OnlineSOS

Website

https://onlinesos.org/

About OnlineSOS supports individuals facing online harassment with actionable resources and advocacy. Our aim is to empower people, particularly journalists, with trusted information to take action. We are also shaping and advancing the conversation around online harassment and abuse. We conduct research, host convenings, and advocate for individuals in an effort to catalyze collaboration among stakeholders.

Founder Liz Lee — Active

Products, Solutions, or Services

Action Plans

A series of web-based resources (downloadable as PDFs) **for targets of online harassment** to identify what they're experiencing and understand immediate and other next steps to take in response. Includes links to additional articles and resources for targets. There are also Action Plans to help users protect themselves online, document your harassment, and manage your professional and emotional wellbeing.

Collection includes:

- Doxxing
- Distribution of False Information (Defamation)
- Mob Harassment
- Distribution of Intimate Images (Nonconsensual Pornography)
- Digital Security
- Emotional Wellbeing
- Proper Documentation (of harassment)
- 2 Factor Authentication (2FA)
- Communicating With Your Employer and Colleagues
- Threat Modeling

Resource List

A list of organizations, legal professionals, and other resources **for targets of online harassment** with a focus on the needs of journalists. People can find helplines for direct service, identity or population-specific support, and resources designed for journalists dealing with online harassment. Resources/Topics include:

- Emergency Funds for Journalists
- Relevant Legal Professionals and Organizations
- Digital and Technical Security Guides and How-Tos
- Mental Health Support
- Resources for Members of the LGBTQ Community
- Professional Organizations for Journalists
- Resources Outside the U.S.
- Research and Reports on Online Harassment

OnlineSOS Blog

Up-to-date tips, news, information, expert interviews, and analysis about online harassment and related issues. Blog includes the Account Safety Cheat Sheet, a one-page PDF, to review where your personal and professional information might be stored in case you'd like to take privacy and security precautions.

Direct Services for Individuals and Journalists — no longer available

Free, professional mental health services and case management for people in crisis and recovery from online harassment and abuse — including a contingency fund. Case work was requested via email helpline, form, and Facebook chatbot (which also could be used as a documentation tool). User research was conducted on the needs of journalists, resulting in the development of a helpline and tailored resources. Repository of real cases and resolutions was developed as a result of case work.

Latest Initiatives or Public Activity

- Fall 2019: Organizing a taxonomy workshop
- Summer 2019: Launching legal research project on stalking cases in online-based relationships
- Spring 2019: Releasing a report about the state of online harassment
- March 2019: Is It Really Just Trolling?
- Active on social media

Stop Online Violence Against Women

Website
https://stoponlinevaw.com/

About "Stop Online Violence Against Women raises awareness and funding to stop online harassment. SOVAW addresses inadequate laws and policies that lack protections for women in particular women of color. SOVAW [provides resources and options] for women and women of color, based on their level of harassment or violence, and hopes to include diverse stories of women who are willing to share their experiences. We highlight and include partnering organizations, legislators & companies who are working together to address this important issue."

Founder Shireen Mitchell — Active

Products, Solutions, or Services
Report on Voter Suppression Through Facebook Ads
An analysis of 3,500 Facebook ads by the Russian Internet Research Agency. The report explains how these ads, released to the public by congress, were used to target black voters and influence them to refrain from voting.

Latest Initiatives or Public Activity

- February 2019: Help Prevent Online Harassment of Women - Fundraising Campaign
- Active on Twitter

Trollbusters

Website
http://www.troll-busters.com/

About TrollBusters is a direct service organize that helps targets of online harassment figure out what to do next, including reporting trolls, learning digital security, and monitoring digital presences. "We are Team TrollBusters. When you spot online violence, online abuse or other troll behavior, send an S.O.S. and we will be your first responders online, supporting you with personal endorsements, just-in-time coaching, and reputation repair services."

Founder Dr. Michelle Ferrier — Active

Products, Solutions, or Services

Research
TrollBusters has developed reports and survey data around the prevalence and harmful effects of online harassment on female journalists and other targets. These are particularly useful for **academics and researchers or media organizations/ employers. Examples include:**

- 2018: Attacks and Harassment: The Impact on Female Journalists and Their Reporting
- 2016: The Progression of Hate

TrollBusters Chatbot - currently defunct (as of writing in March 2019)
A chatbot for people seeking help for online harassment.

Helpline Form
Targets of online harassment or bystanders can report instances of online harassment. In return, TrollBusters can provide "positive messages, virtual hugs, or reputation repair services."

What to Do Infographic
A downloadable infographic for **targets of online harassment** that describes various online harassment tactics and offers steps to respond and/or deal with the situation. It is available in English, Hindi, Turkish, Russian, and Spanish.

Tactics described include:

- DoS
- Doxxing
- Nonconsensual pornography
- Threats
- Libel
- Impersonation

Global Safety Resource Hub

A Google Maps tool that displays organizations addressing online harassment. Documents resources across North America, South America, Europe, Africa, and the Middle East, which can be useful for **targets of online harassment** across the globe, as well as **researchers, other groups, and stakeholders like social media platforms.**

Resources for Journalists List

A list of resources and organizations **for journalists** facing online harassment. The list is organized by situation/tactic.

Resources and topics covered include:

- Hacking and DoS
- Doxxing
- Nonconsensual pornography
- Technical security
- Libel
- Information for non-journalists

Digital Hygiene Course

A collection of 16 learning modules to educate **the general public, targets of online harassment, or high-risk individuals (like journalists)** about digital security**.**

Samples of modules include:

1. Privacy protection on domain names
2. Anonymous "Tor" cloak or VPN
3. Prepare for a DDos attack
4. Two-step verification
5. Privacy plug-ins/cookies
6. Image "hidden pixels"
7. Links and attachments
8. Install patches and updates
9. Use a password manager/strong password
10. Encryption

Latest Initiatives or Public Activity

- Fall 2018 — Report in conjunction with the International Women's Media Foundation: Attacks and Harassment: The Impact on Female Journalists and their Reporting
- Maintain a Paper.li called Troll Tracker

Without My Consent

Website
https://withoutmyconsent.org/

About Without My Consent empowers victims of egregious online privacy violations to lead the fight against online harassment. WMC trains lawyers, law enforcement, and advocates; engages in solutions-oriented work with industry and government; and develops education materials. WMC specializes in support for targets of non-consensual distribution of intimate images (also known as non-concensual pornography or NCP).

Founder(s) Erica Johnstone and Colette Vogele — Both active

Products, Solutions, or Services
Something Can Be Done Guide for **targets of non-concensual pornography**
A web-based guide for targets of non-consensual pornography designed to aid in the legal and takedown process. Topics include: Evidence preservation; Talking to lawyers; Emotional health/care; Restraining order requests; Copyright registration; Content takedowns

Digital Abuse Restraining Order Cheat Sheet for **targets of non-concensual pornography**
A summary of California's Domestic Violence Prevention Act and information and sample language to aid in the request process for targets of nonconsensual pornography.

50 State Project for **targets of non-concensual pornography**
An overview of the laws and legal recourse for nonconsensual pornography in each of the 50 states and the District of Columbia. Provides information on how to take legal action under a pseudonym.

Email Roundup for **individuals** and **legal professionals**
A weekly email newsletter relating to the organization, online harassment and, specifically, non-consensual pornography.

Lawyer Training Module for **legal professionals**
A training presentation or module to educate lawyers on the paths to justice for people experiencing nonconsensual pornography. Includes training about: Client intake; Court orders and settlements; Legislation; Technology used for NCP; Terminology; High-profile cases; Harm caused by NCP

Latest Initiatives or Public Activity

- April 2018 — A New Advisory Helps Domestic Violence Survivors Prevent and Stop Deepfake Abuse

Women's Media Center — Speech Project

Website
http://www.womensmediacenter.com/speech-project

About The WMC Speech Project works collaboratively across organizations and institutions to raise awareness about online harassment and its effect on women's civic and political participation. The WMC Speech Project is dedicated to expanding women's freedom of expression and curbing online harassment and abuse.

Founders Soraya Chemaly and Ashley Judd — Active

Products, Solutions, or Services

WMC Speech Project
Women's Media Center's hub for online harassment resources and articles. The project's goal is to increase the understanding of online misogyny in order to find solutions that allow women to speak freely on the internet. The organization works collaboratively with:

- Advocacy organizations and agencies
- Journalists/media organizations
- Academics/researchers
- Targets of online harassment

Online Abuse 101 Wheel
A graphic that shows how different forms of online harassment are related to various legal issues and offline consequences. The webpage also includes a glossary of the terms used on the graphic for better understanding among **advocacy organizations, journalists / media,** and **the general public.**

Tools and Resources List
A set of tools and resources **for targets of online harassment**, including online guides, supportive organizations, and where to get immediate help.

Research and Statistics Page
Statistics and research findings about online misogyny and online harassment. Topics include:

- Who experiences abuse?
- Who's doing the harassing?
- Intersectionality and Abuse
- Impacts on Users
- Offline Consequences
- The Benefits of Online Engagement

Latest Initiatives or Public Activity

November 2017: The Speech Project's most recent article

However, the Women's Media Center is current and active

Appendix B
A Short List of Online Harassment that Resulted in Violence and Fatalities

This is not a complete list of all cases of real-world harm caused by online harassment, nor does it include cases of bullying among minors. And, of course, there is a limit to what is published in the media and to what can be solely attributed to online harassment. The goal of this resource is to preserve the memory of those affected by online harassment and provide readers with a point of reference to learn about their personal stories.

Lauren McCluskey, 2018

Capital Gazette Tragedy, 2018

Gerald Fischman, Robert Hiaasen, John McNamara, Rebecca Smith, Wendi Winters

Andrew Finch, 2017

Christina Grimmie, 2016

Alice Ruggles, 2016

Christine Belford and Laura Mulford, 2013

Rehteah Parsons and Audrie Potts, 2013

Tyler Clementi, 2010

Epilepsy Foundation Website Hijack, 2008

Appendix C
Research Papers

Click here for a directory of research papers about online harassment and related topics.

Endnotes and Bibliography

Endnotes

1. Maeve Duggan, "Online Harassment 2017," Pew Research Center, July 11, 2017, https://www.pewinternet.org/2017/07/11/online-harassment-2017/

2. This definition of "trolling" is a summary of how we've observed the term used today across mainstream media outlets. For more context and background, please visit https://onlinesos.org/blog/what-trolling-means-online-harassment

3. John Seabrook, "My First Flame," The New Yorker, June 06, 1994, http://www.johnseabrook.com/my-first-flame/

4. Alexander Abad-Santos, "Twitter's Report Abuse Button is a Good, But Small First Step," The Atlantic, July 31, 2013, https://www.theatlantic.com/technology/archive/2013/07why-twitters-report-abuse-button-good-tiny-first-step/312689/

5. Sarah Jeong, The Internet of Garbage, Vox Media, 2018, pp. 16-17, https://cdn.vox-cdn.com/uploads/chorus_asset/file/12599893/The_Internet_of_Garbage.0.pdf

6. Casey Johnston, "Chat logs show how 4chan created #GamerGate controversy," Ars Technica, Sept 10, 2014, https://arstechnica.com/gaming/2014/09/new-chat-logs-show-how-4chan-users-pushed-gamergate-into-the-national-spotlight/

7. Casey Newton and Nilay Patel, "Caroline Sinders Podcast Interview," The Verge, December 12, 2018, https://www.theverge.com/2018/12/4/18125699/vergecast-podcast-interview-caroline-sindersonline-harassment-catalogue-gamergate

8. Bailey Poland, Haters: Harassment, Abuse, and Violence Online, p. 58, 124, November 1, 2016.

9. Ryan Broderick, "Activists Are Outing Hundreds Of Twitter Users Believed To Be 4chan Trolls Posing As Feminists," Buzzfeed News, June 17, 2014. https://www.buzzfeednews.com/article/ryanhatesthis/your-slip-is-showing-4chan-trolls-operation-lollipop

10. Thanks to Shafiqah Hudson for unearthing this Twitter Moment.

11. Shireen Mitchel, Twitter Moment, April 22, 2018, https://twitter.com/i/moments/988069560921284608 More of Shireen's work can be found at digitalsista.me

12. Leigh Alexander, "Online abuse: how women are fighting back," The Guardian, April 13, 2016, https://www.theguardian.com/technology/2016/apr/13/online-abuse-how-women-are-fighting-back. More of Jamia's work can be found on her website at jamiawilson.com

13. Adrian Chen, "The Agency", The New York Times, June 02, 2015, https://www.nytimes.com/2015/06/07/magazine/the-agency.html

14. Julia Carrie Wong, "Anti-vaxx 'mobs': doctors face harassment campaigns on Facebook," The Guardian, February 28, 2019, https://www.theguardian.com/technology/2019/feb/27/facebook-anti-vaxx-harassment-campaigns-doctors-fight-back

15. Data and Society, “Media Manipulation and Disinformation Online”, May 15, 2017. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

16. Lee Ferran, Trish Turner, and Katherine Faulders, “Russia Targeted African-American Vote, Made Instagram ‘Key Battleground’ in Propaganda War: Researchers,” ABC News, December 17, 2018, https://abcnews.go.com/Politics/russia-targeted-african-american-vote-made-instagram-key/story?id=59862038

17. Jamiles Lartey, “Race and Russian interference. Senate report details age-old tactic,” The Guardian, December 24, 2018, https://www.theguardian.com/world/2018/dec/24/race-russian-election-interference-senate-reports

18. “Worldwide Roundup of Journalists Killed, Detained, Held Hostage, or Missing in 2018,” Reporters Without Borders, Accessed on March 08, 2019, https://rsf.org/sites/default/files/worldwilde_round-up.pdf

19. Michael M. Grynbaum, “Trump Discusses Claims of ‘Fake News,’ and Their Impact, with New York Times Publisher,” The New York Times, February 02, 2019, https://www.nytimes.com/2019/02/01/business/media/donald-trump-interview-news-media.html

20. “Human rights experts denounce Trump’s attacks against media,” UN News, August 2, 2018, https://news.un.org/en/story/2018/08/1016222

21. Stephanie Sugars, “An Anatomy of Trump’s Tweets,” Committee to Protect Journalists, Jan 30, 2019, https://cpj.org/x/7638

22. “RDF Publishes Report on Online Harassment of Journalists,” Reporters Without Borders, August 01, 2018, https://rsf.org/en/news/rsf-publishes-report-online-harassment-journalists

23. Maeve Duggan, “Online Harassment 2017,” Pew Research Center, July 11, 2017, https://www.pewinternet.org/2017/07/11/online-harassment-2017/

24. Bailey Poland, Haters: Harassment, Abuse, and Violence Online, p. 173, November 01, 2016.

25. Rachel Thompson, "She will not be silenced," Mashable, June 18, 2018, https://mashable.com/2018/06/18/online-harassment-trolling-women/

26. ADL Report on Online Harassment, 2019, https://www.adl.org/onlineharassment

27. Bailey Poland, Haters: Harassment, Abuse, and Violence Online, p. 52, November 01, 2016.

28. Roni Jacobson, “I’ve Had a Cyberstalker Since I was 12,” Wired, February 29, 2016, https://www.wired.com/2016/02/ive-had-a-cyberstalker-since-i-was-12/.

29. Vegas Tenold, “To Doxx a Racist,” The New Republic, July 26, 2018, https://newrepublic.com/article/150159/doxx-racist

30. Paul Tassi, “Overwatch’s ‘Fake’ Female Player Ellie Scandal,” Forbes, January 06, 2019, https://www.forbes.com/sites/insertcoin/2019/01/06/overwatchs-fake-female-player-ellie-scandal-is-the-mess-that-keeps-on-giving/#5ea69dc628a0

31. Stephen Witt, "The Rise and Fall of a Hip-Hop Supervillain," Rolling Stone, January 16, 2019, https://www.rollingstone.com/music/music-features/tekashi-69-rise-and-fall-feature-777971/

32. David Ingram, "A Grindr harassment suit could change the legal landscape for tech — and free speech," NBC News, January 05, 2019, https://www.nbcnews.com/tech/tech-news/grindr-harassment-suit-could-change-legal-landscape-tech-free-speech-n954976

33. Brendan Koerner, "It Started as an Online Gaming Prank. Then It Turned Deadly," Wired, October 23, 2018, https://www.wired.com/story/swatting-deadly-online-gaming-prank/

34. Patrick Devitt, "13 Reasons Why and Suicide Contagion," Scientific American, May 08, 2017, https://www.scientificamerican.com/article/13-reasons-why-and-suicide-contagion1/

35. Jaclyn Schildkraut, "The media should stop making school shooters famous," Vox, February 22, 2018, https://www.vox.com/the-big-idea/2018/2/22/17041382/school-shooting-media-coverage-perpetrator-parkland

36. Andy Golder, "17 People Who Just Got Verbally Destroyed," Buzzfeed, May 04, 2017, https://www.buzzfeed.com/andyneuenschwander/17-people-who-got-totally-fuckin-burned

37. Jacob Geers, "The 49 Most Hilariously Savage Burns In Social Media History," Thought Catalog, June 10, 2016, https://thoughtcatalog.com/jacob-geers/2016/06/the-49-most-hilariously-savage-burns-in-social-media-history/

38. More research is needed in this area, but feelings of isolation, guilt, anxiety, and distress are well-documented. See DART Center research on journalists and PTSD, https://dartcenter.org/content/covering-trauma-impact-on-journalists; Short, Emma & Guppy, Andrew & A. Hart, Jacqui & Barnes, James. (2015). The Impact of Cyberstalking. Studies in Media and Communication. 3. 10.11114/smc.v3i2.970.; Jessica Vitak, Kalyani Chadha, Linda Steiner, and Zahra Ashktorab. 2017. Identifying Women's Experiences With and Strategies for Mitigating Negative Effects of Online Harassment. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17). ACM, New York, NY, USA, 1231-1245. DOI: https://doi.org/10.1145/2998181.2998337.

39. Rebecca Sheffler, "Someone posted my phone number on craigslist and said I wanted strange men to rape me," Narratively, June 7, 2017, https://narratively.com/someone-posted-my-phone-number-on-craigslist-and-said-i-wanted-strange-men-to-rape-me/.

40. Eliza Romero, "Online Harassment is No Joke," Medium, July 15, 2018, https://medium.com/@AestheticDistance/online-harassment-is-no-joke-aca4ba34edf6

41. Maeve Duggan, "Online Harassment 2017," Pew Research Center, July 11, 2017, https://www.pewinternet.org/2017/07/11/online-harassment-2017/.

42. "Best Practices for Allies and Witnesses," PEN America, 2018, https://onlineharassmentfieldmanual.pen.org/best-practices-for-allies-and-witnesses/

43. Squadbox was developed by PhD student Amy Zhang at MIT's Computer Science and Artificial Intelligence Lab. https://squadbox.org/ Full disclose, Amy Zhang has also worked with OnlineSOS as an advisor.

44. Jaclyn Brzezinski, "I Was Doxed and It Can Happen to You Too," Women AdvaNCe, July 26, 2018, https://www.womenadvancenc.org/2018/07/26/i-was-doxed-and-it-can-happen-to-you-part-two/

45. For more information about Access Now, please visit their website https://www.accessnow.org/

46. Saima Salim, "How Much Time Do People Spend on Social Media? Research Says 142 Minutes Per Day," Digital Information World, January 04, 2019, https://www.digitalinformationworld.com/2019/01/how-much-time-do-people-spend-social-media-infographic.html

47. Mary Meeker, "Internet Trends Report 2018," Kleiner Perkins, May 30, 2018, https://www.kleinerperkins.com/perspectives/internet-trends-report-2018/

48. This figure includes work hours.

49. Aaron Smith and Monica Anderson, "Social Media Use in 2018," Pew Research Center, March 01, 2018, https://www.pewinternet.org/2018/03/01/social-media-use-in-2018/

50. Charlie Hall, "Anti-Defamation League Study Uncovers a Surge in Online Harassment," Polygon, February 13, 2019, https://www.polygon.com/2019/2/13/18223558/adl-anti-defamation-league-study-twitch-discord-harassment

51. Maeve Duggan, "Online Harassment 2017," Pew Research Center, July 11, 2017, https://www.pewinternet.org/2017/07/11/online-harassment-2017/

52. Amnesty International, "Toxic Twitter", March 2018, https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1/.

53. Josh Lowensohn, "Twitter turns its block function into a mute button, leading to user revolt," The Verge, December 12, 2013, https://www.theverge.com/2013/12/12/5205462/twitter-turns-its-block-function-into-a-mute-button#comments

54. Josh Lowensohn, "Twitter turns its block function into a mute button, leading to user revolt," The Verge, December 12, 2013, https://www.theverge.com/2013/12/12/5205462/twitter-turns-its-block-function-into-a-mute-button#comments

55. "Online Hate and Harassment: The American Experience," Anti-Defamation League, 2019, https://www.adl.org/onlineharassment

56. Sarah Perez, "Google Revamps its Security Checkup Feature with Personalized Suggestions for Your Account," TechCrunch, October 16, 2017, https://techcrunch.com/2017/10/16/google-revamps-its-security-checkup-feature-with-personalized-suggestions-for-your-account/

57. "Twitter Health Metrics Proposal Submissions," Twitter, March 01, 2018, https://blog.twitter.com/official/en_us/topics/company/2018/twitter-health-metrics-proposal-submission.html

58. Brian Petrocelli, "Moderators are the Sword, Now Automod is the Shield," Twitch, December 12, 2016, https://blog.twitch.tv/moderators-are-the-sword-now-automod-is-the-shield-df3d8aae32a9.

59. Sady Doyle, "Why Facebook's Harassment Guidelines Fail to Protect Women," Elle, May 26, 2017, https://www.elle.com/culture/career-politics/news/a45567/facebooks-harassment-guidelines-leaked/

60. Joshua Fruhlinger, "Facebook now hiring positions for 'safety,' 'security,' and 'trust'," Thinknum, 2018, https://media.thinknum.com/articles/facebook-is-hiring-around-a-tumultuous-future-for-privacy/

61. Mark Zuckerberg's Facebook Page, May 3, 2017, https://www.facebook.com/zuck/posts/10103695315624661

62. Sarah T. Roberts, "Commercial Content Moderation: Digital Laborers' Dirty Work," Media Studies Publications, 2016, https://ir.lib.uwo.ca/commpub/12

63. "YouTube Director of Trusted Content, Trust and Safety," Linkedin, Accessed March 02, 2019, https://www.linkedin.com/jobs/view/director-trusted-content-trust-and-safety-at-youtube-1029294667/.

64. Antigone Davis, "Detecting Non-consensual Intimate Images and Supporting Victims," Facebook Newsroom, March 15, 2019, https://newsroom.fb.com/news/2019/03/detecting-non-consensual-intimate-images/

65. While platforms, private companies, think tanks, and organizations conduct research of their own about online harassment, this section focuses specifically on researching coming out of academic institutions.

66. "Online Harassment Resource Guide," December, 8, 2018, https://meta.wikimedia.org/wiki/Research:Online_harassment_resource_guide

67. See, for example, Facebook's various Request for Proposals and Research Awards. https://research.fb.com/programs/research-awards/proposals/content-policy-research-on-social-media-platforms-request-for-proposals/; Google AI Challenge https://ai.google/social-good/impact-challenge/; Twitter Health Metrics Proposal Submission

68. See, for example: Penn State (2018) "Viral Deception, Polarization, and Networks Workshop"; MIT (2016) "High Impact Questions and Opportunities for Online Harassment Research and Action"

69. You can read more about Dr. Kumar's work AI to predict toxicity and harassment online here: https://uk.pcmag.com/news-analysis/120114/this-ai-predicts-online-trolling-before-it-happens

70. Casey Fiesler's Internet Rules Lab (University of Colorado, Boulder) looks at these questions extensively: https://caseyfiesler.com/research-group/.

71. "Cambridge Analytica, Explained," Wired, March 22, 2018, https://www.wired.com/amp-stories/cambridge-analytica-explainer/.

72. Kara Swisher, interview with Jameel Jaffer, Recode Decode, podcast audio, November 19, 2018. https://www.recode.net/2018/11/19/18103081/first-amendment-facebook-jameel-jaffer-freedom-speech-alex-jones-decode-podcast-kara-swisher

73. "Cyber Exploitation," State of California - Department of Justice - Office of the Attorney General, 31 July 2017, oag.ca.gov/cyberexploitation.

74. Katherine M. Clark, "H.R.3067 - 115th Congress (2017-2018): Online Safety Modernization Act of 2017," Congress.gov, July 14, 2017, https://www.congress.gov/bill/115th-congress/house-bill/3067.

75. Jada F. Smith, “As ‘Sextortion’ Proliferates, Victims Find Precarious Place in Legal System”, The New York Times, May 10, 2016, https://www.nytimes.com/2016/05/11/us/sextortion-victims-brookings-institution.html.

76. Cyber Civil Rights Initiative model state and federal statutes can be found here https://www.cybercivilrights.org/ccri-model-federal-law/ and here https://www.cybercivilrights.org/model-state-law/

77. For more information about CCRI and their work, please visit their website at https://www.cybercivilrights.org/

78. Zeynep Tufekci, “Big Platforms Could Change Their Business Models,” Wired, December 17, 2018, https://www.wired.com/story/big-platforms-could-change-business-models/

79. Ariana Tobin, Madeleine Varner, and Julia Angwin, “Facebook’s Uneven Enforcement of Hate Speech Rules Allows Vile Posts to Stay Up,” ProPublica, December 28, 2017, https://www.propublica.org/article/facebook-enforcement-hate-speech-rules-mistakes.

80. Facebook Community Standards, as of March 2019.

81. “Twitch Community Guidelines Updates,” Twitch, February 08, 2018, https://blog.twitch.tv/twitch-community-guidelines-updates-f2e82d87ae58.

82. “Community Health Initiative,” Wikimedia, February 20, 2019, https://meta.wikimedia.org/wiki/Community_health_initiative.

83. Susan Benesch, “Launching today: new collaborative study to diminish abuse on Twitter,” Medium, April 6, 2018, https://medium.com/@susanbenesch/launching-today-new-collaborative-study-to-diminish-abuse-on-twitter-2b91837668cc

84. Jessica Vitak, Kalyani Chadha, Linda Steiner, and Zahra Ashktorab, “Identifying Women’s Experiences With and Strategies for Mitigating Negative Effects of Online Harassment,” Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW 17, 2017, doi:10.1145/2998181.2998337.

85. Amnesty International, “Toxic Twitter”, March 2018, https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1/.

86. “EPIC - State Revenge Porn Policy,” Electronic Privacy Information Center, 2019, https://epic.org/state-policy/revenge-porn/.

87. Bailey Poland, “Why Naming Online Harassment Accurately Matter,” Women’s Media Center, April 10, 2016, http://www.womensmediacenter.com/speech-project/why-naming-online-harassment-accurately-matters.

88. Sam Kashner, “Both Huntress and Prey,” Vanity Fair, October 20, 2014, https://www.vanityfair.com/hollywood/2014/10/jennifer-lawrence-photo-hacking-privacy

89. J. Nathan Matias, “High Impact Questions And Opportunities for Online Harassment Research and Action,” Civic Media, September 07, 2016, https://civic.mit.edu/2016/09/07/high-impact-questions-and-opportunities-for-online-harassment-research-and-action/.

90. Blackwell, Lindsay, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe, "Classification and Its Consequences for Online Harassment," Proceedings of the ACM on Human-Computer Interaction1, no. CSCW (2017): 1-19. doi:10.1145/3134659.

91. Nathaniel Gleicher and Oscar Rodriguez, "Removing Additional Inauthentic Activity from Facebook," Facebook Newsroom, October 11, 2018, https://newsroom.fb.com/news/2018/10/removing-inauthentic-activity/.

92. Sara Griffiths, "Can This Technology Put an End to Bullying?" BBC, February 11, 2019, http://www.bbc.com/future/story/20190207-how-artificial-intelligence-can-help-stop-bullying.

93. Andy Greenberg, "Security Isn't Enough. Silicon Valley Needs 'Abusability' Testing," Wired, January 28, 2019, https://www.wired.com/story/abusability-testing-ashkan-soltani/.

94. Neima Jahromi, "The New Zealand Shooting and the Challenges of Governing Live-Streamed Video," The New Yorker, March 16, 2019, https://www.newyorker.com/tech/annals-of-technology/the-new-zealand-shooting-and-the-challenges-of-governing-live-streamed-video

95. Recode/Decode Podcast. Full Q&A with Nicole Wong. Sept 12, 2018, https://www.recode.net/2018/9/12/17848384/nicole-wong-cto-lawyer-google-twitter-kara-swisher-decode-podcast-full-transcript

96. Roger McNamee, "I Mentored Mark Zuckerberg. I Loved Facebook. But I Can't Stay Silent About What's Happening," Time, January 17, 2019, http://time.com/5505441/mark-zuckerberg-mentor-facebook-downfall/.

97. Facebook Q4 2018 and year-end earning statement, accessed online, https://investor.fb.com/investor-news/press-release-details/2019/Facebook-Reports-Fourth-Quarter-and-Full-Year-2018-Results/default.aspx.

98. Andreas Reventlow. "The chilling effects of online harassment and how to respond", December 6, 2016, https://www.mediasupport.org/chilling-effects-online-harassment-address/

99. Some examples include Facebook's research grant requests and Twitter's public health request for proposals (see endnote 54 for more information).

100. Sarah Jeong, The Internet of Garbage, Vox Media, 2018, p. 77, https://cdn.vox-cdn.com/uploads/chorus_asset/file/12599893/The_Internet_of_Garbage.0.pdf

101. In January 2019, Twitter user @everywhereist asked the women of Twitter about what comment or post resulted in online harassment. The tweet received 1,800 responses and 19,000 likes. You can read the full thread here: https://twitter.com/everywhereist/status/1082813365628526592

102. Information about the Netizens film and future screenings can be found at https://www.netizensfilm.com/

103. One popular TED Talk has been Jigsaw director of research and development Yasmin Green's talk about how technology might counter extremism and online harassment. The video garnered 1.2 million views since April 2018. https://www.ted.com/talks/yasmin_green_how_technology_can_fight_extremism_and_online_harassment

Bibliography

1. Abad-Santos, Alexander. "Twitter's 'Report Abuse' Button Is a Good, But Small, First Step." The Atlantic. October 29, 2013. https://www.theatlantic.com/technology/archive/2013/07/why-twitters-report-abuse-button-good-tiny-first-step/312689/

2. Alexander, Leigh. "Online Abuse: How Women Are Fighting Back." The Guardian. April 13, 2016. https://www.theguardian.com/technology/2016/apr/13/online-abuse-how-women-are-fighting-back

3. Benesch, Susan, Cathy Buerger, Tonei Glavinic, and Sean Manion. "Dangerous Speech: A Practical Guide." Dangerous Speech Project. April 08, 2019. https://dangerousspeech.org/guide/.

4. "Best Practices for Allies & Witnesses." Online Harassment Field Manual. Accessed March 2019. https://onlineharassmentfieldmanual.pen.org/best-practices-for-allies-and-witnesses/.

5. Blackwell, Lindsay, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe. "Classification and Its Consequences for Online Harassment." Proceedings of the ACM on Human-Computer Interaction, no. CSCW (2017): 1-19. doi:10.1145/3134659.

6. Blackwell, Lindsay, Tianying Chen, Sarita Schoenebeck, and Cliff Lampe. When Online Harassment is Perceived as Justified. 2018. https://www.cybersmile.org/wp-content/uploads/When-Online-Harassment-is-Perceived-as-Justified-ICWSM18.pdf.

7. Broderick, Ryan. "Activists Are Outing Hundreds Of Twitter Users Believed To Be 4chan Trolls Posing As Feminists." BuzzFeed News. June 17, 2014. https://www.buzzfeednews.com/article/ryanhatesthis/your-slip-is-showing-4chan-trolls-operation-lollipop.

8. Brzezinski, Jaclyn. "I Was Doxed and It Can Happen To You: Part Two." Women AdvaNCe. July 26, 2018. https://www.womenadvancenc.org/2018/07/26/i-was-doxed-and-it-can-happen-to-you-part-two/.

9. "CCRI Model Federal Law." Cyber Civil Rights Initiative. Accessed March 2019. https://www.cybercivilrights.org/ccri-model-federal-law/.

10. "CCRI Model State Law." Cyber Civil Rights Initiative. Accessed March 2019. https://www.cybercivilrights.org/model-state-law.

11. Chen, Adrian. "The Agency." The New York Times. June 02, 2015. https://www.nytimes.com/2015/06/07/magazine/the-agency.html.

12. Chen, Gina Masullo, Paromita Pain, Victoria Y. Chen, Madlin Mekelburg, Nina Springer, and Franziska Troger. "'You Really Have to Have a Thick Skin': A Cross-cultural Perspective on How Online Harassment Influences Female Journalists." Journalism, 2018, 146488491876850. doi:10.1177/1464884918768500.

13. Citron, Danielle Keats. Hate Crimes in Cyberspace. Cambridge, MA: Harvard Univ Press, 2014.

14. "CivilServant." CivilServant. Accessed April 2019. https://civilservant.io/.

15. Clark, Katherine M. "H.R.3067 - 115th Congress (2017-2018): Online Safety Modernization Act of 2017." Congress.gov. July 14, 2017. https://www.congress.gov/bill/115th-congress/house-bill/3067.

16. "Community Health Initiative." Meta. Accessed March 2019. https://meta.wikimedia.org/wiki/Community_health_initiative.

17. "Community Standards." Facebook. Accessed March 2019. https://www.facebook.com/communitystandards/.

18. "Content Policy Research on Social Media Platforms Request for Proposals." Facebook Research. January 30, 2019. https://research.fb.com/programs/research-awards/proposals/content-policy-research-on-social-media-platforms-request-for-proposals/.

19. Corbett, Erin. "Digital PTSD Is Real." The Outline. May 14, 2018. https://theoutline.com/post/4543/journalists-ptsd-online-harassment-digital-ptsd-alt-right-white-nationalists?zd=1&zi=eonignt6.

20. "Covering Trauma: Impact on Journalists." Dart Center. May 26, 2016. https://dartcenter.org/content/covering-trauma-impact-on-journalists.

21. Cyber Civil Rights Initiative. Accessed March 2019. https://www.cybercivilrights.org/.

22. "Cyber Exploitation." State of California - Department of Justice - Office of the Attorney General. July 31, 2017. https://www.oag.ca.gov/cyberexploitation.

23. Davis, Antigone. "Detecting Non-Consensual Intimate Images and Supporting Victims." Facebook Newsroom. March 15, 2019. https://newsroom.fb.com/news/2019/03/detecting-non-consensual-intimate-images/.

24. "Defining 'Online Harassment': A Glossary of Terms." Online Harassment Field Manual. Accessed April 2019. https://onlineharassmentfieldmanual.pen.org/resource-guide-to-combat-online-harassment/defining-online-harassment-a-glossary-of-terms/.

25. Department for Digital, Culture, Media & Sport. Online Harms White Paper. HM Government, 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf

26. DeRuiter, Geraldine. "Women of Twitter: What's the Most Inconsequential Topic or Opinion You've Expressed That Has Resulted in Online Harassment?I Once Got Rape and Death Threats When I Wrote about the Time I Dyed a Gallon of Milk Pink to Prank My Husband." Twitter. January 09, 2019. https://twitter.com/everywhereist/status/1082813365628526592.

27. Devitt, Patrick. "13 Reasons Why and Suicide Contagion." Scientific American. May 08, 2017. https://www.scientificamerican.com/article/13-reasons-why-and-suicide-contagion1/.

28. "Digital Civility Index and Promoting a Safer Internet." Microsoft. 2019. https://www.microsoft.com/en-us/digital-skills/digital-civility?activetab=dci_reports:primaryr6.

29. Doyle, Sady. "Why Facebook's Harassment Guidelines Fail to Protect Women." ELLE. October 08, 2017. https://www.elle.com/culture/career-politics/news/a45567/facebooks-harassment-guidelines-leaked/.

30. Duggan, Maeve. "Online Harassment 2017." Pew Research Center: Internet, Science & Tech. January 03, 2018. http://www.pewinternet.org/2017/07/11/online-harassment-2017/.

31. Electronic Privacy Information Center. "EPIC - State Revenge Porn Policy." Electronic Privacy Information Center. Accessed March 2019. https://epic.org/state-policy/revenge-porn/.

32. "Enrolled House Bill No. 5017." Michigan Legislature - Home. January 02, 2019. http://www.legislature.mi.gov/documents/2017-2018/publicact/htm/2018-PA-0457.htm.

33. "Facebook Reports Fourth Quarter and Full Year 2018 Results." Facebook. Accessed March 2019. https://investor.fb.com/investor-news/press-release-details/2019/Facebook-Reports-Fourth-Quarter-and-Full-Year-2018-Results/default.aspx.

34. Ferran, Lee, Trish Turner, and Katherine Faulders. "Russia Targeted African-American Vote, Made Instagram 'key Battleground' in Propaganda War: Researchers." ABC News. December 17, 2018. https://abcnews.go.com/Politics/russia-targeted-african-american-vote-made-instagram-key/story?id=59862038.

35. Fiesler, Casey. "The Internet Rules Lab." Casey Fiesler. November 13, 2018. https://caseyfiesler.com/research-group/.

36. "Fight Back against Harassment." SQUADBOX. Accessed March 2019. https://squadbox.org/.

37. "From Fake News to Enemy of the People: An Anatomy of Trump's Tweets." Committee to Protect Journalists. January 30, 2019. Accessed April 03, 2019. https://cpj.org/x/7638.

38. Fruhlinger, Joshua. "Facebook Now Hiring Positions for 'safety,' 'security,' and 'trust'." Thinknum Media. March 22, 2018. https://media.thinknum.com/articles/facebook-is-hiring-around-a-tumultuous-future-for-privacy/.

39. Geers, Jacob. "The 49 Most Hilariously Savage Burns In Social Media History." Thought Catalog. June 22, 2016. https://thoughtcatalog.com/jacob-geers/2016/06/the-49-most-hilariously-savage-burns-in-social-media-history/.

40. Gleicher, Nathaniel, and Oscar Rodriguez. "Removing Additional Inauthentic Activity from Facebook." Facebook Newsroom. October 11, 2018. https://newsroom.fb.com/news/2018/10/removing-inauthentic-activity/.

41. Golder, Andy. "17 People Who Got Totally Fuckin' Burned." BuzzFeed. May 05, 2017. https://www.buzzfeed.com/andyneuenschwander/17-people-who-got-totally-fuckin-burned.

42. Green, Yasmin. "How Technology Can Fight Extremism and Online Harassment." TED. April 2018. https://www.ted.com/talks/yasmin_green_how_technology_can_fight_extremism_and_online_harassment.

43. Greenberg, Andy. "Security Isn't Enough. Silicon Valley Needs 'Abusability' Testing." Wired. January 28, 2019. https://www.wired.com/story/abusability-testing-ashkan-soltani/.

44. Griffiths, Sarah. "Future - Can This Technology Put an End to Bullying?" BBC. February 11, 2019. http://www.bbc.com/future/story/20190207-how-artificial-intelligence-can-help-stop-bullying.

45. Grynbaum, Michael M. "Trump Discusses Claims of 'Fake News,' and Their Impact, With New York Times Publisher." The New York Times. February 01, 2019. https://www.nytimes.com/2019/02/01/business/media/donald-trump-interview-news-media.html.

46. Hall, Charlie. "Anti-Defamation League Study Uncovers a Surge in Online Harassment." Polygon. February 13, 2019. https://www.polygon.com/2019/2/13/18223558/adl-anti-defamation-league-study-twitch-discord-harassment.

47. Hegeman, Roxana. "20 Years for Man behind Hoax Call That Led to Fatal Shooting." AP NEWS. March 29, 2019. https://apnews.com/9b07058db9244cfa9f48208eed12c993.

48. Jhaver, Shagun, Sucheta Ghoshal, Amy Bruckman, and Eric Gilbert. "Online Harassment and Content Moderation." ACM Transactions on Computer-Human Interaction25, no. 2 (2018): 1-33. doi:10.1145/3185593.

49. Marwick, Alice, and Rebecca Lewis. Media Manipulation and Disinformation Online. Report. Data & Society Research Institute. 2017. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf.

50. Marwick, Alice E. and Ross W. Miller. Online Harassment, Defamation, and Hateful Speech: A Primer of the Legal Landscape. Fordham Center on Law and Information Policy Report. June 10, 2014. https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=1002&context=clip.

51. Microsoft Digital Civility Index Resource Guide. 2019. https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RWrEDR.

52. "Human Rights Experts Denounce Trump's Attacks against Media | UN News." United Nations. August 02, 2018. https://news.un.org/en/story/2018/08/1016222.

53. "Impact Challenge – Google AI." Google AI. Accessed March 2019. https://ai.google/social-good/impact-challenge/.

54. Ingram, David. "A Grindr Harassment Suit Could Change the Legal Landscape for Tech - and Free Speech." NBCNews.com. January 05, 2019. https://www.nbcnews.com/tech/tech-news/grindr-harassment-suit-could-change-legal-landscape-tech-free-speech-n954976.

55. Jacobs, Julia. "Wikipedia Isn't Officially a Social Network. But the Harassment Can Get Ugly." The New York Times. April 08, 2019. https://www.nytimes.com/2019/04/08/us/wikipedia-harassment-wikimedia-foundation.html.

56. Jacobson, Roni. "I've Had a Cyberstalker Since I Was 12 | Backchannel." Wired. October 11, 2017. https://www.wired.com/2016/02/ive-had-a-cyberstalker-since-i-was-12/.

57. Jahromi, Neima. "The New Zealand Shooting and the Challenges of Governing Live-Streamed Video." The New Yorker. March 18, 2019. https://www.newyorker.com/tech/annals-of-technology/the-new-zealand-shooting-and-the-challenges-of-governing-live-streamed-video.

58. "Jamia Wilson." Jamia Wilson. Accessed March 2019. http://www.jamiawilson.com/.

59. Jeong, Sarah. The Internet of Garbage. Vox Media. 2018. https://cdn.vox-cdn.com/uploads/chorus_asset/file/12599893/The_Internet_of_Garbage.0.pdf.

60. Johnston, Casey. "Chat Logs Show How 4chan Users Created #GamerGate Controversy." Ars Technica. September 09, 2014. https://arstechnica.com/gaming/2014/09/new-chat-logs-show-how-4chan-users-pushed-gamergate-into-the-national-spotlight/.

61. Koerner, Brendan. "It Started as an Online Gaming Prank. Then It Turned Deadly." Wired. November 19, 2018. https://www.wired.com/story/swatting-deadly-online-gaming-prank/.

62. Kumar, Srijan, William L. Hamilton, Jure Leskovec, and Dan Jurafsky. Community Interaction and Conflict on the Web. Working paper. Stanford University. 2018. https://cs.stanford.edu/people/jure/pubs/conflict-www18.pdf.

63. Lartey, Jamiles. "Race and Russian Interference: Senate Reports Detail Age-old Tactic." The Guardian. December 24, 2018. https://www.theguardian.com/world/2018/dec/24/race-russian-election-interference-senate-reports.

64. Lowensohn, Josh. "Twitter Turns Its Block Function into a Mute Button, Leading to User Revolt." The Verge. December 13, 2013. https://www.theverge.com/2013/12/12/5205462/twitter-turns-its-block-function-into-a-mute-button#comments.

65. "Mark Zuckerberg." Mark Zuckerberg - Over the Last Few Weeks, We've Seen… May 03, 2017. https://www.facebook.com/zuck/posts/10103695315624661.

66. Matias, J. Nathan. "High Impact Questions And Opportunities for Online Harassment Research and Action." MIT Center for Civic Media | Creating Technology for Social Change. September 07, 2016. https://civic.mit.edu/2016/09/07/high-impact-questions-and-opportunities-for-online-harassment-research-and-action/.

67. McNamee, Roger. "I Mentored Mark Zuckerberg. But I Can't Stay Silent." Time. January 17, 2019. http://time.com/5505441/mark-zuckerberg-mentor-facebook-downfall/.

68. Meeker, Mary. "Internet Trends Report 2018." Kleiner Perkins. May 30. 2018. https://www.kleinerperkins.com/perspectives/internet-trends-report-2018/.

69. Mitchel, Shireen. "Hacking of 2016 Would Have Never Happened Had Folk #ListenedToBW." Twitter. April 22, 2018. https://twitter.com/i/moments/988069560921284608.

70. "Netizens." Netizens. Accessed March 24, 2019. https://www.netizensfilm.com/.

71. "Online Hate and Harassment: The American Experience." Anti-Defamation League. Accessed March 2019. https://www.adl.org/onlineharassment.

72. OnlineSOS. "Solutions and Resources for Online Harassment." OnlineSOS. Accessed March 2019. https://onlinesos.org/action-center.

73. Perez, Sarah. "Google Revamps Its Security Checkup Feature with Personalized Suggestions for Your Account." TechCrunch. October 16, 2017. https://techcrunch.com/2017/10/16/google-revamps-its-security-checkup-feature-with-personalized-suggestions-for-your-account/.

74. Petrocelli, Brian. "Moderators Are the Sword. Now AutoMod Is the Shield." Twitch Blog. December 12, 2016. https://blog.twitch.tv/moderators-are-the-sword-now-automod-is-the-shield-df3d8aae32a9.

75. Poland, Bailey. "Why Naming Online Harassment Accurately Matters." Women's Media Center. April 10, 2016. http://www.womensmediacenter.com/speech-project/why-naming-online-harassment-accurately-matters.

76. Poland, Bailey. Haters: Harassment, Abuse, and Violence Online. Lincoln: Potomac Books, 2016.

77. Reporters Without Borders. "Worldwide Round-Up of Journalists Killed, Detained, Held Hostage, or Missing in 2018." Accessed March 2019. https://rsf.org/sites/default/files/worldwilde_round-up.pdf.

78. "Research: Online Harassment Resource Guide." Wikimedia. December 08, 2018. https://meta.wikimedia.org/wiki/Research:Online_harassment_resource_guide.

79. Reuters. "Dating App Grindr Defeats Appeal over Harassment Campaign." NBCNews.com. March 27, 2019. https://www.nbcnews.com/tech/tech-news/grindr-defeats-appeal-over-harassment-dating-app-n988156.

80. Reventlow, Andreas. "The Chilling Effects of Online Harassment and How to Respond." International Media Support. December 06, 2016. https://www.mediasupport.org/chilling-effects-online-harassment-address/.

81. Risch, Julian, and Ralf Krestel. Aggression Identification Using Deep Learning and Data Augmentation. Report. Hasso Plattner Institute, University of Potsdam. 2018. https://aclweb.org/anthology/W18-4418.

82. Roberts, Sarah T. "Chapter Eight: Commercial Content Moderation: Digital Laborers' Dirty Work." The Intersectional Internet. doi:10.3726/978-1-4539-1717-6/19.

83. Romero, Eliza. "Online Harassment Is No Joke." Medium. July 15, 2018. https://medium.com/@AestheticDistance/online-harassment-is-no-joke-aca4ba34edf6.

84. "RSF Publishes Report on Online Harassment of Journalists | Reporters without Borders." RSF. August 01, 2018. https://rsf.org/en/news/rsf-publishes-report-online-harassment-journalists.

85. Salim, Saima. "How Much Time Do You Spend on Social Media? Research Says 142 Minutes per Day." Digital Information World. January 04, 2019. https://www.digitalinformationworld.com/2019/01/how-much-time-do-people-spend-social-media-infographic.html.

86. Schaub, Michael. "Dictionary.com Adds New Words, including 'cybermob,' 'Latinx' and 'dad Joke'." Los Angeles Times. April 03, 2019. https://www.latimes.com/books/la-et-jc-dictionary-new-words-latinx-dad-joke-20190403-story.html.

87. Scheffler, Rebecca and Gracey Zhang. "Someone Posted My Phone Number On Craigslist and Said I Wanted Strange Men to Rape Me." Narratively. August 01, 2018. https://narratively.com/someone-posted-my-phone-number-on-craigslist-and-said-i-wanted-strange-men-to-rape-me/.

88. Schildkraut, Jaclyn. "The Media Should Stop Making School Shooters Famous." Vox. March 31, 2018. https://www.vox.com/the-big-idea/2018/2/22/17041382/school-shooting-media-coverage-perpetrator-parkland.

89. Seabrook, John. "My First Flame." The New Yorker. January 24, 2019. https://www.newyorker.com/magazine/1994/06/06/my-first-flame.

90. "Shireen Mitchell." Shireen Mitchell's Biography | Founder, Speaker, Social Analyst, & Diversity Strategist. http://digitalsista.me/.

91. Short, Emma, Andrew Guppy, Jacqui A. Hart, and James Barnes. "The Impact of Cyberstalking." Studies in Media and Communication3, no. 2 (December 2015). doi:10.11114/smc.v3i2.970.

92. Siacon, Aleanna. "Michigan's New Cyberbullying Law about to Take Effect: What to Know." Detroit Free Press. March 25, 2019. https://www.freep.com/story/news/local/michigan/2019/03/25/michigan-cyberbullying-bullying-illegal/3266178002/.

93. Sinders, Caroline. "An Incomplete (but Growing) History of Harassment Campaigns since 2003." Medium. November 26, 2018. https://medium.com/digitalhks/an-incomplete-but-growing-history-of-harassment-campaigns-since-2003-db0649522fa8.

94. Smith, Aaron and Monica Anderson. "Social Media Use 2018: Demographics and Statistics." Pew Research Center: Internet, Science & Tech. September 19, 2018. https://www.pewinternet.org/2018/03/01/social-media-use-in-2018/.

95. Smith, Jada F. "As 'Sextortion' Proliferates, Victims Find Precarious Place in Legal System." The New York Times. December 21, 2017. https://www.nytimes.com/2016/05/11/us/sextortion-victims-brookings-institution.html.

96. Srikishan, Gautam. "The History of Online Harassment before and after Gamergate with Caroline Sinders." The Verge. December 04, 2018. https://www.theverge.com/2018/12/4/18125699/vergecast-podcast-interview-caroline-sindersonline-harassment-catalogue-gamergate.

97. "Standing Against Hate." Facebook Newsroom. March 27, 2019. https://newsroom.fb.com/news/2019/03/standing-against-hate/.

98. Stuart, S. C. "This AI Predicts Online Trolling Before It Happens." PCMag UK. March 19, 2019. https://uk.pcmag.com/news-analysis/120114/this-ai-predicts-online-trolling-before-it-happens.

99. Sugars, Stephanie. "From Fake News to Enemy of the People: An Anatomy of Trump's Tweets." Committee to Protect Journalists. January 30, 2019. https://cpj.org/x/7638.

100. Swisher, Kara. "Full Q&A: Former Google Lawyer and Deputy U.S. CTO Nicole Wong on Recode Decode." Recode (audio blog), September 12, 2018. https://www.recode.net/2018/9/12/17848384/nicole-wong-cto-lawyer-google-twitter-kara-swisher-decode-podcast-full-transcript.

101. Swisher, Kara. "Should the First Amendment apply to Facebook? It's complicated." Recode (audio blog), November 19, 2018. https://www.recode.net/2018/11/19/18103081/first-amendment-facebook-jameel-jaffer-freedom-speech-alex-jones-decode-podcast-kara-swisher

102. Tassi, Paul. "Overwatch's Fake Female Player 'Ellie' Scandal Is The Mess That Keeps On Giving." Forbes. January 06, 2019. https://www.forbes.com/sites/insertcoin/2019/01/06/overwatchs-fake-female-player-ellie-scandal-is-the-mess-that-keeps-on-giving/#5ea69dc628a0.

103. Tenold, Vegas. "To Doxx a Racist." The New Republic. July 26, 2018. https://newrepublic.com/article/150159/doxx-racist.

104. "The Cambridge Analytica Story, Explained." Wired. March 22, 2018. Accessed April 03, 2019. https://www.wired.com/amp-stories/cambridge-analytica-explainer/.

105. Thompson, Rachel. "She will not be silenced: The most harassed women online share

why they're not logging off." Mashable. June 18, 2018. https://mashable.com/2018/06/18/online-harassment-trolling-women/.

106. Tobin, Ariana, Madeleine Varner, and Julia Angwin. "Facebook's Uneven Enforcement of Hate Speech Rules Allows Vile Posts to Stay Up." ProPublica. March 09, 2019. https://www.propublica.org/article/facebook-enforcement-hate-speech-rules-mistakes.

107. Tufekci, Zeynep. "Yes, Big Platforms Could Change Their Business Models." Wired. December 17, 2018. https://www.wired.com/story/big-platforms-could-change-business-models/.

108. "Twitch Community Guidelines Updates." Twitch Blog. February 08, 2018. https://blog.twitch.tv/twitch-community-guidelines-updates-f2e82d87ae58.

109. "Twitter Health Metrics Proposal Submission." Twitter. March 01, 2018. https://blog.twitter.com/en_us/topics/company/2018/twitter-health-metrics-proposal-submission.html.

110. Vitak, Jessica, Kalyani Chadha, Linda Steiner, and Zahra Ashktorab. "Identifying Women's Experiences With and Strategies for Mitigating Negative Effects of Online Harassment." Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW 17, 2017. doi:10.1145/2998181.2998337.

111. "Warren Center Workshop: Viral Deception, Polarization and Networks." The Warren Center for Network and Data Sciences. Accessed March 2019. https://warrencenter.upenn.edu/events/viral-deception-polarization-and-networks-workshop/.

112. "What Does Trolling Really Mean? It's Not Trivial." OnlineSOS. Accessed March 2019. https://onlinesos.org/blog/what-trolling-means-online-harassment.

113. Witt, Stephen. "Tekashi 69: The Rise and Fall of a Hip-Hop Supervillain." Rolling Stone. February 01, 2019. https://www.rollingstone.com/music/music-features/tekashi-69-rise-and-fall-feature-777971/.

114. Wong, Julia Carrie. "Anti-vaxx 'mobs': Doctors Face Harassment Campaigns on Facebook." The Guardian. February 28, 2019. https://www.theguardian.com/technology/2019/feb/27/facebook-anti-vaxx-harassment-campaigns-doctors-fight-back.

115. "YouTube Hiring Director, Trusted Content, Trust and Safety in San Bruno, CA." LinkedIn. Accessed March 2019. https://www.linkedin.com/jobs/view/director-trusted-content-trust-and-safety-at-youtube-1029294667/.

Thank you to everyone who contributed their insights, knowledge, support, and feedback to produce this report. We couldn’t have done it without you.

OnlineSOS is a nonprofit organization that supports individuals facing online harassment through actionable resources and advocacy. Visit OnlineSOS.org for more infomation.