Defining “Online Abuse”: A Glossary of Terms

15-19 minutes ⋮ 4/11/2018

The first step to combatting online abuse is developing a shared language to identify and describe it.

The phenomenon has many names: cyber harassment, cyberbullying, trolling, flaming, etc. Some terms are used interchangeably, others have been drained of meaning. **PEN America prefers the terms *online harassment* or *online abuse*, which we define as the “pervasive or severe targeting of an individual or group online through harmful behavior.”**

- **Severe** because even a single incident of online abuse, such as a death threat or the publishing of a home address, can have serious consequences.
- **Pervasive** because, while some individual incidents of online abuse, such as insults or spam, may not rise to the level of abuse, a steady drumbeat of incidents, or a coordinated onslaught, does.
- **Online** includes email, social media platforms (such as Twitter, Facebook, Instagram, and TikTok), messaging apps (such as Facebook Messenger and WhatsApp), blogging platforms (such as Medium, Tumblr, and WordPress), and comments sections (on digital media, personal blogs, YouTube pages, and Amazon book reviews).

Below you’ll find definitions of many of the abusive tactics that writers and journalists face, as well as examples and tips on what to do. We hope you’ll explore the rest of this Field Manual for more detailed guidance.

Definitions

Astroturfing

Definition: Astroturfing is the dissemination or amplification of content (including abuse) that appears to arise organically at the grassroots level and spread, but is actually coordinated (often using multiple fake accounts) by an individual, interest group, political party, or organization.

Example: During the 2017 general election in Kenya, bloggers intimidate their journalists, judges and members of civil society using different hashtags, according to a research conducted by the non-profit Mozilla Foundation, inside the shadowy world of disinformation for hire in Kenya. The research goes ahead to reveal that the influencers are paid to directly harass politicians’ opponents.

What to do: Astroturfing is effective because harassers go to great lengths to make fake accounts seem real. Nevertheless, it’s a good idea to check for signs that an account might be fake (see the guidance on fake accounts below). You can try deploying a supportive community to help you report accounts, block and mute, and document the abuse. If you’re considering investigating the astroturfing campaign to expose and discredit it, take a look at these guidelines for practicing counterspeech.

Concern Trolling

Definition: Abusers pose as fans or supporters of a target’s work and make harmful and demeaning messages and comments masked as constructive feedback.

What to do: Because concern trolls are trying to get your attention and waste your time, counterspeech may be counterproductive and blocking could escalate the abuse. Muting—which enables you to hide specific abusive content (by user, keyword, etc.) so you don’t have to see it—might be more effective. Be

sure to report any content that crosses over from the annoying to the abusive and consider rallying a supportive cyber community to have your back.

Cross Platform Harassment

Definition: Cross-platform harassment is coordinated and deliberately deployed across multiple social media and communications platforms, taking advantage of the fact that most platforms only moderate content on their own sites [adapted from WMC].

What to do: There is no easy way to deal with coordinated cross-platform harassment. It's critically important to tighten your cybersecurity to protect yourself from hacking and doxing. To cope with the volume and reach of the attacks, it helps to rally a supportive cyber community to share the burden of documenting, reporting, blocking and muting the abuse.

Cyberbullying

Definition: An umbrella term, cyberbullying encompasses many harassing behaviors, but boils down to "willful and repeated harm inflicted through the use of computers, cell phones, and other electronic devices" [Source: Cyberbullying Research Center]. The term is primarily used in relation to children and young adults.

What to do: Visit Cyberbullying.org for the best resources and information related to cyberbullying.

Cyber-Mob Attacks (aka Dogpiling)

Definition: When a large group of abusers collectively attacks a target through a barrage of threats, slurs, insults, and other abusive tactics.

Outrage/Shame Mobs
A form of mob justice focused on publicly exposing, humiliating, and punishing a target, often for expressing opinions on politically charged topics or ideas the outrage mob disagrees with and/or has taken out of context in order to promote a particular agenda.

Example: Ricochet editor and politically-conservative columnist Bethany Mandel experienced a surge of anti-Semitic trolling from self-identified white nationalists via Facebook and Twitter after publicly declaring her opposition to Donald Trump.

What to do: Trying to navigate cyber-mob attacks can feel like an exhausting game of whack-a-mole. If reporting the abuse isn't getting you anywhere, consider asking a member of your support community to monitor and report the abuse on your behalf while you take a break. Other options include: launching a counterspeech campaign to reestablish a narrative or reclaim a hashtag associated with your username; making a statement on social media alerting your social network to the negative activity; and/or temporarily taking a break from or going private on your social media accounts until the worst of the harassment has passed.

Cyberstalking

Definition: In a legal context, "cyberstalking" is the prolonged and repeated use of abusive behaviors online (a "course of conduct") intended "to kill, injure, harass, intimidate, or place under surveillance with intent to kill, injure, harass, or intimidate" a target [See: 18 U.S. Code § 2261A].

Example: Over a 15-year period, a freelance journalist at *Scientific American* was the target of cyberstalking from a man who would go on to steal her identity and threaten her career. Read her story at *Wired*.

What to do: Cyberstalking is a federal offense, and many states have cyberstalking laws on the books. If you're comfortable contacting law enforcement or seeking the advice of a lawyer, you might wish to take legal action against a cyberstalker. Other strategies include blocking your stalker on social media, documenting every harassing incident that occurs in relation to cyberstalking, making sure your online accounts are protected if you anticipate identity fraud, and enlisting your support community.

Deepfake

Definition: The use of "a form of artificial intelligence called deep learning" to make manufactured images, audio, and/or video that appear real. [Source: Guardian] These images, audio, and/or video are "mimicking speech or facial expressions so as to make it appear that someone has said or done something they haven't." [Source: The Boston Globe].

Example: A deepfake porn video created using the image of investigative journalist Rana Ayyub was shared more than 40,000 times in an attempt to humiliate and silence her. She was brave enough to share her story.

What to do: If a deepfake image or video of you starts circulating online, report it—and the accounts posting or sharing it—wherever possible. Try a reverse image search using all or part of the photo to figure out where else it may have spread and continue reporting. To protect yourself from further violations of your privacy, take some time to research what information is available about you online and try to reduce it or limit its visibility. You may also consider speaking out to discredit the deepfake and enlisting a support community to help you with research, reporting, and counterspeech.

Denial of Access

Definition: Leveraging the "features of a technology or platform to harm the target, usually by preventing access to essential digital tools or platforms" [Source: Data & Society].

Mass Report (aka False Reporting)
Abusers coordinate to falsely report a target's account as abusive or otherwise harmful to try to get it suspended or shut down.

Message Bombing (aka Flooding)
Abusers flood an individual or institution's phone or email accounts with unwanted messages meant to limit or block the target's ability to use that platform.

Example: In 2017, a flood of emails sent by bot accounts shut down the servers at ProPublica in a retaliatory attack against ProPublica journalists who had written a controversial article about the relationship between tech companies and extremist websites. The attack prevented the company's employees from accessing important emails and interfered enormously with the news outlet's day-to-day operations.

What to do: Immediately report the incident to the social media platform, phone provider, internet company, or email provider where the harassment is taking place. If necessary, create a new and/or temporary email address or username to inform your colleagues, family, and friends that you have been message bombed and no longer have access to your usual accounts. You can find more in-depth information on dealing with online harassment via direct messages here and on talking to friends and family about online abuse here.

Denial of Service (DoS) Attacks

Definition: A cyberattack that temporarily or indefinitely causes a website or network to crash or become inoperable by overwhelming a system with data. DoS attacks can prevent you from accessing your own devices and data, and they can compromise sensitive information stored on your devices.

Distributed Denial of Service (DDoS)
When an attacker takes control of multiple users' computers in order to attack a different user's computer. This can force the hijacked computers to send large amounts of data to a particular website or send spam to targeted email addresses.

Example: In 2016, the BBC suffered a targeted DDoS attack in its U.S. offices, which also caused limited access to Reddit, Twitter, Etsy, GitHub, SoundCloud, and Spotify.

What to do: Because DoS attacks target email addresses, websites, and online accounts, it's essential that you contact the necessary providers to report the abuse. Check out this DDoS Incident Response Cheat Sheet for more information.

Dog Whistling

Definition: Using words or symbols with a double (or coded) meaning that is abusive or harmful, sometimes to signal a group of online abusers to attack a specific target [PEN America and IWMF, Totem Project].

Example: In 2016, white supremacists on Twitter began using triple parentheses—an (((echo)))—around an individual's name to identify them as Jewish and instigate a coordinated campaign of abuse. Jewish writers and journalists banded together to reclaim the symbol, proactively adding triple parentheses to their own Twitter handles.

What to do: For dog whistling to work as abusers intend, only the abusers can know of the double-meaning of the terms or symbols deployed. One potentially effective way to respond is through counterspeech. Consider reclaiming the symbol or word or rallying a supportive cyber community to expose the dog whistle and undermine its power. Here's some guidance for practicing counterspeech safely. If you're concerned that the dog whistle might unleash a cyber mob or threaten your safety, you can find more information about those tactics elsewhere on this page.

Doxing (aka Doxxing)

Definition: The publishing of sensitive personal information online—including home address, email, phone number, social security number, photos, etc.—to harass, intimidate, extort, stalk, or steal the identity of a target. Short for "dropping docs," doxing was a revenge tactic among '90s computer hackers, according to HTML.com.

Example: After reporting on the police officer involved in the shooting of Michael Brown in Ferguson, Missouri, two reporters for *The New York Times* were forced to flee their homes when their personal addresses were posted online in retaliation for their coverage.

What to do: Check out the Protecting Information from Doxing section of this Field Manual for tips on preparing for and preventing doxing. If you've already been subjected to doxing, immediately report the dox to the platform on which it appears, and do your best to assess the threat level to your safety. If you believe that the doxed information could fall into the hands of someone intent on harming you, please consider involving your local law enforcement immediately.

Hacking

Definition: The unauthorized intrusion into a device or network, hacking is often carried out with the intention to attack, harm, or incriminate another individual by stealing their data, violating their privacy, or infecting their devices with viruses. When hacking is used to perform illegal activities or intimidate a target, it is a cybercrime.

Example: Fancy Bear, a Russian hacking unit, has targeted hundreds of journalists, including

independent Russian reporters, at least 50 *New York Times* journalists, and several reporters at The Daily Beast, among other journalists who report on intelligence, national security, and Russian troll farms.

What to do: Practicing rigorous cyber security is critical to protecting yourself from hacking.

Hashtag Poisoning

Definition: The creation of an abusive hashtag—or the hijacking of an existing hashtag—which is then leveraged as a rallying cry for cyber mob attacks. [Source: Adapted from RSF]

Example: In 2015, Feminist hashtags #TakeBackTheTech and #ImagineAFeministInternet were overwhelmed by a coordinated flood of misogynistic messages and memes in an attempt to "destroy" the campaigns.

What to do: Rally a supportive cybercommunity to reclaim a hijacked hashtag or create a new supportive hashtag, as comedian and actress Leslie Jones' fans did with #justiceforleslie to counter the coordinated campaign of racist and misogynist abuse she was subjected to in 2016. In 2020, gay men across Twitter turned the tactic back on the purveyors of white supremacist hate by overwhelming #ProudBoys with positive images of themselves with their partners.

Hateful Speech

Definition: Expression that attacks a specific aspect of a person's identity, such as their race, ethnicity, gender identity, religion, sexual orientation, disability, etc. Hateful speech online often takes the form of ad hominem attacks, which invoke prejudicial feelings over intellectual arguments in order to avoid discussion of the topic at hand by attacking a person's character or attributes.

Example: In 2016, comedian Leslie Jones was the target of a widespread trolling campaign built on racist, misogynist messaging, which culminated in her being subjected to revenge porn, doxing, and hacking. (The incident also resulted in notorious troll Milo Yiannopoulos being kicked off Twitter for his targeted, publicly racist abuse—thanks in part to a #LoveforLeslieJ Twitter campaign launched by her supporters.)

What to do: Depending on the level of threat and intimidation couched in these attacks, you may wish to block or mute a user, engage in counterspeech, or, in some cases, even consider directly confronting your troll. If you don't feel safe responding to or blocking a user, turn to your support community and make sure you're practicing self-care. If you've been named in a threat of violence or sexual intimidation and are afraid for your safety, please consider contacting law enforcement.

Nonconsensual Intimate Images (aka Revenge Porn)

Definition: Nonconsensual pornography is "the distribution of private, sexually-explicit images [or videos] of individuals without their consent" [Source: Cyber Civil Rights Initiative].