

Tech Policy Design Lab: Online Gender-Based Violence and Abuse

Outcomes & Recommendations



3 Executive Summary

6 Introduction

8 Workshop Approach

11 Recommendations

13 Curation

21 Reporting

27 Next Steps

Executive Summary

The internet is one of the most powerful platforms for opportunity, sharing ideas and building community the world has ever seen, but women are too often silenced and censored because they fear the repercussions of speaking out. Online gender-based violence and abuse (OGBV) is a grave threat to progress on gender equality. When women have less space online, they have less space in newsrooms, boardrooms and the halls of political power.

The Web Foundation's Gender and Data Rights team has created a dedicated workstream focused on developing solutions to OGBV as part of the Web Foundation's commitment to [tackle digital inequality](#).

The Web Foundation's work in this area is firmly rooted in the [UN Women Technology and Innovation for Gender Equality Action Coalition Blueprint](#) which states: 'By 2026, a majority of countries and tech companies demonstrate accountability by implementing policies and solutions against online and tech-facilitated GBV and discrimination.' Specifically, companies are being asked to work towards preventing and eliminating online gender-based violence by 'developing better tools to prevent and respond to abuse', including 'improving tools that avoid harm and user care services.'

In the first stage of this work, we convened four multi-stakeholder consultations between March 2020 and February 2021. The consultations brought together some of some of the world's largest tech platforms (Facebook/Instagram, Twitter, Google/YouTube, and TikTok), civil society organisations, government representatives and academia to generate evidence, build relationships and engage in constructive dialogue around women's experiences of online gender-based violence and abuse.

We then used the insights and evidence gathered during our consultations to develop a series of three policy design workshops throughout April 2021. Led by two service design firms, [Feminist Internet](#) and [Craig Walker](#), the workshops applied design-thinking and co-creation to develop prototypes that centre the experiences of women most impacted by online abuse.

Specifically, the workshops focused on the experiences of highly visible women online (e.g. politicians, journalists, activists) from around the world, and provided an opportunity for participants to innovate around content curation and reporting flows.

Workshops consisted of 20-25 participants, including 2-3 representatives from each tech company from both product and policy teams, as well as 15-20 participants from civil society, policy and regulatory bodies, and academia.

Participants produced 11 prototypes during the workshops that can help create safer online experiences for women and tackle OGBV.

The prototypes also helped inform detailed recommendations and suggestions for technology companies around two main themes:

- 1 Greater control over who can comment or reply to posts as well as more choice over what women see online, when they see it and how they see it
- 2 Improved reporting systems so women can be better supported when they do receive violent or abusive content

In this report

We recognise that tech companies are already taking some steps to address OGBV on their platforms and our focus is to consider ways to further refine this work with illustrative examples. We emphasise the need to provide more opportunities for women to curate their online experiences and to improve the reporting process, while recognizing that this focus is not exhaustive. Areas for future work could include image-based abuse, content moderation, and better understanding the motivations and experiences of perpetrators.

Chatham House Rule

All workshops were convened under a modified Chatham House Rule and this report does not reveal implicitly or explicitly the identity of workshop participants apart from whether they are representatives of civil society organisations, government and regulatory bodies, academia or tech platforms

Recommendations

Curation

Build better ways for women to curate their safety online

This could be achieved by:

- 1 Offering more granular settings (e.g. who can see, share, comment or reply to posts)
- 2 Using more simple and accessible language throughout the user experience
- 3 Providing easy navigation and access to safety tools
- 4 Reducing the burden on women by proactively reducing the amount of abuse they see

Reporting

Implement improvements to reporting systems

This could be achieved by:

- 1 Offering users the ability to track and manage their reports
- 2 Enabling greater capacity to address context and/or language
- 3 Providing more policy and product guidance when reporting abuse
- 4 Establish additional ways for women to access help and support during the reporting process

“The world has made important progress on gender equality thanks to the unceasing drive of committed champions everywhere, but I am seriously concerned that online harms facing women and girls — especially those of colour, from LGBTQ+ communities and other marginalised organisations — threaten that progress.”

Sir Tim Berners-Lee

Web Inventor and Co-Founder, The World Wide Web Foundation

Introduction

Women and girls are experiencing a surge of violence and abuse online and the digital world has become a new front in the assault on the freedom and power of women. Being a woman online means being disproportionately at risk of abuse. When you're a woman of colour, and a Black woman in particular, LGBTQ+ or from other marginalised organisations — the abuse is often far worse.

Globally 38% of women have reported being personally subject to online violence according to the latest numbers from the **Economist Intelligence Unit**. This rises to 45% for Gen Zs and Millennials. This is consistent with a **Web Foundation survey** which found 52% of young women and girls had experienced abuse online.

This abuse isn't less harmful because it happens behind a screen. Impersonation, defamation, threats of physical and sexual violence, the non-consensual sharing of images – all of these can have devastating consequences for the reputations and the physical and mental health of those targeted. It can cost women their jobs and damage relationships.



As part of our work to tackle this crisis of online gender-based violence and abuse the Web Foundation convened a series of four multi-stakeholder consultations between March 2020 and February 2021 focusing on [women activists](#), [women in public life](#) and [young women](#).

The consultations brought together some of the world's largest tech platforms (Facebook/Instagram, Twitter, Google/YouTube, and TikTok), civil society organisations, researchers and other experts to generate evidence, build relationships and engage in constructive dialogue around women's experiences of OGBV on social media platforms.

Here are the key takeaways:

- 1 Being a woman in public life and visible online makes you a target for different forms of abuse.
- 2 Online gender-based violence and abuse can have offline impacts.
- 3 Online gender-based violence and abuse disproportionately impacts women with overlapping and intersecting identities.
- 4 Not all journalists and politicians experience abuse and violence in the same way.
- 5 While many companies have developed innovative privacy and security tools, it can be difficult for women in public life to find and use these features.
- 6 Reporting systems need to be improved to better support women.
- 7 Researchers need better transparency and access to data.
- 8 Online gender-based violence and abuse against women journalists and politicians has a generational impact.

These are significant, global challenges but we firmly believe that we can build solutions where there is meaningful multi-stakeholder participation; where people and their online experiences drive policy and product design; and where solutions take into account the full diversity of those who use digital tools.

As a result of the consultations, we recognised the need to create a space where civil society organisations, government representatives, academia and tech platforms could work together constructively to create product and policy solutions that can shape a better web for women. We decided to use the important topic of OGBV targeted at highly visible women as a pilot for the Web Foundation's [Tech Policy Design Lab](#), which brings together governments, companies and civil society in a collaborative environment and uses design thinking to develop policy frameworks, products and services that address the most challenging tech policy issues of our time.

We used the insights and evidence gathered during our consultations to develop a series of three workshops. In collaboration with two service design firms, Feminist Internet and Craig Walker, we applied design-thinking and co-creation techniques to develop user experience prototypes that centre and respond to the experiences of women disproportionately impacted by online abuse - politicians, journalists and activists from around the world.

Civil society organisations, government representatives, academia and some of the world's largest tech platforms were invited to co-create policy and product solutions to mitigate OGBV. We asked participants to innovate around content curation and reporting flows. As a result, 11 prototypes were developed to address key problems that underpin women's experiences of abuse and harassment online.

Our consultations and workshop series brought together over 120 experts from over 35 countries. The workshops consisted of 20-25 participants, including 2-3 representatives from each tech platform from both product and policy teams, as well as 15-20 participants from civil society organisations, government and regulatory bodies, and academia.

The following section outlines the outcomes and analysis of our multi-stakeholder workshops, using the prototypes developed as illustrative examples for possible ways to allow for greater curation and improved reporting systems. We spotlight the ideas that were co-created by participants alongside an analysis of the outcomes and recommendations.

“

How will platforms contribute to normalising the idea that any form of online abuse is unacceptable and should not be an accepted norm?

- Workshop Participant, CSO

Workshop Approach



The workshops applied design thinking and co-creation to develop concrete solutions that centre the experiences of women most impacted by online abuse.

We challenged product and policy experts from some of the world's largest tech platforms, civil society organisations, government and regulatory bodies and academia to collaborate and create user experience prototypes for tackling OGBV. Specifically, these workshops provided an opportunity to innovate around content curation and reporting flows.

Participants worked in mixed small groups to collaboratively design solutions for a specific scenario made up of a persona, a 'human problem' and a fictional app.

The scenarios were based on the lived experiences of women who face abuse online, with special consideration given to overlapping and intersecting identities (i.e., race, ethnicity, sexual orientation, gender identity, etc.).

We know that if we can solve problems facing these women, then we create a safer online experience for everyone.

Personas

The five personas focus on highly visible women with intersecting identities. They were developed based on evidence from our consultations, the project team's expertise and desk research.

[View the personas in full here](#)



Amy

Body positivity influencer, US

Amy experiences a broad range of hurtful and abusive messages in her comments and replies, from diet advice and 'concerns' about her health, to multiple abusers saying 'go kill yourself'. The abuse targets all aspects of her identity, and is increasingly racist and xenophobic, particularly in the last year, which has seen a surge in anti-Asian racism online.



Paula

University student, Brazil

After a bad personal experience on a date, Paula created a short video about consent. She posted it to raise awareness, and expected a few of her friends to see it, but it went viral, with 100,000 views within 24 hours, and 1 million views within 3 days.



Yvonne

Politician, UK

Yvonne experiences large volumes of abuse across platforms. The abuse often focuses on personal characteristics including Yvonne's gender, race and physical appearance, as well as a perceived lack of competence.



Karishma

Journalist, India

Karishma receives abusive replies when she posts on social media platforms. She is concerned about continuing to speak about certain topics, especially as she has witnessed more established journalists experiencing threats of violence and knows a colleague who had their home address published on a social media platform.



Mouna

Activist, Tunisia

Mouna's experience of online abuse has been infuriating and overwhelming. YouTube videos of them attending protests have been repeatedly posted on the platform without them knowing. This results in abusive comments that often refer to their gender identity, and/or attempt to discredit them and the NGO they work for.

“

The persona helps you ground your thinking in a human experience, and consider how someone's identity and location may make it specific.

- Workshop Facilitator

Human Problems

Drawn from lived experiences of abuse against women online, the human problems gave participants a specific scenario to focus on.

“

The human problem is the central focus - this is what you are trying to address in your design.

- Workshop Facilitator

“

I've tried to report loads of abusive comments, but I haven't heard back from the platform in a while and I don't know how to check if I did it right.

I am under siege by a growing online mob. I need a way to protect myself and stop it, but I don't know how.

I've been attacked but I don't have any evidence because the abusive comments have been removed by the perpetrator or the platform.

I feel frustrated and helpless because I don't know if reporting the abuse will make any difference. I don't even know if anyone will think it's abuse!

People are sending me hurtful and abusive messages. It's overwhelming and I need a way to protect myself.

Fictional Apps

Fictional platforms were also created for the participants to build their solutions for. These apps gave participants the freedom to design possible solutions and the space to experiment - participants had the opportunity to think about and design solutions for what is necessary rather than just what is currently possible.

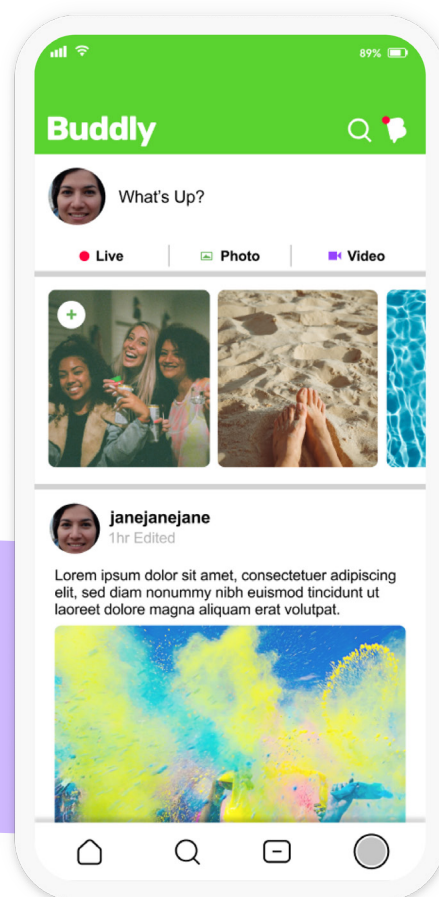
The fictional apps also encouraged collaboration, and avoided any one person from being the authority on a particular app:

1. *Buddly*: a social media platform.
2. *Picus*: an image-sharing app.
3. *TVez*: a video-sharing app.
4. *Zowie*: a short video-sharing app.
5. *Mycro*: a microblogging app.

“

The fictional app helps you locate your design in a specific online space.

- Workshop Facilitator



Recommendations

This report outlines our analysis of the workshops' outcomes, spotlights innovative features from the prototypes and makes recommendations for tech companies based on the two key themes – Curation and Reporting.

We recognise that many tech companies are already taking some steps to address OGBV on their platforms and we have focused on ways to further refine that work through illustrative examples in the form of product prototypes.

In each thematic area that follows, you will see:

- 1 **Recommendations**, which include one key recommendation for each theme.
- 2 **Design suggestions**, which detail how to achieve each recommendation.
- 3 **Illustrative examples**, which provide product examples from the workshops that provide inspiration for how to achieve the design suggestions.
- 4 **Considerations**, which product and policy teams may need to address when developing solutions. These take into account both technical challenges, the possibilities for misuse of product features by bad actors, and policy changes or updates that may be necessary.

Curation

This theme is about how to give women more choice over what they see online, when they see it and how they see it, as well as greater control over who can comment or reply to their posts.

Reporting

This theme explores how reporting processes can be improved so women can better manage and track their reports, and so that companies can respond to reports of abuse more effectively.

Curation

Recommendation

Build better ways for women to curate their safety online

The consultation evidence showed that existing user controls do not give women enough options to avoid and manage abuse in replies and comments. Women expressed a need for greater control over who can comment or reply to their posts, as well as more choice over what they see online, when they see it and how they see it.

Key stakeholders noted that this challenge presents opportunities for users to reduce the burden of managing abuse they receive. It is also a chance to think in new ways about how communities of users might more actively manage abuse together.

Design Suggestions

- 1 Offering more granular settings (e.g. who can see, share, comment or reply to posts)

Design features that give women greater control over the types of comments they see online and more options over who can respond to or share their posts.

- 2 Using more simple and accessible language throughout the user experience

Ensure that any steps taken to address or help users who are experiencing abuse are clearly communicated and easily accessible.

- 3 Providing easy navigation and access to safety tools

Ensure that safety tools already available are as easy to find and use as possible.

- 4 Reducing the burden on women by proactively reducing the amount of abuse they see

Help women manage online abuse but also ensure they are not overwhelmed by the feeling that dealing with abuse is solely up to them.

Design Suggestion 1

Offering more granular settings (e.g. who can see, share, comment or reply to posts)

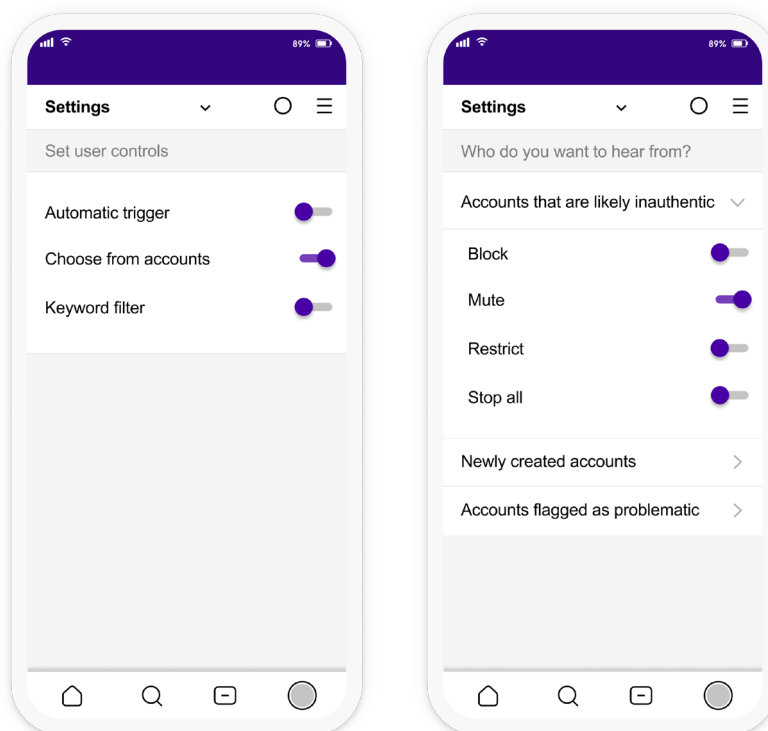
To achieve this you could:

- a Allow users to set times for disabling/enabling specific controls such as turning off comments and replies.

This could happen in advance - i.e a journalist anticipates an article she is posting will prompt a spike in abuse. It could also happen in the moment - i.e. someone is experiencing a spike in comments and wants a break from receiving replies.
- b Allow users to set up automated controls, so platforms can automatically perform tasks such as blocking accounts with previous violations or automatically muting accounts from certain types of users.

Prototype Example

Calm the Crowd offers users more granular controls when a spike in unusual activity is detected. [View this prototype in full here ↗](#)



Calm the Crowd

Offering more granular settings (e.g. who can see, share, comment or reply to posts)

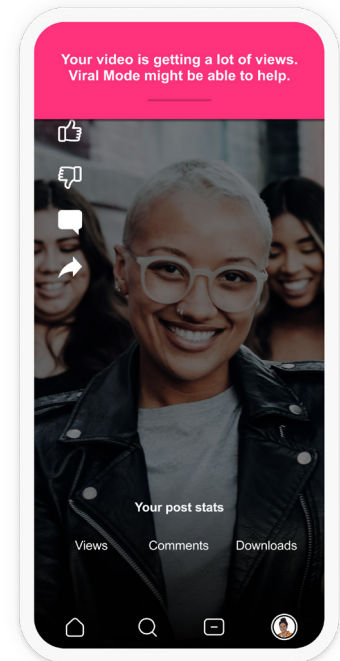
To achieve this you could:

- c Provide users with data that can help them understand when their account is receiving more attention than usual so they can make informed decisions about how to respond.

Such metrics could include an analysis of view count, shares, comments, and download count of any content they have posted. Users could receive notifications when their content is getting more attention than usual, or when their account is experiencing unusual patterns of activity. These analytics could be collated into a summary report.

Prototype Example

Viral Notification automatically offers a Viral Mode where users can turn off comments or downloads when their content is shared or viewed at scale. [View this prototype in full here ↗](#)



Viral Notification

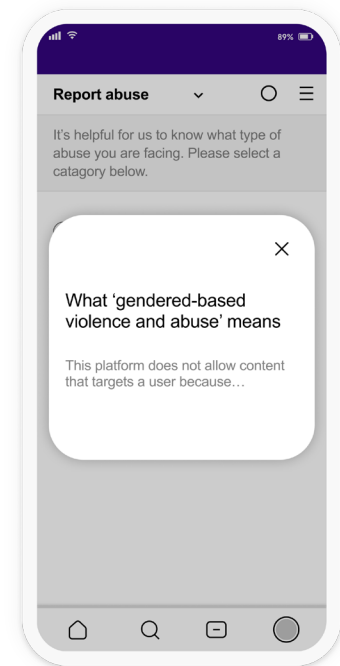
Considerations for this suggestion

- Spikes in use may look different for different users. For some, even 10-20 comments could be overwhelming, whilst for others much higher volumes of comments may be considered fairly normal. It may be challenging to adjust parameters on a case by case basis.
- Platforms may have already implemented similar features in the backend. For example, new accounts or suspected bad actors may not be allowed to comment or reply.

Using more simple and accessible language throughout the user experience

To achieve this you could:

- a Review language used in policies and product features with women from around the world to ensure it can be understood by a wide range of people who experience abuse in different cultural and geographical contexts.
- b Consider tone of voice throughout the user journey to make it feel human rather than technical or impersonal.
For example, 'This does not violate our community standards' could be rephrased using more considerate/understanding language, such as, 'From what we could observe, this doesn't seem to violate our policies. Would you like to provide more info?'. Or, when a user reports content they are able to hover over different categories of abuse and see a short definition or explanation of what each category entails.
- c Provide clear, real-time explanations of safety features within the user experience and explain definitions of harassment clearly during the reporting process.
For example, a hover function provides text that explains the function of a safety mode feature.



Reporting 2.0

Prototype Example

Reporteroo provides a transparent reporting process that allows easy, real-time access to relevant information, considers content in local languages and offers the ability to add context to a report.

[View this prototype in full here ↗](#)

Prototype Example

Reporting 2.0 enables users to hover over content to get more information on a specific policy or community guideline.

[View this prototype in full here ↗](#)

Considerations for this suggestion

- Tone matters - think about the audience, including how the age or region of a user, for example, might impact their interpretation or understanding. Think about how the names of product features might feel for different audiences.
- Ensure policies and community guidelines feel 'human' and refrain from language that is dense, technical or impersonal, particularly as this may also not translate well in other languages or contexts.

Providing easy navigation and access to safety tools

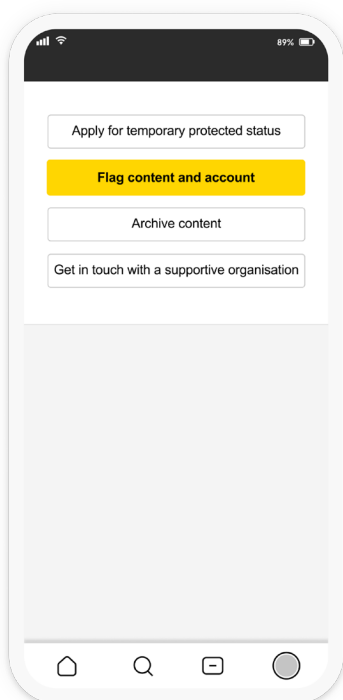
To achieve this you could:

- a Build an option for users to activate a suite of safety controls that can be activated in one click (either at any time or specifically during an emergency).

These features could include disabling comments, activating a profanity or keyword filter or disabling tags. The option can be accessed from Settings, or by simply clicking a button available on different pages throughout the platform - i.e. feeds page, posts page or profile page.

Prototype Example

Gateway provides a button (switched on in Settings) that enables users to signal to the platform that they feel under attack and are receiving harmful and abusive comments/replies. The platform actively shields the user from potentially abusive comments and content. Users are also able to archive any flagged content and are prompted to get in touch with civil society organisations that can provide further support and guidance. [View this prototype in full here ↗](#)

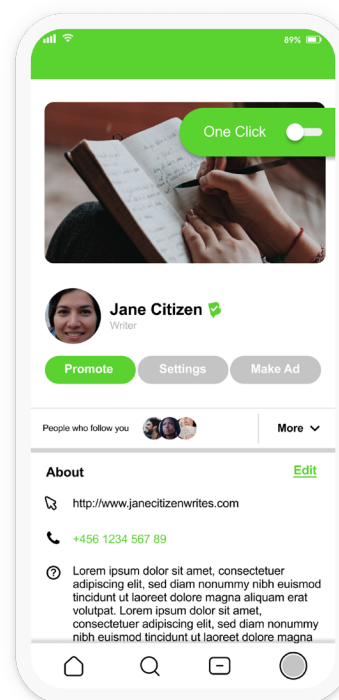


Gateway

Prototype Example

One Click provides a quick toggle that enables predefined Settings to protect the user. The toggle can be accessed from both the settings and the profile page.

[View this prototype in full here ↗](#)



One Click

Providing easy navigation and access to safety tools

To achieve this you could:

- b** Prompt users to activate safety tools if they experience a spike in comments or replies.

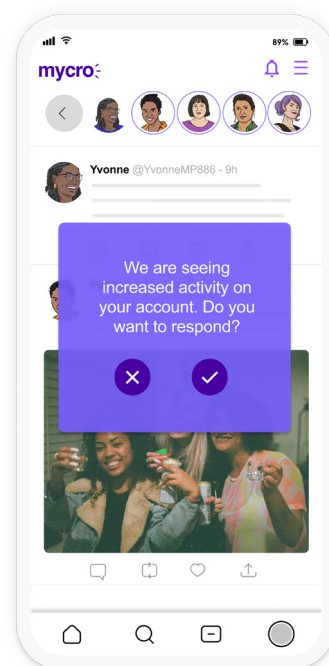
For example if a platform detects unusual activity or unusual engagement patterns on an account, they could provide a notification to the user so that they can review the activity and decide how to respond.

- c** Give users a tour of controls when they sign up and during instances of online abuse.

For example, when a user makes a report, offer them the choice to watch a pop up video explaining the reporting process before they begin.

Prototype Example

Calm the Crowd offers users more granular controls when a spike in unusual activity is detected. [View this prototype in full here >](#)



Calm the Crowd

Considerations for this suggestion

- Be strategic about when you prompt users to check their safety and when you make buttons available - if they're there all the time users may not even notice them.
- It is very challenging to define abuse in a globalised way. More work needs to be done to work with regional experts and youth experts, for example, to understand what language is appropriate and accessible.

“

It's critical for the platform to be able to distinguish between a 'pile on' of attacks vs. something going viral that's positive.

- Workshop Participant,
Government / Regulatory Representative

Reducing the burden on women by proactively reducing the amount of abuse they see

To achieve this you could:

a Continue to improve support to users through automated tools.

For example, data analytics could be used to identify accounts that are likely inauthentic, newly created, that participate in coordinated attacks or that users may have already flagged as problematic. AI could be used to identify words that seem abusive for the user and automatically notify them if any of the terms appear in their comments.

Prototype Example

Image Shield provides a personalised AI system that learns new words, phrases, and slang related to a particular protected characteristic faster than existing/more general models.

[View this prototype in full here ↗](#)

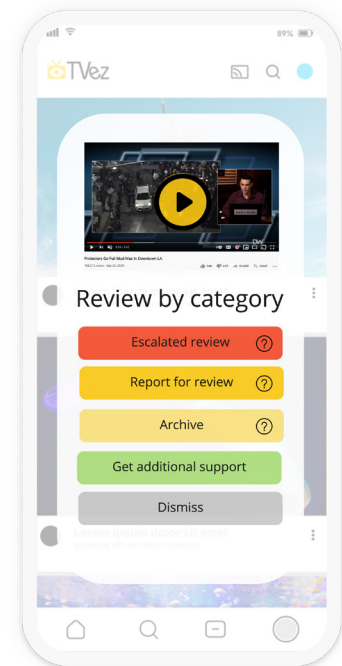


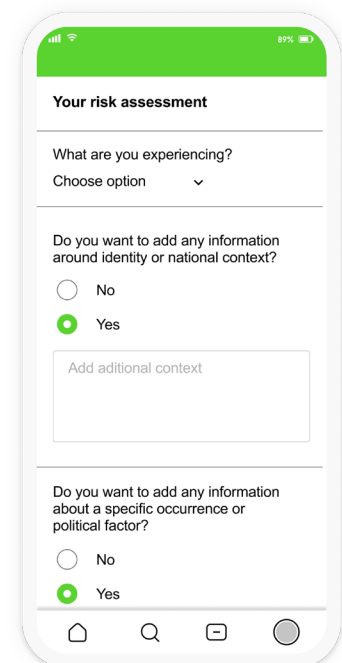
Image Shield

b Improve keyword and/or profanity filters for the replies and comments women see, including user-curated filters.

For example, along with words that companies may already flag as hate speech, users may want to add words to block or mute that are relevant to their national or local context or words in other languages. Users are also allowed to make time-specific adjustments based on events going on in their area (e.g. organised protests/marches) or attacks they expect to happen (e.g. after posting specific content).

Prototype Example

Wellbeing Check-Up prompts users to enter keywords that the platform can scan for and subsequently hide, block or mute. Users are specifically prompted to enter context-specific keywords based on identity, national events, languages they may be exposed to or other relevant events. [View this prototype in full here ↗](#)



Wellbeing Check-up

Reducing the burden on women by proactively reducing the amount of abuse they see

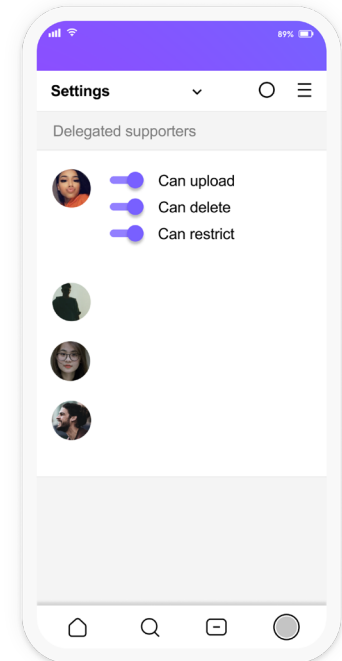
To achieve this you could:

- c Improve ways for users to share the burden of managing abuse they receive with trusted contacts.

For example, users could assign different roles to trusted contacts, giving them authority to upload or delete content, or restrict and delete comments. Posts uploaded by trusted contacts could have the option to be marked with a verified 'trusted contact' badge. This would help reduce security risks for people who already ask trusted contacts to help them with managing abuse informally, for example by sharing passwords.

Prototype Example

Com Mod offers users tools to define trusted users, assign them different roles to moderate comments, and approve or reverse the action taken by the trusted users. The comments can also be deleted or restricted in bulk. [View this prototype in full here ↗](#)



Com Mod

“

I would also like an option for the company to be flagged when trusted supports are being used so they can document what types of attacks require additional community support

- Workshop Participant, CSO

Considerations for this suggestion

- People are already using informal ways for trusted contacts to help them manage their content, for example through sharing passwords or adding administrators to public pages. Sharing passwords is a security risk and platforms should provide users with alternative options to help them manage abuse and reduce this risk.
- Trusted contacts will also need to be protected, as they could also be targeted when helping someone manage abusive content.
- If trusted contacts can manage content, there needs to be careful consideration around bulk delete functions, as politicians or other public figures could also use them to mass delete dissenting/critical voices.
- It is important that companies do not shift too much responsibility onto people who volunteer to help others manage abuse.
- If creating features that use AI and automated controls to manage abuse, platforms should be transparent about how they work. The ways that accounts are analyzed, classified and detected should be clear to users.

Reporting

Recommendation

Implement improvements to reporting systems

[Evidence from our consultations](#) showed that reporting flows do not always adequately consider context and often do not offer the option to add nuanced information when needed. Women are not able to effectively manage and track the progress of their reports and feel frustrated when they can't check the status of a report, particularly when they do not hear back from platforms.

Key stakeholders noted that this challenge presents opportunities for companies to increase user trust by being transparent about reporting decisions and processes, and to build principles of procedural fairness into the process. On a broader level, improving reporting processes and related policies is an opportunity to normalise the idea that any form of online abuse is unacceptable and reports of abuse made to the platform will be managed adequately.

Design Suggestions

1 Offering users the ability to track and manage their reports

Offer clearer guidance for women when they report abuse. This includes clarity around platform policies and community guidelines, content violations, and how product features can support women when they report abuse.

2 Enabling greater capacity to address context and/or language

Provide women with the ability to add context when reporting abusive content. When women are not able to include additional context around the abuse they are reporting, the moderator reviewing the report will not be able to make the right determination about her report. This is particularly important for marginalised communities who may have context/language cues that moderators aren't trained to recognise.

3 Providing more policy and product guidance when reporting abuse

Shift the focus and develop ways reporting processes could be shared by trusted contacts/communities.

4 Establish additional ways for women to access help and support during the reporting process

Give women more control over, and transparency about, the reporting process and build greater trust with users by being clearer about the status of reports.

Offering users the ability to track and manage their reports

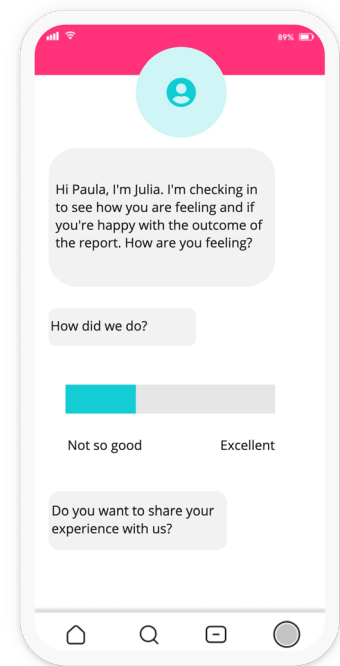
To achieve this you could:

- a Ensure reporting categories are aligned with community guidelines/policies, so that users are not confused by conflicting terminology during the process.
- b Rationalise outcomes of reporting processes by better explaining relevant policies and guidelines, rather than simply linking to them.
- c Create an avenue for users to provide feedback about the reporting experience and outcomes.

For example, allow users to rate their experience and provide the option to fill in a feedback form.

Prototype Example

iMatter provides friendly and fast support through a mix of chatbots and resources; the chatbot guides users through platform resources and options for community support; the tool provides a way for users to rate responses from the platform and feed back on the way a report has been dealt with. [View this prototype in full here](#)



iMatter

Considerations for this suggestion

- As many reports are not classed as policy violations, platforms could build more trust and better manage expectations by being more transparent about the enforcement of their rules alongside the reporting process.
- Consider 'feature fatigue'. While users benefit from greater transparency about processes, too much information may be overwhelming and counterproductive in improving the reporting process.

“

Re-think reporting categories. There are many reasons to report something that aren't currently a possibility inside most apps (e.g. grooming).

- Workshop Participant, Tech Platform

Enabling greater capacity to address context and/or language

To achieve this you could:

- a Provide opportunities for users to provide additional context when reporting, so it is easier to explain to platforms why the content they are reporting is abusive.
For example, if a user is told a report is not in violation of their terms and services, they could be invited to provide more information. Companies can include prompt questions and/or free text fields to guide and enable users to ensure they are only providing only relevant details.
- b Allow users to easily flag if they are translating content they are reporting, or if they want to have the content translated by the company.
- c Allow people to flag if they are reporting in the same language as the abuse, and if not, offer robust translation options, such as allowing people to specify which language they are translating to and from and built-in automatic translation.
- d Provide guidance on how to use reporting systems and information about the outcomes of reports accessible in multiple languages.

Prototype Example

Reporteroo is a reporting dashboard that provides transparency and accountability around the reporting process during and after reporting abusive content. If a report is unsuccessful the outcome is communicated to the user in a way that offers the opportunity for them to provide more context.

[View this prototype in full here ↗](#)

Prototype Example

Reporting 2.0 gives users the ability to report in the language they feel most comfortable with and has a functionality that translates the report into a specific language. [View this prototype in full here ↗](#)

“

If there is an unsatisfactory result, that shouldn't be the end. Rather than saying 'this was rejected', we should say: 'What we have seen does not show a violation of our policy. Would you like to provide more details?'

- Workshop Participant , Academia

Considerations for this suggestion

- Free text fields for adding context may not work well for moderators, because users may not be precise or may provide too much text. Providing prompt questions may help address this.
- Allowing users to add extra context may mean reports take longer to review.
- Filling in free text fields could re-traumatise targets of abuse, so it is important to provide options to report effectively without having to fill these in (such as dropdown menus).
- Bad actors might try to abuse reporting processes, for example by coordinating pile-on reporting to get a user blocked. Systems should be developed to protect against misuse attempts.

Providing more policy and product guidance when reporting abuse

To achieve this you could:

- a Allow users to delegate reviewing content and reporting abuse to trusted contacts.

This could be at an account level or a more granular level, i.e. per post or for a specific amount of time. Allowing trusted contacts to report could be used as a pre-emptive measure; for instance if a trusted contact sees something before the user does, they can start the reporting process straightaway. Trusted contacts should also be able to report and follow the status of their report.

- b Give users the ability to easily approve or reverse action taken by trusted contacts.

- c Allow trusted contacts to opt-in to a process of supporting selected users on an ongoing basis or during specified time periods.

For example, users could allow members of their network to opt-in to help report content during a wave of abuse.

- d Provide the option to connect to trusted contacts and verified organisations to seek support in reporting abusive content.

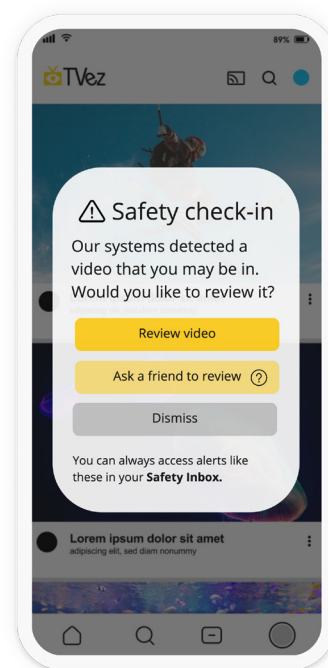


Image Shield

Prototype Example

Image Shield gives users the ability to activate a button that allows their community members to opt-in and support reporting.

[View this prototype in full here ↗](#)

Considerations for this suggestion

- As noted previously, trusted contacts need to be protected, as they could also be targeted. [see Curation, suggestion 4](#)
- Overly relying on trusted contacts could shift the burden from one group of users to another, so it is important that platforms continue to reduce the burden on all users.
- Additional privacy measures such as account verification or two-factor authentication may be needed when trusted contacts are deployed to report abuse, so that they and the user are protected.
- Safeguards need to be in place in case a trusted contact deletes content the user does not think is harmful.

Establish additional ways for women to access help and support during the reporting process

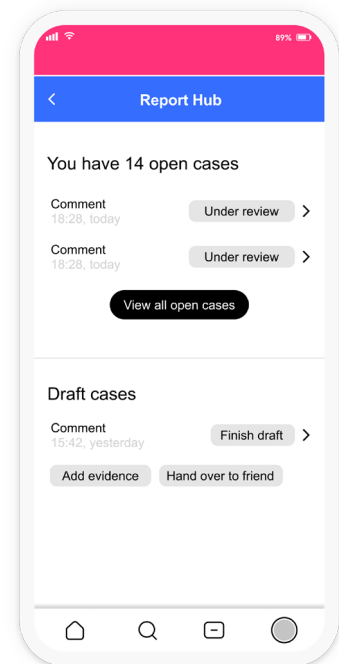
To achieve this you could:

- a** Give users more control over tracking their reports.
For example, allow them to track reports based on the type of content reported; see whether a report is open or closed; reopen a closed report or easily appeal a decision
- b** Enable the ability to archive reported content.
Features may include an automatic evidence-gathering tool to capture and archive proof of abusive content and the ability to save the content in a format that can be shared with employers, other organisations or enforcement officials. Content could be saved in a 'vault' along with a time and date stamp, platform details, and the type of abuse being flagged. A button provided next to comments could allow users to save the commentary for evidence.

Prototype Example

Gateway provides a safety feature in Settings that allows the user to capture and archive abusive content.

[View this prototype in full here ↗](#)



Report Hub

Prototype Example

Report Hub provides options to save and gather evidence of an abusive event quickly as well as the ability to mass report comments.

[View this prototype in full here ↗](#)

“

People find it more valuable when they know what the process is, even if they are not happy with the outcome of the report.

- Workshop Participant, Tech Platform

“

It is not just about giving information but setting some expectations on what will happen and committing to certain milestones.

- Workshop Participant, CSO

Offer users the ability to track and manage their reports

To achieve this you could:

- c** Build a dashboard to allow users to see what is happening with a report they have submitted and how long it may take for them to receive a response.
For example, users are able to track key milestones in the reporting process and are shown where in the process their report is.
- d** Ensure the reporting dashboard/hub is easy to find from key pages on the platform (e.g. home page, profile page).
- e** Enable users to more easily report multiple abusive content at once.
For example, provide a mass upload template for people who need to manage a pile-on of abuse in which more than one person is attacking the user.
- f** Be careful not to retraumatise users through the reporting process.
For example, offer the option to toggle on/off the visibility of report content.

Reporteroo

Prototype Example

Reporteroo's reporting dashboard can be accessed through the profile page; users can track their reports based on the type of content, check whether a report is open or closed, see statistics on the outcomes of any reports and easily reopen a report or appeal a decision. A mass upload template is available for people who need to report multiple pieces of abusive content at the same time.

[View this prototype in full here ↗](#)

Considerations for this suggestion

- It is important to consider how content can be archived in a way that can be used as valid evidence in a legal process.
- Archives could be legally problematic if they store content featuring children or minors, or that is found to contain child sexual abuse material.
- Transparency should come with accountability. If the user is offered clarity on what to expect, they should also be able to trust that those expectations will be met.
- Creating a dashboard that shows no one has responded to a request could be demoralising, so it is important that reporting processes are improved as well as displayed more clearly.

Next Steps

“While we can’t quickly unwind the sexism that drives abuse, we can redesign our digital spaces and change the online environments that allow this misogyny to thrive.”

Azmina Dhrodia,

Web Foundation Senior Policy Manager for Gender and Data Rights

The Web Foundation will continue our ground-breaking work at the intersection of gender and data rights, including further developing our dedicated workstream focused on solutions to online gender-based violence and abuse.

Our long-term vision is to continue to bridge the gap between technologists, policymakers, researchers, and those who use online services - applying the right mix of expertise and real experience to create effective, workable product and policy solutions to some of the biggest technology challenges of our time.

We will do this through the Web Foundation’s Tech Policy Design Lab - part of a new phase of tech policy development where solutions are developed on the basis of sound evidence - not reactionary politics; where there is meaningful multi-stakeholder participation; where people and their online experiences drive policy and product design; and where solutions take into account the full diversity of those who use digital tools and the prevalent and prominent problems women face online.

[Check here for regular updates on this work](#)

With thanks to

All of the organisations and stakeholders involved in this process. **[View the full list here](#)** »

Azmina Dhrodia, Eve Kraicer and Emily Sharpe at **[the Web Foundation](#)**

Dr. Charlotte Webb, Conor Rigby and Ruth Marshall-Johnson at **[Feminist Internet](#)**

Kernow Craig and Fiona Morrisson at **[Craig Walker](#)**

Gunes Kocabag at **[Outsight International](#)**

Priyanka Jain, Janice Dean & Paula Kawakami at **[3X3 Design](#)**

Report and website design
[Conor Rigby](#)

To find out more about this project and for more information on how to use this work at your organisation please contact:
ogbv@webfoundation.org