DIGITAL RESPONSES TO CRISES:

AN ACTION PLAN FOR PLATFORMS AND CSOS CONFRONTING ONLINE THREATS

NDI
NATIONAL DEMOCRATIC INSTITUTE

DIGITAL RESPONSES TO CRISES:

AN ACTION PLAN FOR PLATFORMS AND CSOS AND CONFRONTING ONLINE THREATS

SEPTEMBER 2023

Website: www.ndi.org

ACKNOWLEDGMENTS

These findings and recommendations are the result of feedback and input from many people dedicated to safeguarding digital safety and information integrity. The development of this action plan was led by the National Democratic Institute (NDI) with input from Chihhao Yu (Taiwan Information Environment Research Center), Dhanaraj Thakur (Center for Democracy and Technology), Emmanuel Maurice Diokh (Responsible Internet Without Borders Senegal), Farzaneh Badiei (Digital Medusa), Francisco Gonzalez (Movilizatorio), Khattab Hamad, Márcio Vasconcelos (Technology and Equity Institute), and Sarah Oh (T2). NDI would like to thank the representatives from civil society and industry who participated in a global roundtable discussion for sharing their experiences, feedback, and ideas. Leadership and direction were provided by NDI's Senior Advisor for Information Strategies, Daniel Arnaudo, and Democracy and Technology Program Officer, Rachelle Faust. NDI staff who contributed to the action plan and research include Matthew Graydon, Kaleigh Schwalbe, and Calvin Johnson. Recognition is also due to colleagues on NDI teams in Washington, DC, Taiwan, and Ukraine who supported the roundtable discussion. NDI gratefully acknowledges the support of the National Endowment for Democracy (NED) for providing the funding for this work.

About the National Democratic Institute

The National Democratic Institute (NDI) is a nonprofit, nonpartisan, nongovernmental organization that responds to the aspirations of people around the world to live in democratic societies that recognize and promote basic human rights. Since its founding in 1983 as one of the four core institutes of the National Endowment for Democracy, NDI and its local partners have worked to support and strengthen democratic institutions and practices by strengthening political parties, civic organizations and parliaments, safeguarding elections, and promoting citizen participation, openness and accountability in government.

CONTENTS

Executive Summary

Drawing from a review of actions taken by platforms during recent crises, guidance and recommendations on crisis response published by other organizations, and discussions with a wide range of NDI partners from global civil society and industry, **this action plan examines gaps, needs, and areas of opportunity for platforms and presents a set of recommendations for strengthening digital responses to future crises.**

It is clear from reviewing these sources that there are misalignments between how platforms are responding to crises and what actors on the ground actually need. Specifically:

- Platforms are heavily invested in policy and product interventions to crises, but **external stakeholders want more engagement and transparency.**
- **Transparency, human rights impact assessments, and data access were all identified as priorities by civil society,** but are not among the main actions taken by platforms.

Feedback from stakeholders included a range of recommended actions that platforms can take to better protect users throughout the different phases of a crisis. These actions include:

Overarching	Before a crisis
• Engage with external partners • Provide greater transparency • Understand human rights risks • Provide safe access to data • Equitably respond to global crises • Ensure accountability and fairness	• Establish monitoring and early warning systems • Map policy and product intervention “levers” • Build and adequately resource crisis-focused teams • Invest in localized knowledge
During a crisis	**After a crisis**
• Identify and protect vulnerable groups and individuals • Enforce policies consistently • Ensure data can be preserved for accountability • Consider secondary impacts and unintended consequences	• Define de-escalation criteria and regularly reassess • Communicate product or policy changes • Learn from past responses • Enable external audits

Introduction

Crises and other critical democratic moments, such as political events, present unique and heightened risks to the information space. These risks can be seen in recent crises like the invasion of Ukraine, the Tigray War in Ethiopia, and the Taliban takeover of Afghanistan, where mis- and disinformation, hate speech, coordinated manipulation, and other harmful online behavior and content have shaped and often exacerbated armed conflict and civil unrest. Prioritizing the safety of vulnerable groups and individuals and the integrity of the information space during these fragile periods is essential.

Technology companies that own and operate social media platforms have the potential to serve as powerful allies in civil society's work to protect the information space during crises. These companies control significant access to the tools and data needed to understand the online environment and promote internet freedom and information integrity. Additionally, they have the ability to provide safe and accessible mechanisms for participation in democracy, especially in closed information spaces. While technology companies are responding more quickly to crises and other critical democratic moments, NDI has observed that learned best practices are seldom implemented beyond high profile incidents, and rarely outside of the United States, Europe, and other high-value markets. There is a clear need for a more consistent, informed, and cohesive approach, with industry, civil society, and, where appropriate, government working in partnership.

The following plan presents a range of actions that technology companies can take to strengthen digital responses to crises, drawn from research into best practices as well as feedback from civil society, government, and industry stakeholders. This guidance complements NDI's existing interventions on ending online violence against women and recommendations for expanding civil society researchers' access to platform data. By adopting relevant recommendations from this plan and collaborating with civic partners, technology companies can enable greater protection for their users at times when they need it most.

Methodology

To develop this action plan, NDI hosted a roundtable discussion with more than fifty representatives from global civil society and industry to gather feedback and identify recommendations and additional lines of investigation. NDI also reviewed and identified gaps and areas for improvement in previous platform responses to crisis moments, as well as guidance and recommendations developed by other organizations working in this space. In parallel with this research, NDI conducted additional in-depth consultations with experts from civil society and industry based in Africa, Asia, North America, and South America.

Application

These recommendations are not intended to be exhaustive, but instead offer a starting point to inform swift and effective platform responses in crisis contexts. The technical implementation of the recommended actions will inevitably differ depending on the type of crisis, the country where the crisis occurs, and the internal structures and function of the platform, and has thus not been specifically prescribed in this action plan. NDI encourages additional research into best practices for digital responses to crises and potential implementation strategies based on these findings and recommendations.

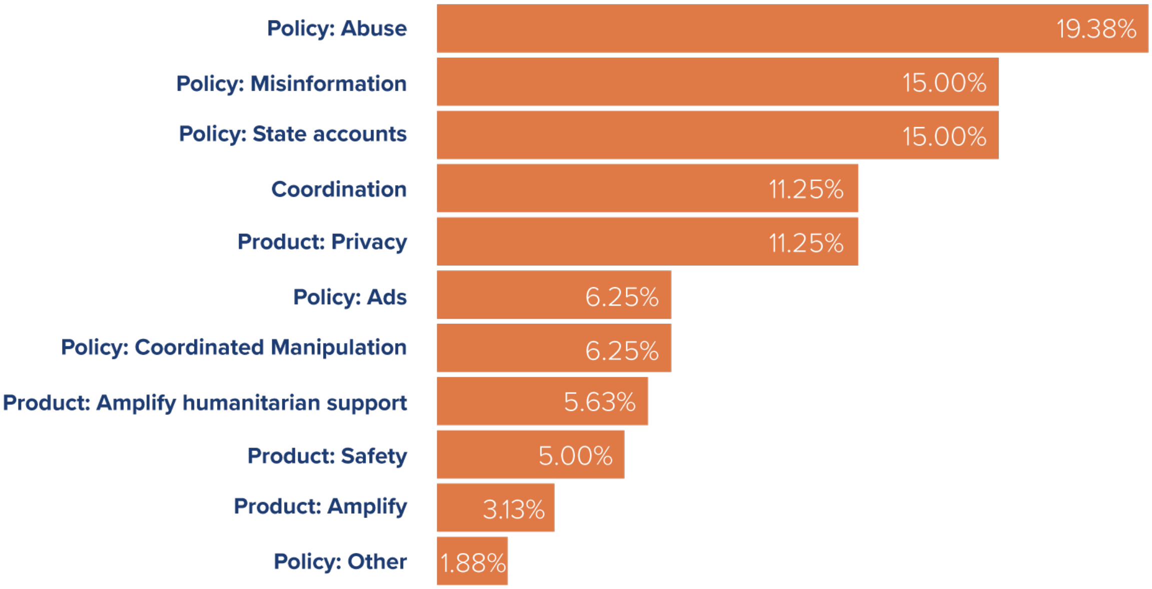
Definitions

- For the purposes of this action plan, a **crisis** is generally defined as "an unexpected situation that impacts the lives of many citizens in a country or region." This includes armed conflicts, civil unrest, and natural disasters.
- **Critical democratic moments** include crises, as defined above, as well as political events (e.g. elections and referenda).
- **Technology companies** are primarily defined under this action plan as companies that own or operate social media platforms, but can also include companies that have developed other digital technology products.
- A **platform** is a software or hardware system, e.g. the Instagram social media platform owned by the technology company Meta.

SUMMARY OF FINDINGS

How have platforms responded to crises?

Based on a review of publicly available information, the actions taken by the major tech platforms[1] in response to recent crisis events (specifically the full-scale invasion of Ukraine, civil conflict in Ethiopia and Myanmar, and the Taliban takeover of Afghanistan) can be grouped into five categories. These are: 1) **enforcement of abusive behavior policies,** such as hate speech or incitement to violence; 2) **enforcement of misinformation policies;** 3) **enforcement of policies related to government and state media accounts,** including removing or limiting the reach of those accounts; 4) **coordination** with internal stakeholders (for example through establishing a dedicated task force) as well as external partners; and 5) **product interventions focused on protecting user privacy.**



Additional platform interventions tracked across these crisis events included enforcement of policies related to advertising (e.g. pausing advertising in impacted countries) and coordinated manipulation (e.g. removing bot networks and other forms of inauthentic behavior). Platforms also launched product features focused on directing people to humanitarian support resources, protecting user safety, and amplifying credible information.

Examples of specific interventions

- Twitter launched Search and Home Timeline prompts to surface critical digital safety and security resources in English, Ukrainian, and Russian at the outset of the full-scale Russian invasion of Ukraine.
- Google locked the Gmail accounts of former members of the Afghan government to prevent sensitive information from being seized following the Taliban takeover of the country.
- Meta expanded its hate speech enforcement to encompass a more extensive list of slurs across the four main Ethiopian languages during the Tigray War in Ethiopia.

[1] See the Appendix for a full list of the platforms and events reviewed.

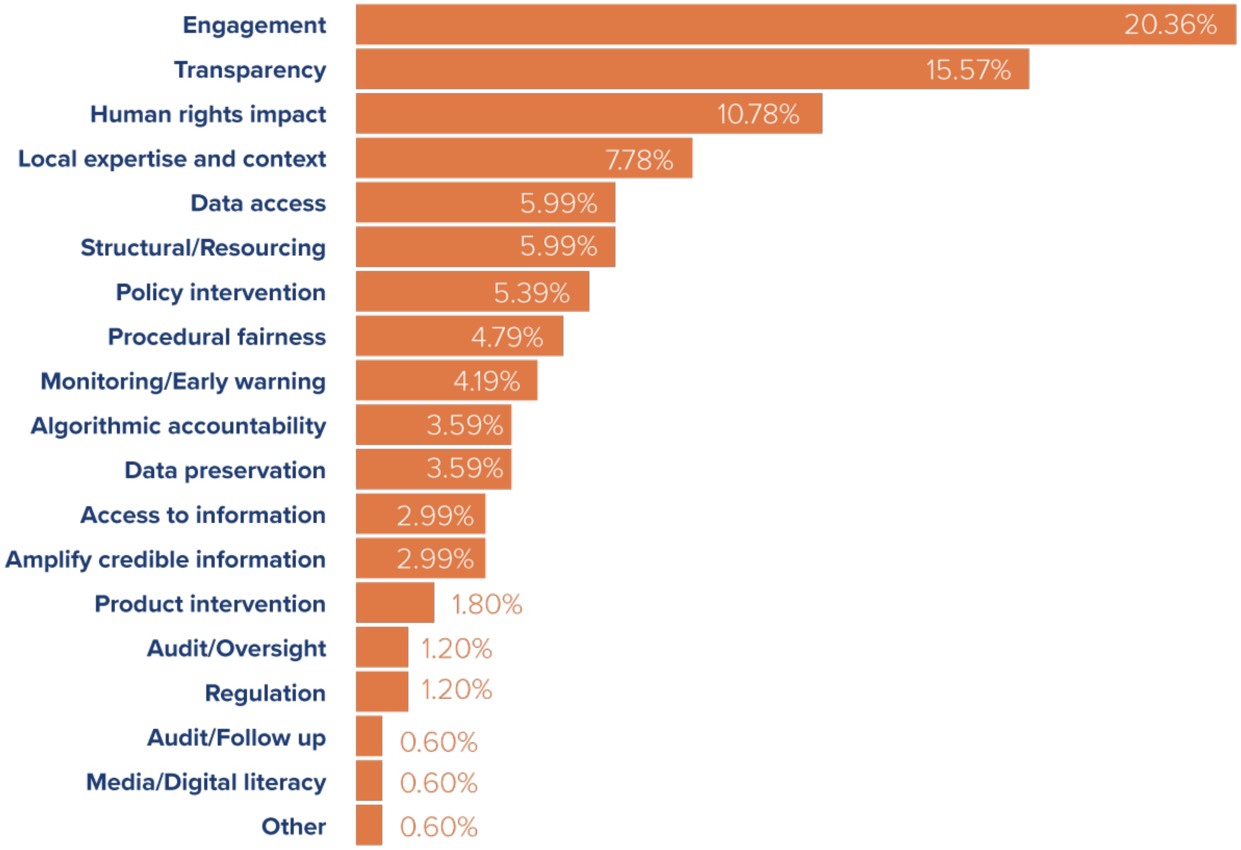
Takeaways

- **Information sharing from platforms regarding their responses to crises and other critical democratic moments has been inconsistent.** While platforms shared details regarding their responses to the full-scale invasion of Ukraine and the Tigray War in Ethiopia in blog posts and official statements, for example, notably less information was available on actions taken in response to the Taliban takeover of Afghanistan. In the case of other crises, such as the 2023 civil conflict in Sudan, no publicly available information from the major platforms was identified at all.
- **There is a lack of insight into the impact and effectiveness of platform interventions.** Specific data on interventions shared by platforms is often presented without broader context, e.g. "X pieces of content removed" or "launched Y safety resource." Without additional information on the scale of the problem, it is difficult to judge whether or not these interventions have sufficiently addressed the targeted issue and if resources are being invested effectively.

Civil society recommendations to platforms

Between June and July of 2023, NDI conducted a series of individual and group discussion sessions with stakeholders from civil society and industry to capture their experiences and feedback on how platforms have responded to recent crises and other critical democratic moments. In addition, NDI reviewed existing recommendations and guidance related to digital responses to crises in order to identify additional key themes.

The findings from these discussion sessions and published recommendations were coded by theme, as reflected in the chart below. A full breakdown of these findings can be found in the Appendix.



The major themes identified through this survey of civil society were:

- **Engagement.** Engagement was the issue most cited as a critical gap or need by a significant margin. Stakeholders called out the need for more consistent and accessible communication with platforms, including through dedicated reporting channels as well as regular briefings and consultations. Contact with frontline policy and operations staff was also emphasized. Several stakeholders noted additional challenges related to engagement and communication stemming from recent tech industry layoffs, and described situations where their platform contacts "suddenly disappeared." One civil society respondent also made the broader observation that "platforms don't see civil society as equal collaborative partners."
- **Transparency.** The need for greater transparency from platforms, both to enable independent research and for purposes of accountability, was also a leading theme. Published recommendations sought greater visibility into how platforms make and enforce content moderation decisions, particularly in relation to algorithmic ranking and automated enforcement; disclosures made to government and law enforcement; and exceptional measures taken during crisis moments. A civil society respondent also highlighted the need for trust and safety processes to be transparent and open source so that they can be replicated by emerging platforms.
- **Human rights impact.** Recommendations focused on ensuring that platforms conduct ongoing and meaningful human rights due diligence. In line with the previous theme of transparency, there were also calls to make human rights impact assessments fully public and transparent. Specific considerations for protecting the rights of vulnerable and marginalized groups, such as women in politics, were emphasized as well.
- **Local expertise and context.** Multiple stakeholders called out gaps related to platforms' understanding of local language and cultural context. "Companies may have data on likes, etc., but they have gaps when it comes to language expertise and cultural understanding," said one respondent. Recommendations emphasized hiring staff with the necessary local and regional expertise and language skills to effectively enforce policies. The important role that local organizations can play in filling this gap was also highlighted, with another respondent noting that "even the wealthiest tech companies can't have the range of [government and civil society organizations]."
- **Data access.** Data access remains an area of concern among stakeholders, with one civil society respondent noting that "platforms are not very cooperative in terms of extending access to data sets and there are inherent limitations to data sets that don't fully satisfy research needs." Furthermore, stakeholders called out the need for more inclusive application programming interface (API) policies that would extend access to civil society and journalists in addition to academia. Published recommendations called for platforms to expand API and dataset access, while also ensuring that the users of this data are adequately vetted.
- **Structural/Resourcing.** Recommendations emphasized the need for platforms to devote greater financial and human resources to crises and, more broadly, to content moderation in global majority countries. Several recommendations also focused on the importance of developing dedicated cross-functional teams for rapid crisis response.
- **Policy intervention.** Specific recommendations related to policy development and enforcement included several calls for platforms to more consistently identify and remove repeat instances of violative content. A proposal to lower thresholds for actioning content and accounts during a crisis was also made. One civil society respondent pointed out the need to shift from content-level to actor-level moderation, as problematic content is often created by the same actors.

- **Procedural fairness.** Ensuring that users had visibility into moderation decisions affecting their content and accounts, and the right to appeal those decisions, was another prominent theme. Recommendations also emphasized the importance of free expression and aligning platform remediations (e.g. content removal) proportionally with the degree of harm.
- **Monitoring/Early warning.** The need for platforms to proactively monitor high-risk areas and situations that may become crises came up frequently. Specifically, recommendations called for platforms to develop clear indicators and early warning systems for potential crises, in alignment with the crisis-focused resources highlighted above.
- **Algorithmic accountability.** Lastly, recommendations called attention to the risks presented by automated moderation tools and the need for a consistent "human in the loop" to ensure accuracy. In line with the previous themes, the need for transparency into how algorithmic enforcement tools are developed and operated was also underscored, along with the necessity of appeals systems.

Recommendations by time period (ranked by frequency cited)

Pre-crisis	*During a crisis*	*Post-crisis*	*Overarching*
Human rights impact	Engagement	Transparency	Engagement
Structural/Resourcing	Transparency	Access to information	Transparency
Monitoring/ Early warning	Data preservation	Audit/Oversight	Local expertise & context
	Policy intervention		

Additional feedback

- **Bright spots.** Several stakeholders highlighted successful bilateral engagement with tech platforms as a bright spot. For example, one civil society respondent noted that it has been relatively easy for them to reach contacts at most of the major platforms via email when they have an issue to discuss. Some stakeholders also pointed to instances where platforms directly or indirectly adopted their recommendations, as in the case of a civil society respondent who cited a platform policy change that came about as the result of a blog post their organization published. Funding support from platforms, particularly for fact checking and media literacy initiatives, was identified as an additional bright spot.
- **Areas of opportunity.** Stakeholders emphasized the effectiveness of regulation (provided that it respects free expression and other individual rights) as a forcing function for more consistent communication and engagement as well as the adoption of more robust policy and product protections. Stakeholders also noted that change can be driven by pressure from a platform's community of users, such as the push by public health authorities and civil society organizations to introduce verification tools on Telegram during the coronavirus outbreak. In addition, the role that civil society can play as a trusted intermediary between platforms and user communities was identified as an area of opportunity.

- **Other insights.** Additional notable feedback from stakeholder consultations included the observation that while social media and cybersecurity are often the focus during crises, other areas of the tech sector can play a role as well. For example, during the 2022 Iranian protests, the mobile game "Clash of Clans" was banned by Iranian authorities in an attempt to prevent its chat feature from being used by activists for coordination. Stakeholders also noted the importance of cross-sector alignment on defining a "crisis," as definitions varied across regions and contexts. Lastly, a civil society respondent emphasized the need for all stakeholders supporting digital responses to crises to avoid duplication and redundancy by focusing on existing solutions, saying "most of the time there is already a solution to a problem, but stakeholders often seek to come up with a new solution. It's important to put ego aside and focus on impact."

Alignment between interventions and recommendations

When comparing the main areas of platform intervention with the gaps and needs cited in the feedback and recommendations from civil society, some clear areas of alignment and misalignment become evident.

Platform interventions	*Recommendations*
Ranked by frequency cited	
1. Policy intervention 2. Product intervention 3. Amplify credible info **4. Engagement** **5. Local expertise and context**	**1. Engagement** 2. Transparency 3. Human rights impact assessments **4. Local expertise and context** 5. Data access

Takeaways

- **Platforms are heavily invested in policy and product interventions to mitigate online harms during crises, but external stakeholders want more engagement and transparency.**
- **Platform interventions overlapped with civil society recommendations across two categories: engagement and local expertise and context.** However, it is clear that even where platforms are taking actions that align with feedback and recommendations, there is more work to be done. For example, while platforms regularly highlight engagement with civil society and partners as a key component of their response to crises, it is also the gap or need most frequently cited by stakeholders.
- Actions related to **transparency, human rights impact assessments, and data access were all identified as priorities by civil society,** but do not appear among the top actions taken by platforms.

ACTION PLAN FOR PLATFORMS

Drawing from the civil society recommendations and feedback outlined in the previous section, as well as the learned experiences of NDI and its partners, the following action plan includes specific steps that platforms can take to better protect users and respond to challenges during crises and other critical democratic moments. The recommended actions are arranged by time period (before, during, and after a crisis, and overarching), as outlined below.

Overarching		
Engage with external partners	Understand human rights risks	Equitably respond to global crises
Provide greater transparency	Provide safe access to data	Ensure accountability and fairness
Before a crisis	**During a crisis**	**After a crisis**
Establish monitoring and early warning systems Map policy and product intervention "levers" Build and adequately resource crisis-focused teams	Identify and protect vulnerable groups and individuals Enforce policies consistently Ensure data can be preserved for accountability Consider secondary impacts and unintended consequences	Define de-escalation criteria and regularly reassess Communicate product or policy changes Learn from past responses Enable external audits

Before a crisis

1. **Establish monitoring and early warning systems.** While some sudden onset crises are difficult to predict, others present warning signals that can be detected in advance. These signals can come in the form of external events, such as an increase in political violence, or indicators that are internal to platforms such as a spike in abuse reports from users. To maximize time for advanced preparation, platforms need to develop monitoring systems that can consistently detect these signals and ensure they are communicated to the necessary internal teams. External vendors can also be engaged to provide this service, although systems are often most effective when they integrate internal data.

Case study: Ukraine

While there was uncertainty surrounding Russia's intentions in the lead up to the full-scale invasion of Ukraine in 2022, the possibility of expanded conflict was clear from the heightened rhetoric and buildup of troops along the border beginning in October 2021. This provided platforms with several months to anticipate potential risks and begin prepositioning policy and product resources for activation if and when needed.

A former Twitter employee highlighted the benefits of this advanced planning mentality, which reduced the platform's implementation time for a key user safety intervention from weeks to hours by anticipating needs and prepositioning language, design, and engineering assets.

2. **Map policy and product intervention "levers."** Platforms should map out in advance the policy and product intervention levers that can be activated during a crisis. These interventions can vary widely depending on the nature of the platform and the needs of a given crisis, and it is important to ensure that the appropriate interventions are applied. For example, while disabling location features may make sense for protecting user privacy and safety in the context of an armed conflict, those same features could be vital for directing humanitarian aid during a natural disaster. See the previous section for additional examples of platform interventions from past crises.

3. **Build and adequately resource crisis-focused teams.** Crises can touch on many aspects of a platform's operations, from policy and engineering to sales and human resources. It is crucial to determine the necessary stakeholders, establish crisis-focused teams, and provide adequate financial support for staffing and coordination. This internal cohesion is also important for providing clear and unified messaging to external partners. Ideally, these teams would be solely focused on crises, although smaller platforms can also consider designating crisis focal points within existing teams.

    a. Given the similarities in policy and product interventions as well as the stakeholder teams, elections-focused resources can be potentially leveraged for crisis response.

4. **Invest in localized knowledge.** Effectively addressing safety risks in the context of a crisis requires nuanced understanding of local context, culture, and language. Hate speech, for example, can be highly context-specific and include slang or codewords that only someone with local knowledge can detect (see the case study below). Platforms should ensure that this risk is mitigated by:

    a. Hiring staff with local expertise and language skills for regions that are likely to be impacted by crises.

    b. Building relationships and communication channels with NGOs, activists, and other representatives from at-risk regions and localities.

Case study: Ethiopia

Hate speech in the context of ethnic conflict in Ethiopia provides one example of the complexities of local language and culture that require specialized knowledge for effective on-platform enforcement.

Peace Tech Lab's Ethiopia Hate Speech Lexicon identifies numerous terms that have hateful and inflammatory connotations but may appear innocuous to an uninformed outsider. For example, while the literal meaning of the term ወላሞ (Wollamo) in Amharic is to "eat a fig (tree)," in practice "the attributes associated with this term of crime and theft promote prejudices against southern communities."

During a crisis

1. **Identify and protect vulnerable groups and individuals.** Certain groups and individuals are particularly vulnerable during crises. These include people like activists and journalists as well as members of ethnic minorities. Aid workers are another group that faces acute risks from misinformation, targeted harassment, and other online harms during a crisis. It is vital for platforms to map these vulnerable actors and support them with appropriate policy and product interventions, which can include:

    a. Policies specifically addressing the targeting of certain vulnerable groups, such as aid workers.

   b. Proactively providing heightened monitoring and support to the accounts of vulnerable groups and individuals.

2. **Enforce policies consistently.** If policies are not enforced consistently, platforms run the risk of exacerbating the very harms they are aiming to mitigate. Detecting and removing hate speech in one local language and not another, for example, can put users at risk and also lead to perceptions of bias. Tying into the earlier point regarding localized knowledge, platforms need to ensure that they have insight into local contexts in order to enforce policies consistently and effectively.
3. **Ensure that data can be preserved for accountability.** During a crisis, content posted online can provide key evidence for war crimes investigations and other forms of justice and accountability. Platforms should ensure that investigators from governments, international organizations, and civil society (including NGOs, academia, and media) are able to communicate their data preservation needs and, where appropriate and in line with user privacy rights, have access to platform data.
   a. Communicate regularly with external partners to understand and anticipate their data preservation needs.
   b. Create systems to allow vetted partners to flag content for potential preservation.

4. **Consider secondary impacts and unintended consequences.** In the heat of the moment during a crisis, it can be easy to lose sight of the potential secondary impacts and unintended consequences of policy or product interventions. It is important to consider these outcomes and safeguard against the possibility of creating or exacerbating harms. For example, while some government accounts may be antagonists in the context of a crisis, and platforms may consider deamplifying or removing them, these same accounts may also provide vital information to citizens on government services

Case study: Syria

The Syrian Archive, a project run by the nonprofit organization Mnemonic, has tracked over 350,000 videos from the conflict in Syria that have been removed by social media platforms for violating content moderation policies. These videos include documentation of chemical attacks, attacks on hospitals and medical facilities, and destruction of civilian infrastructure. The loss of these public records can impact potential criminal case building as well as human rights research.

After a crisis

1. **Define de-escalation criteria and regularly reassess.** Equally important to knowing when to activate a crisis response is understanding the changes in risk factors that can allow for safe de-escalation. Having a clear process in place helps to ensure that resources are allocated effectively across global crises. Platforms should establish consistent and transparent criteria for determining when crisis resources are deactivated, and regularly re-evaluate situations to ensure that risk factors have not re-emerged.

2. **Communicate product or policy changes.** Ensure that any product or policy changes that have been deactivated or reverted following a crisis are transparently communicated to users. For example, if a product feature was turned off as a safety measure during a crisis, users should be informed as to when and why it has been turned back on.

3. **Learn from past responses.** Effective responses to crises demand a learning approach. This includes utilizing tools such as retrospectives and after-action reports following a crisis response. These evaluations should consider what actions were and were not effective as well as any opportunities for action that were missed. Improving future responses relies on this critical reflection.
    a. For protracted crises, consider a regular cadence of retrospectives to identify successes and gaps and allow for course correction.

4. **Enable external audits.** In addition to conducting internal retrospectives, platforms should provide external partners with transparency and data access to enable independent auditing of actions taken during a crisis. Platforms should also commit to engaging with external auditors and adopting relevant recommendations.

Overarching

1. **Engage with external partners.** Siloed efforts cannot effectively address the complex safety and integrity threats created by crises. While not a panacea, collaboration across sectors is essential given the unique strengths and resources that each stakeholder group brings to the table. As noted in the previous section, for example, platforms need the local context and expertise offered by civil society organizations in order to consistently enforce their policies. Similarly, civil society can benefit from the unique insights into on-platform behavior and trends that only platforms can provide. More frequent and open communication between all stakeholders also helps to bridge gaps in understanding between industry, civil society, and government that often hinders productive collaboration. Platforms can foster closer engagement by:
    a. Creating channels for communication such as reporting lines and regular meetings.
    b. Participating in multi-stakeholder forums, both within industry (e.g. the Digital Trust & Safety Partnership) and across sectors.
    c. Facilitating direct relationships with on-the-ground policy and operations teams.

Engagement with governments

While government can be an important partner in the "whole of society" approach to protecting the information space, collaboration with government should be transparent and ensure respect for user privacy and other rights.

See NDI's *A Practical Guide for Civil Society Organizations During a Crisis* for additional guidance and best practices related to civil society engagement with government.

2. **Provide greater transparency.** By providing greater visibility into systems and decision-making processes, platforms can enable research that creates a better understanding of threats and mitigations. Transparency also fosters critical checks and balances that can help platforms improve their approaches to crises through identifying gaps and areas for improvement. Keep in mind, however, that transparency actions must be balanced with the risk of divulging information that can be exploited by bad actors. Actions that platforms can take to improve transparency include:
    a. Providing visibility into how policies are developed and enforced (particularly in relation to automated enforcement mechanisms).
    b. Consistently communicating actions taken in response to crises, and the mechanisms underlying these approaches.
    c. Sharing data related to policy and product intervention efficacy, as further outlined below.

3. **Learn from past responses.** Consideration of human rights risks and impacts needs to be a fundamental part of platform approaches to crises. This includes conducting human rights impact assessments prior to, during, and following crises as well as embedding the protection of human rights into policy and product design. Specific ways that platforms can better understand and safeguard human rights include:
   a. Incorporating staff with expertise on human rights into the policy and product development process.
   b. Soliciting regular feedback and guidance from external human rights experts.
   c. Adhering to global norms and standards like the UN Guiding Principles on Business and Human Rights.

4. **Provide safe access to data.** In line with the importance of greater transparency described above, platforms can strengthen their responses and resilience to crises by enabling independently vetted partners to access data for research and accountability purposes. Platforms should build relationships with these partners to understand their data needs and seek to develop equitable solutions for sharing data. At the same time, platforms should ensure that any access to user data is provided in accordance with global best practices on privacy and data protection, and that partners are adequately vetted.
   a. Open data access to journalists, civil society organizations, and other researchers beyond academia.
   b. Ensure that data access is globally equitable, including by providing low- or no-cost API access options for vetted partners.

5. **Equitably respond to global crises.** Crises can impact any country or region, and platforms should ensure that response resources are allocated equitably. Potential harm to safety and security, particularly to vulnerable groups and individuals, should be the predominant factor determining platform interventions in a given country or crisis situation.

Case study: Meta's Trusted Partner program

A 2023 report by Internews on Meta's Trusted Partner program found "a significant disparity of service between Ukraine and other countries, including countries that are also experiencing major armed conflict, internal displacement, and political disinformation." The report notes that while Ukrainian partners generally received a response from Meta within 72 hours, equivalent reports from Ethiopia relating to the Tigray War went unanswered for several months.

6. **Ensure accountability and fairness.** Crises present heightened risks that can lead platforms to err on the side of policy over-enforcement. Any changes to standard policy enforcement during a crisis should be transparently communicated to users. In addition, users should have the ability to understand why their content or account has been the subject of policy enforcement as well as the means to appeal those decisions.
   a. Be especially attentive to appeals from NGOs, activists, journalists, and other vulnerable users in a crisis context.

Focus on: Emerging platforms

Many platforms can have an impact during a crisis, no matter their size, age, or the type of product or service they offer. Although emerging platforms often lack the same resources as more established platforms, there are still actions they can take to protect users and reduce risk during a crisis. Priority actions for emerging platforms include (but are not limited to):

- Engaging with external partners.
- Understanding human rights risks.
- Developing policy and product intervention levers.
- Enforcing policies consistently.

Emerging platforms are also encouraged to share information and best practices to collectively develop crisis response tools and resources, ideally from an open-source perspective.

RECOMMENDATIONS FOR CIVIL SOCIETY

In addition to recommendations for platforms, feedback from partners identified several areas of focus for civil society organizations:

1. **Build coalitions and speak with one voice.** Partnerships between like-minded organizations can amplify advocacy efforts. These partnerships can be particularly valuable for smaller organizations. By speaking with one voice, civil society organizations can more efficiently and impactfully convey their needs and concerns to industry.

2. **Learn from other sectors.** Effective coalitions have emerged in other online safety spaces, such as child safety and countering violent extremism, that can serve as models for collaboration. Civil society organizations working to protect the information space during crises can learn from these existing coalitions to leverage lessons learned and best practices.

3. **Mix closed-door and open advocacy.** Civil society organizations should adopt a mix of behind-the-scenes collaboration and public engagement with platforms in order to maximize impact. Closed-door advocacy encourages more open sharing of feedback and ideas, and provides a platform for constructive dialogue. Simultaneously, open advocacy can help to raise awareness, mobilize support, and exert external pressure.

APPENDIX

Research findings

Digital Responses to Crises: Research Findings

List of recommendations reviewed

Title	Organization	Date
Tech and Journalism Crisis and Emergency Mechanism: consultation's key takeaways	Global Forum for Media Development	Mar 2023
Declaration of principles for content and platform governance in times of crisis	Access Now	Nov 2022
Regarding the creation of a unified position ("one voice") of Ukraine to global tech platforms to fight against disinformation and fakes	Center for Strategic Communication	Jul 2022
How online conspiracies about Syria cause real-world harm	Institute for Strategic Dialogue	Jun 2022
Twitter, YouTube ignore takedown requests by the Ukrainian Government	Reset	Jun 2022
Facebook's Content Moderation Failures in Ethiopia	Council on Foreign Relations	Apr 2022
Civil Liberties Groups Urge Social Media Platforms to Better Protect Free Flow of Information in Crisis Zones	Electronic Frontier Foundation	Apr 2022
How Big Tech's Content Moderation Policies Could Jeopardize Users in Authoritarian Regimes	Open Internet for Democracy Initiative	Feb 2022
Policy Recommendations: Internet Freedom	Freedom House	2022
Social media: A tool for peace or conflict?	Stockholm International Peace Research Institute	Aug 2021
Social Media and Conflict: Understanding Risks and Resilience	Mercy Corps	Jul 2021
Social media companies need better emergency protocols	Brookings Institute	Jan 2021