Social media in peace mediation

a practical framework

June 2021

While mediation remains a fundamentally human-led endeavour, the importance of digital technology in peace processes is growing. Mediators and their teams need to be prepared to use digital tools and address digital risks. This especially applies to social media, given its widespread and rapidly increasing use. According to one estimate[1], by the start of 2021, there were some 4.2 billion active social media users – 53% of the global population – of whom women users accounted for 45.9%. In 2020, 490 million new users joined social media, a growth rate of 13.2%. On average, users have accounts on eight different platforms and spend two-and-a-half hours per day on social media. As a result, social media increasingly shapes political and social interactions and, indeed, people's perceptions of reality.

Social media also plays an increasing role in armed conflicts. It impacts how conflict actors communicate with one another and with the public; how information is disseminated; how the outside world perceives conflicts; and, most fundamentally, how armed conflicts are fought. While the impact varies depending on the context and the stakeholders involved, all types of conflicts are affected, even those taking place in areas with limited access to digital technologies.

Social media also impacts efforts to prevent, manage and resolve conflicts. The United Nations has led efforts to understand these dynamics and develop responses as part of its work on the relationship between digital technologies and international peace and security. The UN Secretary-General has highlighted the importance of social media in his Strategy on New Technologies and reiterated it in his Data Strategy 2020–22.[2] In the area of peace mediation, the Mediation Support Unit of the UN Department of Political and Peacebuilding Affairs (DPPA) launched the CyberMediation Initiative[3] and published the Digital Technologies and Mediation Toolkit[4] to explore the various uses of digital technologies, including social media.

---

1   Data based on Digital 2021: Global Overview Report.

2   The infographic "The UN, Peacekeeping and Digital Technologies", produced by the Center for International Peace Operations (ZIF) in Berlin, provides a useful overview of the various UN initiatives applicable to peacekeeping and more broadly.

3   DPPA, together with swisspeace, the Centre for Humanitarian Dialogue, and Diplo, established the CyberMediation Initiative. It has been further developed to become the CyberMediation Network.

4   DPPA and the Centre for Humanitarian Dialogue jointly developed the toolkit. The accompanying report and an interactive website are available here.

These initiatives all shed light on the fact that the use of social media by both conflict parties and the public is changing the mediation environment in significant ways. For mediators and their teams, this presents both opportunities (for example, facilitating communication and analysis) and challenges (for example, limiting confidentiality and opening a new contested terrain with different stakeholders trying to influence public perceptions of a peace process). What is clear is that mediators need to engage with this new reality in order to tap into the opportunities it presents, while being aware of the related risks and attempting to "Do No Harm".

Against this background, the DPPA Mediation Support Unit and swisspeace teamed up for an evidence-based conversation about the impact of social media on peace mediation, and in 2020 organized a series of workshops and focus group meetings with mediation practitioners, researchers, technology experts and representatives from social media companies. Participants in these consultations prioritized four issues: leveraging social media for gender-sensitive analysis of a conflict-affected context; using social media for communication purposes; understanding and responding to conflict parties' use of social media; and understanding and addressing social media as a source of mis- and disinformation, all within the context of mediation processes.

The report is structured around these four issues. Each section outlines the importance of a specific issue, summarizes the state of play with regard to its significance, makes practical suggestions for mediators and their teams, and poses questions for further consideration and analysis. Illustrative examples are provided in text boxes.

For the purpose of the report, "social media" is defined as comprising social networking websites as well as instant messaging and voice applications that are both publicly accessible and private. The report focuses primarily on the role of social media in Track 1 mediation, or peace mediation, involving official representatives of conflict parties. It provides a snapshot of the current state of play in terms of the role of social media in peace mediation, with the understanding that this and related issues require constant monitoring and observation to keep up with the dynamic and rapidly developing character of social media.

1 Using social media for analysis

Why is this important?

Social media is commonly understood as a communication tool, but it is just as useful as a listening device. For mediators and their teams, social media can be extremely valuable for accessing knowledge about events that are under-reported or censored in traditional media. It can also increase situational awareness and shed light on aspects of conflicts that conventional political analysis may not fully cover. These include trends in narratives about a conflict event, perceptions of key issues among different groups, an overview of influential stakeholders and connections between them, insights into power dynamics, and information about campaigns to spread rumours and disinformation, in particular defamation and hate speech targeting women. These factors can complement existing conflict analyses and thus increase the preparedness of mediators and their teams.

What do we know?

The social media landscape is heterogeneous, contextual and developing very rapidly. If mediators want to utilize social media for analysis and avoid bias, they need to understand the specific ecosystem in which they are working, which is distinct for each country and region. They should consider the following factors when analysing the digital operating environment in a peace mediation context:

- internet use and digital access, including the gender divide in digital literacy and access to the internet;[5]

- demographic, geographic and gender disparities, and the various languages and dialects in internet and social media use;

- the prevalence of particular social media platforms in a given country or region and these platforms' policies and terms of service regarding content moderation and online behaviour;

- regulatory frameworks in a given country or region related to social media;

- the degree of state control of the internet and online content as set out in national policies and legislation relevant to human rights, national security, internet censorship, etc.;

- the cyber capabilities and intent of state and non-state actors, as well as of foreign powers that may seek to influence conflict dynamics;

- influential voices and actors inside and outside the country (including in the diaspora and in states supporting one or other conflict party);

- specific uses of social media in the country or region in question – for example, as a news source – by various political forces, by protest movements, as a form of weaponization, by armed groups for recruitment and representation purposes, etc.; and

- past cases of conflict being fuelled by social media, including trending hashtags and posts that went viral around a specific political or social event.

---

5   According to the OECD report Bridging the Digital Gender Divide, the global internet access rate for women is about 45%, compared to about 51% for men.

Not everybody is on social media, and relatively few users produce most of the content. There are also myriad inauthentic coordinated campaigns, for example by a set of bots run by outside actors. Another caveat is that social media records interaction of some kind, which results in passive or silent behaviour online being perceived as "non-data", that is, effectively invisible. Some posts, in particular those stirring emotions, produce a disproportionate amount of interaction. This means that social media data is not "neutral" and can lead to skewed analysis and misperceptions. Mediators and their advisors need to be aware of this and combine social media analysis with other, more traditional approaches for a more complete picture.

Conflict actors use social media platforms tactically to advance their goals. They may use one platform for disseminating text, another to share video material and a third for internal communication. Or they may use one platform to communicate with international audiences and another to communicate with local constituencies, using different languages in each case. Mediators and advisors need to be aware of this kind of cross-domain use when selecting specific platforms for analysis and when considering which platforms are most useful in disseminating key messages.

Mediation teams also need to be aware of the different types of platforms, including open platforms (for example, Twitter), closed platforms (for example, Signal), and platforms that have both open and closed components (for example, Facebook). Some social media data can be easily accessed for analysis, some can be purchased, and other data is either not immediately available or not available at all. The analysis of open-source data that is accessible on public platforms is more straightforward, although other data might also be available. In any case, as mentioned above, mediators and advisors need to be conscious that social media data is not neutral, nor does it provide an "objective" assessment of public perceptions. Social media data often reproduces and amplifies existing biases, including the biases of those programming the tools used to gather the data.

A variety of methods can be used to analyse social media data. These range from a mediation team's simple monitoring of conflict actor public communications to the use of automated quantitative methods. The latter rely on large amounts of data, but still require context-specific

human-led interpretation. The following are examples of dynamic tools for information gathering and analysis that are particularly useful for mediators and their teams:

- hashtag and keyword analysis, network analysis, user traffic analysis and the analysis of trends over time;

- stakeholder or social network analysis, which can identify connections among actors in a conflict setting; and

- sentiment or perception analysis, which reveals the sensibilities of certain groups in societies affected by conflict and is based on large quantities of data.

Social media analysis is primarily undertaken by research institutions and for-profit companies. Mediators and their teams can use the methods these organizations rely on, but the interpretation and explanation of results need to be adapted to a specific context and to serve the needs of the mediator working in that context. Sensitivity to context will require an understanding of, for example, the languages and dialects spoken, prevalent social and cultural norms, and the different levels of internet access and technical capacities of the actors using social media. In this regard, mediators can draw on the tools developed by humanitarian[6] and peacebuilding organizations.[7]

---

6   See, for example, International Committee of the Red Cross, International Federation of Red Cross and Red Crescent Societies, and UN Office for the Coordination of Humanitarian Affairs, How to Use Social Media to Better Engage People Affected by Crises, online report, September 2017.

7   See, for example, International Alert, Realising the Potential of Social Media as a Tool for Building Peace, online policy paper, November 2020. See also a series of policy briefs on social media, technology and peacebuilding published by the Toda Peace Institute.

Practical suggestions for mediators and their teams:

→ Develop a "dashboard" that provides an overview of the social media ecosystem in the country and region concerned (see the list above for the various ecosystem components), including sex-disaggregated data where available. This dashboard should be established in the preparatory phase of a mediation process and periodically updated.

→ Be aware that social media analysis is not a goal in and of itself. Mediation teams should set clear objectives and choose the appropriate methods they will utilize accordingly.

→ Be guided by international human rights and gender equality norms as well as data protection and privacy standards and establish protocols regarding transparency of data sources. External consultants and organizations working on behalf of mediation teams, including NGOs and private companies, should also be guided in their activities by existing norms and standards.

→ Build capacity within mediation teams to create awareness of the potential value (and limitations) of social media analysis. Teams should be in a position to decide whether to conduct analysis in-house or to collaborate with external partners. In any case, social media analysis should be combined with other conflict-analysis methods.

→ A mediation team should decide early on whether it will use social media analysis in its work and make sufficient resources available to do so, in particular for time-consuming analytical methods that rely on the analysis of large quantities of data.

→ Make social media analysis competency part of the standard toolbox of mediation support units, both in the UN and beyond.

→ Conduct risk assessment to identify the pros and cons of social media analysis and the various social media analysis tools that are available.

→ Identify complementary sources of data that address gaps in key stakeholder contributions to discourse and take into account particular risks to women and marginalized groups who speak up on social media.

→ Consider outsourcing the in-depth analysis of social media. Partnerships with research institutions and NGOs specializing in "peace tech" are promising in this regard. Public–private partnerships with for-profit companies can also be explored, provided that ethical and legal considerations related to privacy, data security, impartiality and archiving are duly considered.

Support for social media analysis from the UN DPPA Innovation Cell

Through its Innovation Cell, DPPA developed in-house approaches to social media analytics in support of conflict prevention, peace mediation and peacebuilding initiatives. These efforts have focused on developing tailored methodologies that can be used in the social media-based analysis of political affairs in a conflict-affected context. They include initiatives to better understand public mobilization in the digital sphere, analyse social media voices, and strengthen the UN's public messaging accordingly. For example, DPPA piloted social media analysis in the context of Arabic dialects to expand early-warning capacities, monitor hate speech and counter incitement to violence. The Innovation Cell also supports special political missions and field presences in conducting baseline studies scoping the use of social media, internet use and other relevant parameters. Lessons learned have been summarized in the UN DPPA checklist for social media analysis. Responding to the growing demand for support, DPPA launched the social media reporting application Sparrow in early 2021. This online tool allows users to analyse trends in publicly available Twitter data, including popular keywords, hashtags, tweets and the level of engagement of accounts related to areas of concern.

ICG analysis of social media fuelling conflict in Cameroon

A 2020 report by the International Crisis Group (ICG) analysed the role of social media in fuelling political and ethnic tensions in Cameroon, in particular after the disputed 2018 elections. The report shows how both pro- and anti-government activists use Facebook in particular to spread inflammatory discourse, hate speech and misinformation, including calls for violence against specific ethnic groups. Facebook officials reportedly attempted to curtail such practices. However, the qualitative analysis conducted by ICG suggests that inflammatory content remains widespread. To de-escalate conflict and more effectively combat hate speech, the report recommends that Facebook should improve its capacity to evaluate inflammatory content in Cameroon, verify the pages of key institutions and influencers, and strengthen relationships with Cameroonian civil society to solicit independent views on identifying hate speech.

2 Using social media for communication

Why is this important?

If mediators want to promote pro-peace narratives and inform the broader public about the peace process, they need to use diverse media channels, including social media. Indeed, negotiating parties and the population as a whole in conflict-affected countries increasingly communicate via social media. Non-state armed groups in particular have used social media in new and sophisticated ways. For mediators, social media offers the potential not just for broadcasting information but for dynamic social engagement. This points to the need to synchronize public communication with a broader outreach strategy in the service of a mediator's mandate and the objectives of a given peace process.

What do we know about it?

Communication via social media presents challenges for mediators in terms of preserving the confidentiality of the process and avoiding a situation where polarizing voices shape the narrative around a peace process. Depending on the style of the mediator, the sensibilities of the parties and the specific phase of a given process, mediators may therefore decide to keep a low profile and engage with the media through generic press statements. However, this comes with risks, as the narrative about a peace process is increasingly shaped in the social media sphere, where parties are expected to constantly feed their constituencies and followers with information regarding the process. Actors fuelling conflict may also assert themselves through social media. In these contexts, having a verified and authenticated social media presence allows mediators to try to proactively shape the narrative and counter divisive statements and incorrect information.

Social media communication offers additional benefits for mediators. An agreement on the ground rules for social media use in a mediation process can serve as a confidence-building measure and help clarify what parties can communicate about the proceedings and outcomes of a mediation process. At critical moments – for example, when an agreement has been reached – mediators and parties can jointly promote a coherent narrative and prevent rumours from spreading.

In addition, communication via social media offers mediators the possibility to address specific actors. Messages can, for example, be targeted to influential actors and decision-makers. They can also be tailored to speak to specific groups – for example, religious and ethnic minorities, women's groups, youth, or members of other social groups. Social media can provide an avenue for political expression for communities not directly engaged in peace processes. It can create space for women to raise their voices and signal their priority demands for the inclusion of particular issues in peace talks. To effectively use social media, mediators need to know who uses particular platforms and how to reach certain groups. They also need to synchronize social media communication with other outreach and inclusivity mechanisms.

More generally, online communication has most impact when the message is consistent with the overall narrative of why peacemaking is necessary. Communication should be frequent and grounded in the strategic use of images, videos, hashtags and keywords. Visual content performs better on social media because the algorithms of most platforms tend to rank it higher and make it more visible and available. Novel approaches such as social media advertisements and working with influencers could further increase the impact of a mediator's message in the social media sphere.

Practical suggestions for mediators and their teams:

→ Weigh the pros and cons of using social media for active communication purposes during a mediation process. Depending on the level of confidentiality, the nature of the process and the style of the mediator, public engagement is sometimes not possible. Social media silence is a viable option in specific cases, provided the risks are carefully assessed and alternative communication channels are both available and effective.

→ Assess the value of communicating via social media during a particular phase of the peace process. For example, it is probably not appropriate to use social media to communicate during confidential pre-talks, but when talks progress, it may be a useful, even necessary means of communication. Social media is particularly valuable after a peace agreement is signed as a way to advocate in favour of the settlement and inform the broader public of its terms,

advantages, etc. This applies in particular in the context of a referendum or some other confirmatory act.

→ Ensure that the decision on whether and how a mediator communicates via social media is part of a broader communication strategy. Hybrid approaches – using social media together with traditional media and in-person forms of engagement – offer many advantages, especially in contexts with low internet access.

→ When engaging actively with the various stakeholders via social media, be aware of the broader social media ecosystem, including gender disparities in social media use and strategies to address populations in key gaps, as well as the specific resources and expertise required within mediation teams to use social media effectively.

→ Mediators' spokespersons and communication teams can use private social media channels to share information in a targeted fashion – for example, with journalists from major national or international media outlets or civil society leaders, including women-led civil society. This can help to shape a positive narrative of the peace process and build allies among opinion makers.

→ Use visually appealing content, including charts, icons, pictures and videos, when communicating on social media. These tools can help to convey and explain complex topics and therefore increase reach and engagement levels. However, since pictures and videos can easily be misconstrued or send unintended messages (for example, through facial expression, unrelated background, etc.), caution should be exercised in order to preserve the perception of impartiality of the mediator among all stakeholders. Mediators should also be aware that visual diplomacy requires a particular skill set and sensitivity.

→ If appropriate, consider working with external partners and influencers who are able to promote pro-peace narratives in online conversations, including women in media and women-led civil society.

Strategic messaging in the context of UN-led mediation processes

UN political offices or missions leading a mediation process have preferred to communicate mainly through standard public information tools and practices, such as press conferences and official statements. More recently, however, some offices have started to utilize social messaging applications for non-sensitive communication as a tool for their strategic outreach and communication activities. For example, some offices communicate via a WhatsApp group with an extensive network of people with specialized knowledge of the conflict and the wider region, including official international and national actors, journalists, activists, and academic and think-tank researchers, among others. This helps these offices to crowdsource and confirm relevant news in greater detail, add context to various news reports, exchange news articles, and instantly share press releases and information about peace process-related events. It also allows UN political offices to receive feedback in an informal yet trusted environment, because the appropriately managed use of the WhatsApp group can ensure a degree of confidentiality.

Pro-peace messages by the UN Office of the Special Envoy for Yemen

The UN Office of the Special Envoy of the Secretary-General for Yemen (OSESGY) uses social media platforms to listen to views and gather feedback on the peace process, and to engage in direct outreach with various Yemeni constituencies. OSESGY operates its social media platforms in a highly polarized environment that is sceptical and distrustful of the UN-facilitated peace process. On Twitter, for example, OSESGY promotes specific pro-peace hashtags in Arabic and English, such as #ShapingPeaceTogether and #StockholmAgreement. When relevant, it uses hashtags promoted by humanitarian actors working in Yemen – for example, #YemenCantWait – to express support for the humanitarian cause and reiterate the call for an urgent end to the conflict. OSESGY maintains an active presence on multiple social media platforms such as Twitter, Facebook and YouTube, and also opened a Telegram channel in October 2020, allowing the Office to disseminate information about the Special Envoy's activities instantly and widely.

Keeping civil society abreast of the Syria peace process

In 2016, the UN Office of the Special Envoy for Syria (OSE-S) established the Civil Society Support Room (CSSR), an inclusion mechanism for Syrian civil society on the margins of the intra-Syrian talks. Through the CSSR process the OSE-S has engaged with hundreds of Syrian civil society actors based inside Syria and in the diaspora. In 2021, through the CSSR implementing partners swisspeace and NOREF Norwegian Centre for Conflict Resolution, OSE-S launched a CSSR website for the purpose of enhancing the sharing of information about and transparency of civil society consultations as an adjunct to the official UN information channels. The website includes an interactive component that allows for intra-civil society dialogue, opportunities for engagement and advocacy, and the possibility for civil society actors to communicate their own inputs and perspectives to OSE-S on issues relevant to the mandate of the UN Special Envoy set forth in UN Security Council Resolution 2254.

Social media consultations with Yemeni women

In March 2021, in partnership with the Office of the Special Envoy of the Secretary-General for Yemen, Build Up held ten focus group consultations over WhatsApp with 93 women from 11 governorates across the country. The aim was to identify women's opinions, perspectives and insights about headlines relating to peace and conflict in Yemen, and how the issues that were identified affected women's daily lives. This was preceded by a mapping process to identify diverse networks of women and by a parallel mapping of conversations on social media to validate and expand on the topics identified. The objective of the consultations was to deepen understanding of concepts mentioned elsewhere by women, determine patterns, bridge narratives, and create dialogue about women's perspectives on areas of interest and significance to them, while reaching out to participants from areas not typically covered by peace process-related consultations in Yemen.

3 Using social media to engage with conflict parties

Why is this important?

There are several reasons why social media matters for mediators when they work with conflict parties. Most obviously, social media offers an easy and cheap way to engage with the parties, allowing mediators to increase the quality and quantity of their outreach and helping to build trust in their efforts and in the process overall. Social media also matters insofar as the online behaviour of conflict parties can exacerbate tensions, widen the political divide between them, and obstruct a negotiation process. When this is the case, mediators may need to engage the parties and address social media as an issue at the negotiation table. This could include, for example, drawing up proposals for protocols and guidelines in which parties agree to moderate their social media behaviours during the negotiations.

What do we know about it?

Mediators' use of social media to communicate with parties has become widespread, in particular using instant messaging platforms. These platforms are used for private communication, but they increasingly serve to spread information publicly – for example, through WhatsApp groups – which is why they can be included in the category of social media. Their use offers many advantages for mediators and their teams, but also poses risks and challenges to the safety of interlocutors and the confidentiality of information and documents. This gives rise to a new set of "Do No Harm" considerations, of which mediators and advisors in general are not yet sufficiently aware. Using social media in contacts with parties may also make it more difficult to archive a mediation team's communications and ensure institutional memory, for example when there are staff changes in mediation teams.

Aside from direct communication, mediators may need to engage with conflict parties to ensure their actions on social media do not undermine a peace process. Problematic behaviours can include confidentiality breaches, discriminatory and misogynistic narratives, escalatory rhetoric, hate speech, posturing, and the targeting of individuals and

their constituencies. Conflict parties can use human or technical proxies to make it difficult to trace messages back to them. They can disseminate disinformation about peace talks or leak confidential information on social media, which often results in mutual accusations. The effects can be significant. For instance, when confidential information is leaked, it can seriously weaken trust among negotiating parties and undermine the public image of both peace efforts and the mediator.

Some mediators have facilitated negotiations among conflict parties on social media behaviour, resulting, for example, in stand-alone codes of conduct or in a protocol within a broader agreement about ground rules in negotiations. Guidance on social media use has also been introduced into peace agreements. In such cases, parties have agreed to a set of measures to guide their online activity, both in terms of what they should proactively do and where restraint needs to be exercised. Such an approach is already manifest in the context of elections, where some political parties have agreed on permissible online behaviours governed by international standards and domestic norms.

Such protocols or codes of conduct are also relevant to mediation processes, particularly where the social media behaviour of parties or their proxies has been problematic and where there are limited incentives to prevent it. In addition to reducing the potential harms of negative social media use, these kinds of agreements have the potential to build trust in both the mediator and the wider peace process, and among the parties. They can help level the playing field by holding all parties, whether they are powerful or not, to the same standards. However, caution should be exercised to ensure that any such agreement is inclusive, respects existing obligations, including with regard to human rights and fundamental freedoms, and does not benefit one party over another.

It is important for mediators to be aware that moderating social media behaviours in any context, let alone that of an ongoing conflict, is highly complex. Agreements are difficult to implement, especially in fragmented settings where negotiators representing conflict parties do not have control over their constituencies and where different views vis-à-vis the peace process exist within a group. Another challenge relates to disputes over the interpretation of the agreement. Because mediation relies on the consent of the parties, mediators will also have to balance issues such as parties' accountability for social media behaviours that contradict the terms of the agreement with the need to secure their continued participation in the process.

In addition, agreements on social media behaviour mostly cover open rather than closed platforms, and usually do not consider the social media behaviour of proxies and external parties. Finally, such agreements do not generally consider the content moderation policies – let alone the decision-making procedures – of the actual owners of the social media platforms. These have proved to be highly unpredictable, particularly in periods of escalated tensions.

Practical suggestions for mediators and their teams:

→ Agree with parties on which platform or application to use for communication purposes. It may be easiest to "meet where they meet" and use platforms that the parties are already familiar with. Regardless of what is decided, due consideration should be given to security risks, preferably within the framework of a broader risk management strategy.

→ Develop appropriate protocols to ensure the secure and effective use of social media platforms and applications to communicate with conflict parties and inform the parties accordingly in order to promote trust and transparency.

→ Ensure that important social media communications with parties are archived in a timely manner.

→ Where appropriate, proactively identify entry points to help the parties and their constituents agree on a set of acceptable and unacceptable behaviours on social media. This could be within the

framework of a voluntary protocol (for example, a code of conduct or something similar) or within a broader agreement.

→ If no formal protocol has been agreed, promote measures to protect the confidentiality of peace talks and clarify with the parties what information can be shared on social media, when to share it, and how to address disinformation when it occurs.

→ Encourage factual, consistent, inclusive and positive messaging on social media around both the peace process and the individuals and groups involved in it.

→ Offer strategic communication support to negotiating parties, especially in situations when the parties have asymmetrical access to social media and significantly different capacities with regard to its use.

Engaging actors in preparation for political dialogue in Libya

In Libya, the Acting Special Representative of the UN Secretary-General, Stephanie Williams, told the UN Security Council in May 2020 that "social media is another theatre of the Libyan conflict". To support this statement, she highlighted that the UN Support Mission in Libya (UNSMIL) engaged actors from across the political spectrum on social media and had produced a code of ethics to tone down hateful or other unacceptable content on social media. When negotiations in the Libyan Political Dialogue Forum (LPDF) started, the UN worked with participants to establish such a code, defining expectations of the Forum's members' behaviour on social media. The code constituted a moral commitment for LPDF members and promoted an online environment that was conducive to the Forum's progress. In late 2020 and early 2021 Ms Williams engaged around one thousand political, societal, youth and women leaders virtually via Zoom and Facebook in preparation for the LPDF.

Countering hate speech on social media in the Libya ceasefire agreement

The Agreement for a Complete and Permanent Ceasefire in Libya, signed in Geneva on 23 October 2020, contains a clause in which the parties commit to combatting hate speech, with a particular focus on social media. In article 5 of Section II of the agreement, the parties made a commitment to:

"Halt the currently rampant media escalation and hate speech by audio-visual broadcasting channels and websites. The judicial and competent authorities shall be called upon to take the necessary measures to ensure serious and deterrent prosecution of these channels and websites. [The UN Support Mission in Libya] also calls for necessary measures to be taken to ensure that the administrations of social media applications shall take the necessary action regarding these platforms. To this end, the Joint Military Committee (JMC) decided to establish a sub-committee to follow up on hate speech and pursue the necessary actions. The JMC also decided to address a direct message to all audio-visual broadcasting channels not to broadcast any media material that includes such type of rhetoric."

Social media code of conduct to prevent conflict during local elections in Indonesia

In Indonesia the Jakarta-based organization Saraswati and the Centre for Humanitarian Dialogue (HD) facilitated a social media code of conduct around local elections in December 2020. The code covered a range of commitments, such as the need for parties to fact-check their content, use authentic social media accounts, and disseminate accurate information. The code outlined the responsibilities of political parties, social media platforms, media companies and civil society by setting out an objective set of standards of online behaviour during local elections. A coalition of 12 civil society organizations drafted and publicized the code of conduct, and the Electoral Commission and representatives from Facebook attended its launch. A private social media management firm was commissioned to monitor the implementation of the code of conduct in two key constituencies during the final stages of the campaign.

4 Addressing mis- and disinformation on social media

Why is this important?

Various stakeholders in conflicts – be they conflict parties, local actors opposed to a settlement or external actors – increasingly use mis- and disinformation to achieve their aims. As a result, negotiators, mediators, advisors and, indeed, entire peace processes are increasingly being attacked on social media. This is especially the case during major political events such as elections and high-profile rounds of peace negotiations. The issues of whether and how to counter these attacks have become serious challenges in peace processes. If left unaddressed, the spreading of mis- and disinformation on social media can undermine mediation efforts by deepening conflict lines, delegitimizing peace talks, and undermining the mediator as a person and the reputation of the organization she or he represents. The fact that many conflicts are internationalized creates additional problems – for example, when external interveners engage in social media disinformation campaigns aimed at supporting or undermining a particular conflict party.

What do we know about it?

"Misinformation" involves the spreading of false information without necessarily malicious intent, while "disinformation" represents a deliberate attempt to disseminate harmful rumours and false information, and sometimes even hate speech. Mis- and disinformation can be general or used to target a specific individual or group. It can be spread through a variety of channels, but increasingly via social media. The UN has recognized countering disinformation and hate speech on social media as a major human rights challenge, as indicated in the UN Secretary-General's Strategy and Plan of Action on Hate Speech,[8] as well as in thematic reports of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression.[9]

When in doubt as to whether information relayed via social media is accurate, mediators and advisors can rely on experts to verify the

---

8   United Nations Strategy and Plan of Action on Hate Speech, guidance report, September 2020.

9   See, for example, a thematic report on online hate speech released in October 2019.

information. To this end, they can, for example, consult relevant lists on Twitter, seek the advice of well-informed civil society organizations, including women-led groups and investigative journalists, or use specific fact-checking applications and websites. They can also use the services of private companies that specialize in analysing social media and authenticating information.

Disinformation poses specific difficulties for mediators and their teams, especially when the source cannot be verified. Setting the record straight can be useful but can also give oxygen to an artificial controversy and divert from a mediator's principal mandate. If the source of disinformation becomes known to the mediation team, the mediator should weigh the advantages and disadvantages of revealing the identity – either publicly or privately – of the perpetrating party. This is one of the factors that the mediator should take into account when designing a response, which can be implemented either online, or offline, or both.

When deciding on an appropriate response, it is important to know whether the source of mis- and disinformation is authentic (for example, when civil society groups launch a coordinated social media campaign against peace talks) or whether it is inauthentic (for example, when the content is generated by external third parties using human or auto-generated bots or other tactics). Mis- and disinformation can be countered through a hybrid approach, using open or closed social media channels in addition to outreach with traditional media, for example through statements, communiqués and interviews. The lower a population's access to the internet, the stronger the reliance on traditional media.

A peace process can benefit from the removal of social media posts that directly or indirectly incite violence, as per the terms of service of the respective social media platforms and – importantly – domestic legislation. For content falling short of inciting violence, policies governing content removal and moderation vary across platforms. Whatever the case, the willingness of social media platforms to remove conflict-fuelling content has often fallen short, particularly in countries where many conflicts that require mediation currently take place. In these contexts, social media companies are often much too slow or unresponsive in removing harmful content, which is left in place to fuel conflict and undermine peace talks.

As an initial remedial step, some social media companies have set up "trusted partners" networks that are supposed to respond more quickly to these kinds of threats. After much pressure, companies have also tried to establish some form of external oversight of the content of their platforms, such as Facebook's Oversight Board. Because social media is playing an increasingly central role in contemporary conflicts, much more action by social media companies is needed to ensure that the platforms they run do not fuel conflicts and obstruct efforts to end them. There is also potential for mediation organizations to develop and maintain channels of communication to social media companies to facilitate addressing harmful social media content. Some collaborations of this kind have been established, but they remain ad hoc and insufficient.

Practical suggestions for mediators and their teams:

→ Partner with specialized NGOs to monitor social media in order to have better information, be in a position to detect cases of mis- and disinformation early, and potentially act to counter or respond to such information.

→ Encourage conflict parties to refrain from spreading harmful social media content.

→ Partner with local and national organizations that can help to identify and flag harmful content and encourage them to report it or take other types of action against it.

→ Set up a system to verify the accuracy of information relayed via social media or other channels. In complex cases, consider partnering with private companies specializing in this area, as well as with credible local news sources or institutions.

→ Set up a risk-management checklist to inform decisions on whether to ignore or react to mis- or disinformation. Relevant factors to consider can include whether there is a reasonable chance that problematic content will be rectified or removed, the scale of the threat to the peace process, and whether the source is known.

→ When the source of mis- and disinformation is known, contact those posting this material, either directly or via proxies, and ask them to rectify or stop disseminating the information. When the source is local, rely on local partners and influencers; if it is international, activate diplomatic networks, for example through a "Group of Friends" or other ad hoc grouping.

→ Provide training for local journalists and bloggers to raise awareness about mis- and disinformation.

→ Inform social media platforms about problematic content, using their terms of service as a guide. Identify whether the social media platforms have a country or regional representative and establish contact with this representative. Participate in a relevant "trusted partners" network to speed up response time.

→ If necessary, raise disinformation issues, including gender-based hate speech, with national authorities on the basis of relevant national legislation, while being mindful of human rights implications. Take steps to protect interlocutors who may be targets of disinformation.

→ Social media companies have a responsibility to appropriately manage the content of their platforms and should be more proactive in removing content that fuels conflict and promotes disinformation. Organizations with a strong mediation mandate, particularly the UN, should engage with these companies to encourage more proactive responses on their part. For example, mediators and their teams can help identify problematic terms and content associated with a particular conflict.

→ Beyond peace mediation, establish "rules of the road" for how UN entities engage with social media companies in order to ensure a coherent and norms-based approach to dealing with these actors.

Countering disinformation in the UN-led peace negotiations in Yemen

The Office of the Special Envoy of the Secretary-General for Yemen (OSESGY) has often been the target of individual and concerted attacks on social media and pays special attention to disinformation campaigns. Using human and automated scanning methods, OSESGY is able to detect and analyse social media disinformation campaigns targeting the Special Envoy or the peace process as a whole. This allows the Office to weigh response options with a view to reducing the harm such campaigns can cause to the Special Envoy's mediation mandate. OSESGY also adopts a mixed approach to fact-checking and verifying information, combining social media news cross-referencing with inquiries among journalists and well-informed contacts. In cases of disinformation, the Office posts corrections, as appropriate, or works with trusted Yemeni interlocutors to set the record straight when false information about the UN-led peace process circulates on social media or when the Special Envoy is the target of baseless attacks.

Hybrid offline-online approach to countering disinformation against the UN in CAR

In the Central African Republic (CAR) there are frequent social media attacks on the peacekeeping mission in the country – the UN Multidimensional Integrated Stabilization Mission in the Central African Republic (MINUSCA) – which also provides good offices and supports the country's peace process. A 2020 UN Panel of Experts report highlighted that MINUSCA staff had been subjected to a "well-orchestrated, targeted campaign" involving "the use of a fictitious local NGO and several social media accounts" to level accusations of misconduct against the mission staff, including of trafficking weapons to certain armed groups. This disinformation campaign led to a request by the CAR foreign minister to relocate four of the mission's staff members. MINUSCA adopted a hybrid offline-online approach in a response that was specifically tailored to a context with limited internet access. In addition to defending its staff, the mission used its own social media platforms and radio station to set the record straight and issue press releases addressing local warring parties in order to deflect accusations of partiality. It also used targeted, mass text messaging to address communities that did not have access to internet or social media platforms and engaged in high-level diplomatic discussions with the host government to resolve the issue.

Collaboration with Facebook in the context of the UN-facilitated Libyan political dialogue

Ahead of the launch of the UN-facilitated Libyan Political Dialogue Forum (LPDF) in 2020, the UN Support Mission in Libya established a partnership with Facebook known as the "Trusted Partner" service. This allowed UNSMIL to address hate speech, incitement to violence, and mis- and disinformation. As a result of UNSMIL reports, Facebook removed dozens of harmful social media posts that attacked activists, youth and peace promoters. During the course of the LPDF negotiations, UNSMIL and Facebook tightened the monitoring of dangerous narratives aimed at damaging the reputation of the LPDF process through fake news and misinformation about the process and about LPDF members, especially women. Thanks to this collaboration, many posts and pages promoting hate speech and attacking the UN-led process were removed. At the same time, the UN and LPDF members used social media to call for reconciliation and dialogue. This proved to be crucial for the success of the process and the protection of LPDF members and peacemakers, especially women.

Verification of envoy's social media account

Since 2018 the DPPA Mediation Support Unit (MSU) has been working to establish anticipatory relations with social media companies in order to be able to communicate quickly and effectively with appropriate company representatives should the need arise. Amid an already volatile situation in early 2021, the work of the UN Special Envoy on Myanmar was being complicated by disinformation regarding claims that she was about to arrive in the country. One such false report resulted in crowds gathering outside a UN office in Myanmar. The Envoy's Office asked MSU for assistance in getting her account verified by Twitter in order to establish an authoritative source of information about her statements and movements. MSU was able to make use of established relationships to pass on information regarding the urgency of the situation to Twitter, and the Envoy's account was quickly verified.

Combatting hate speech during a key moment of a peace process

In November 2019, a group of leading Libyan journalists, bloggers and social media influencers attended a workshop on "Combating Hate Speech and Ethics of Journalism" organized by UNSMIL and held in Cairo. Participants, including eight women, who represented various Libyan media outlets that broadcast from both inside and outside Libya reached a set of recommendations to combat hate speech during crucial moments of the peace process:

1. Refrain from posting or promoting disinformation or fake news that fuel hate speech on social media;
2. Respect intellectual property rights;
3. Respect privacy of others;
4. Accept differences, respect diversity and others' points of view;
5. Raise awareness on the importance of enacting legislation on negative consequences of hate speech in traditional and social media platforms;
6. Respect human values and principles in all social media content;
7. Urge Libyan media to promote an inclusive national narrative;
8. Join hands and network with UNSMIL, local and international organizations licensed by the State of Libya to counter hate speech;
9. Uphold professional standards, especially objectivity, accuracy, credibility and impartiality;
10. Protect sources [of information];
11. Establish and empower press trade unions;
12. Establish a mechanism to monitor and track hate speech in traditional media and media;
13. Urge civil society organizations and social figures/influencers to address hate speech;
14. Urge the competent local and international organizations to provide psychosocial support to victims of hate speech;
15. Provision of support to media personnel by the State of Libya.

Outlook

The social media landscape varies significantly from context to context, and behaviour on the various platforms continues to evolve. Mediators and their teams need to be aware of the risks that conflict parties' social media use can pose to a mediation effort, while at the same time remaining proactive in their response. The most important factor is to incorporate a "Do No Harm" lens and risk-management approach to all social media engagement. This means considering the effects – both direct and indirect – of social media use on interlocutors, negotiators and team members, and on the mediation process as a whole. Also critical is mediators' knowledge of the environment in which they are working, which makes an analysis of the digital ecosystem in a conflict-affected context indispensable. Finally, for the UN or any other mediation organization, a collaborative approach that includes engaging in particular with social media companies is crucial to leverage the positive potential but also mitigate the negative impact of social media use in peace mediation.

This report is part of a broader reflection on the role of digital technology in conflict prevention, mediation and peacebuilding. Critical to all of these engagements is a realization that while mediation remains a fundamentally human-led endeavour, digital technologies – including social media – are playing an increasingly important role in peace processes, and mediators and their teams need to be prepared to deal with any related issues that arise. Another common element is the need for additional research to understand social media dynamics in conflict-affected contexts and the associated implications for mediation efforts, and to refine and test practical suggestions in this regard. To substantiate the issues covered in this report, research into how social media influences dynamics in conflict theatres and at the negotiation table would be particularly valuable.

Published material on social media and peace mediation

DPPA (United Nations Department of Political and Peacebuilding Affairs) and Centre for Humanitarian Dialogue, Digital Technologies and Mediation in Armed Conflict, <u>Digital Toolkit report</u>, March 2019.

DPPA, Social Media Analysis, <u>manual</u>, May 2020.

Harlander, Jonathan and Maude Morrison, Social Media Codes of Conduct: Reflections for Mediators, <u>online article</u>, Centre for Humanitarian Dialogue, August 2020.

Hattotuwa, Sanjana, The Janus Effect: Social Media in Peace Mediation, <u>online article</u>, ICT4Peace Foundation, July 2018.

Hirblinger, Andreas, Digital Inclusion in Mediated Peace Processes: How Technology Can Enhance Participation, <u>Peaceworks Report no. 168</u>, United States Institute of Peace, September 2020.

Hirblinger, Andreas, Maude Morrison and Helena Larrauri, "Digital Analysis: Peacemaking Potential and Promise", <u>chapter</u> in Accord 29, Conciliation Resources, September 2020.

Jenny, Joëlle, Rosi Greenberg, Vincent Lowney and Guy Banim, Peacemaking and New Technologies: Dilemmas & Options for Mediators, <u>Mediation Practice Series report</u>, Centre for Humanitarian Dialogue, December 2018.

Kavanagh, Camino, Digital Technologies and Civil Conflicts: Insights for Peacemakers, <u>Conflict Series Brief no. 4</u>, European Union Institute for Security Studies, February 2021.

Lanz, David and Ahmed Eleiba, The Good, the Bad and the Ugly: Social Media and Peace Mediation, <u>Policy Brief 12/2018</u>, swisspeace, December 2018.

Majcin, Juraj, "Social Media Challenges to Peace-making and What Can Be Done about Them", <u>article</u>, Groningen Journal of International Law, vol. 6, no. 2, 2018.

Mandelik, Florence and Ayat Mohamed, Can You Hear Me? Taking Mediation Online: Reflections from Practice, <u>NOREF Brief</u>, NOREF Norwegian Centre for Conflict Resolution, September 2020.

El Mawla, Mira and Krystel Tabet, Take Back the Analysis: Five Things You Can Actually Learn about a Conflict Context from Social Media, <u>online article</u>, Build Up, February 2021.

Morrison, Maude and Adam Cooper, In War Zones, Social Media Disinformation Is Costing Lives, <u>online article</u>, Wired, April 2021.

Wittig, Katrin and Sausan Ghosheh, "The Role of Social Media in Early Peacemaking: Help or Hindrance?", <u>chapter</u> in Accord 29, Conciliation Resources, September 2020.