MEDIATION   **December 2018**
PRACTICE SERIES

8

Peacemaking and new technologies

Dilemmas & options for mediators

Joëlle Jenny, Rosi Greenberg, Vincent Lowney and Guy Banim

Edited by Jonathan Harlander

"hd | Centre for
Humanitarian
Dialogue

Mediation for peace

The Mediation Practice Series is a project of the Centre for Humanitarian Dialogue (HD). We value the feedback of mediation practitioners and researchers on the format and content of this publication. Please write to mediationsupport@hdcentre.org for suggestions and feedback.

Foreword

The Mediation Practice Series

The Mediation Practice Series (MPS) was initiated in 2008 as part of the Centre for Humanitarian Dialogue's (HD) efforts to support the broader mediation community. The series draws on feedback from mediators, including HD practitioners, who tell us they and their teams often lack adequate insight into other peace processes. In the past few years, the international community has significantly strengthened the support available to mediators and their teams. HD is committed to contributing to this effort and to the improvement of mediation practice.

Based on the shared view that mediators often confront similar dilemmas, although mediation differs widely across peace processes, HD is producing a series of decision-making tools that draw on the comparative experience of mediation processes. Each publication in the series will give readers a concise overview of relevant challenges and options, and help them prepare for the potential demands of mediation processes.

Although these publications cannot replace practical experience, it is our hope that they can contribute to a more systematic learning process. The forthcoming publications in this series will be made freely available on HD's website and will be disseminated through our network and those of our partners. *Peacemaking and new technologies* is the eighth publication in this series.

The Mediation Practice Series

Contents

Essential points for practitioners

- Regardless of the development and use of information and communication technologies (ICT), the practice of mediation still relies first and foremost on the trust built between a mediator and conflict parties, and the ability to generate and maintain buy-in to peace processes.

- Mediation teams have a responsibility to be generally literate about the technologies present in the mediation environment and their effect on the mediation process, and to make informed choices about their use. At the very least, mediators need to understand the risks associated with these technologies, and how to mitigate them.

- The use of ICT should never be assumed to be fully secure. Several mediators report that accepting the threat of information disclosure through network monitoring by governments is often the only realistic approach.

- It is critical that mediators apply all basic cyber-hygiene rules, including installing robust anti-virus programmes, applying updates and security patches as soon as they are released by the vendors, and not opening suspicious attachments or hyperlinks.

- If they want to fully eliminate threats, mediation team members must not carry or communicate with digitally-enabled devices.

- ICT leaves all parties involved in the conflict vulnerable to hacking, the loss of confidentiality, and the capability for rapid dissemination of confidential information, potentially for nefarious goals.

- The interconnectedness of people through digital media and technology has enabled the weaponisation of information and disinformation on an unprecedented scale and in a way that would not have been possible in the past.

- At the same time, online space provides new channels through which mediators can monitor rapidly evolving conflict trends, and can interact with, and shape the narratives of, conflict parties.

Peacemaking and new technologies

Dilemmas & options for mediators

1 | Introduction

Rapid advances in information and communication technology (ICT) are having a profound impact on political, economic and social life. The development of new technologies facilitating communication and information exchange – including instant messaging, social media and data analytics software – has been accompanied by a dramatic increase in smartphone ownership and usage. As of January 2018, more than 4 billion people were connected to the internet and more than 3.2 billion (approximately half of the world's population) were using social media.[1] Many parties to conflict and local populations use these technologies, as do many mediators.

Mediation teams are consumers of new ICT and need to choose which, if any, ICT platforms or applications to use. Mediators consequently have a responsibility to understand how these technologies work and the implications of their use. The choices they make are not neutral in terms of balancing political risk and the physical safety of both conflict parties and peacemakers.

New technologies can disrupt conflict dynamics and peace processes. There are two issues around the use of ICT which are of particular concern to the practice of mediation:

- Mediation has traditionally been a communication process largely based on face-to-face interaction and in-person trust-building. New technologies can close some information and cognitive gaps

but they can also introduce new cognitive biases and 'noise' which can potentially challenge assumptions and methods used in mediation processes.

- Mediators use their own analysis to offer conflict parties a structure within which they can 'agree to disagree' about what is true, and where shared narratives can be developed which will plausibly resonate with the wider public. This analysis was previously developed from internal reporting (for example, within a foreign ministry or multilateral body such as the UN and EU), professional journalism and expert publications as well as first-hand information from direct contacts. The proliferation of new sources of information in an increasingly interconnected world changes this analytical landscape.

What does this 'digital revolution' mean for mediators? What does a mediator need to know about digital security? Most mediators understand that their use of a particular technology may have positive and negative impacts on the structure and security of a peace process. However, mediators are not sure where to turn for advice on enhancing digital security and addressing phenomena such as fake news. While there is a growing body of literature on the use of technology in humanitarian aid[2] and development[3], as well as in online dispute resolution[4], little has yet been written about the effects of technology on peace mediation and its potential to aid in the prevention and resolution of violent conflict and in peacebuilding efforts.

This Mediation Practice Series publication will help mediators frame discussions about the impact of innovations in ICT on conflict and mediation. It highlights some of the opportunities and risks around using online sources to analyse a conflict. It also offers ethical guidelines and a threat assessment framework to explore the risks and benefits of ICT applications which may be used to communicate privately and with the wider public.

This publication is based on a review of the existing literature on this issue; interviews carried out in October, November and December

2017; and informal exchanges with mediators, academics and experts. To avoid any potential negative impact on ongoing peace processes or the reputation of interviewees, most quotes in this publication are anonymous.

Mediators cannot simply follow the ICT choices made by the conflict parties. At the very least, mediators need to understand the risks associated with these technologies, and how to mitigate them. There is no simple answer to the question of which ICT tools to use. Further research and discussion among practitioners is needed to help mediators navigate the ICT landscape and keep up with innovations, particularly in relation to surveillance, encryption, disinformation and the analysis of so-called 'big data'.[5]

2 | ICT, conflict and peacemaking

Our definition of ICT includes all technical means used to handle information and facilitate communication. This includes both computer and network hardware, as well as software.[6] A subset of the new ICT is social media, for which we use the Merriam Webster definition: "forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos)."[7]

Potential impact of ICT on the dynamics of warfare

ICT has "radically changed the quantity and quality of information available to individuals, groups and governments (and the way they transmit it); that means much of what we know about civil war dynamics will also change."[8] Through the rapid increase in the availability of technological tools, messages crafted by parties with a stake in the conflict, and their associated broader narratives, reach more people and spread faster than ever before.[9] This creates the potential for greater intercultural exchange and peacebuilding, a better informed citizenry, and quicker access to information allowing for more accurate mapping of peace and conflict dynamics. A

number of public and private sector initiatives, using social media handles such as #peacetech and #ICT4peace, have sought to develop and harness this potential.[10]

However, the new ICT is also a tool for those propagating hate speech. Rumours that are spread quickly over social media can create and fuel animosity.[11] Beyond this, cyber-attacks have become common, turning digital space into a tool of war itself.[12] Some have even argued that the increase in armed conflict since 2010 is correlated with the digital revolution.[13] It is suggested that ICT has brought a shift in the balance of power away from governing elites, a reduction in the barriers for establishing rebel groups, and an increase in impact once these rebel groups are established. ICT offers a greater incentive to frame goals in global and extremist terms; easier access to financing; and the potential for the more rapid spread of ideas.

Using disinformation and propaganda as a tool of political conflict is nothing new. But the rise of social media has increased the global reach of such operations. The conduct of disruptive operations – which are below the threshold of war but aim to destabilise another country or group – has been greatly amplified through the proliferation of new technology. Social media is a perfect medium for manipulating a group's feelings, fears and ideas about their own collective identity. The interconnectedness of people through digital media and technology has enabled the weaponisation of information and disinformation on an unprecedented scale and in a way that would not have been possible in the past.

Whether the technology should be understood as a force for peace, a root cause of conflict, or only as a multiplier, catalyser or accelerator, any mediator will need to consider these dynamics.

Mediators must be prepared to operate in an environment where the possibilities of misinformation and disinformation have profoundly changed.[14] The recent investigation into Russian interference in elections in the United States and in Europe, and allegations that

similar techniques were used in several political scandals in Africa[15], has shown how social media such as Twitter and Facebook can be manipulated by the creation of botnets (groups of automated accounts that repeat posts to increase visibility on Twitter) and fake groups spreading disinformation. This type of information operation will become more common, especially as tools for fabricating video and photographic evidence are spreading.[16]

Ability of ICT to polarise and divide

The business model of the tech giants that has shaped the internet to date has become known as the 'attention economy'. Content is generally free for the public because profit is generated by advertising revenues and by collecting online behavioural data on those visiting the website. For this to work, people are enticed to stay as long as possible on the web pages they visit: the easiest, and hence most profitable, way to do this has been to harness the insights of behavioural science in order to design apps and content that appeal to emotion – in particular fear, anger and outrage.[17] One interviewee noted: "it is a struggle to find even-handed reporting. Before, everyone relied on the same controlled sources but now people are following certain streams that reinforce views and do not give a view of the other side."

The conflict-multiplying potential of platforms that are encouraging citizens to become more impulsive in the way they reason and communicate has barely begun to be considered, let alone understood. Increased access to technological tools by elites and media outlets had been assumed to increase their capability to influence yet may be doing the opposite.[18] As one interviewee put it, "ICT sometimes makes people lose touch with reality. People create something online and recipients are not able to analyse it. They just believe one person and it can create more tension in society as a whole." In addition, armed groups now use new technologies to recruit as well as to incite and enact violence. For example, Daesh used social media to promote its destruction of cultural sites and to recruit new fighters.[19]

Has social media driven violence and conflict in Myanmar?

Myanmar, a country of 53 million people, has experienced a dramatic transition in information access. Ten years ago journalism was heavily controlled and state-censored, and a SIM card cost the equivalent of €2,500. As recently as 2014, less than one percent of the total population had internet access. Today there are 30 million registered Facebook accounts.[20] Many users see the platform as a primary means for accessing and sharing information. Within this context, and with posts frequently mistaken for news, there has been a proliferation of harmful misinformation and hate speech in Myanmar. This has had a significant effect: accusations have been made that Facebook contributed to the Rakhine crisis of August/September 2017, which saw more than 600,000 Rohingya fleeing from Myanmar to Bangladesh.[21] Specifically, it appears to have made possible a rapid polarisation between communities.

In investigating the crisis, the Chair of the UN Independent Fact-Finding Mission on Myanmar said that social media had "substantively contributed to the level of acrimony and dissension and conflict, if you will, within the public. Hate speech is, of course, certainly part of that."[22]

In an effective use of new technologies to stimulate debate on the issue, Myanmar author, historian and former UN official Thant Myint-U tweeted:

@thantmyintu 5 April "I wonder if he [Zuckerberg] or anybody else at his company have thought at all about all the ways Facebook these past few years have shaped Myanmar's entire political landscape as well as the prospects for any future peace or democracy."

@thantmyintu 6 April "Wasn't referring only or even primarily to hate speech. Facebook as the only social media platform in Myanmar, where a telecoms revolution accompanied the first political liberalization in half a century, has profoundly shaped post-junta dynamics, w/different winners and losers."

@thantmyintu 8 April "Interesting to imagine Myanmar's political opening if say 10 years earlier, before social media; would there have been flowering of old media? A different kind of post-junta political discourse? More dissent on the street v. online? Different expectations, types of mobilization?"

The online Myanmar newspaper Irrawaddy (itself a product of internet freedom but also the victim of a number of deliberate hacks) has a sobering forecast but asks the right question, "there is no doubt that the next major conflict in Myanmar will start on social media or in cyberspace. But who will take the blame, and how can the problem be resolved?"[23]

Potential ways in which ICT impacts mediation methods

Researchers in the field of Online Dispute Resolution write, "what we usually fail to take into account, as we talk, text, phone, mobile phone, e-mail, instant message, Tweet, status change and video-conference our way through life, is that the messages we transmit and receive, and the way we transmit and receive messages, are affected – and sometimes fundamentally altered – by the media we have chosen to convey them through. Understanding the ways in which the medium affects the message – 'media effects' – is crucial."[24]

In addition, ICT disrupts traditional goals of mediation and can create new possibilities:

- **Mediation as an effort to close information gaps between parties**. According to game theory, war can be an outcome of a communication failure whereby information asymmetries lead actors to behave irrationally (i.e. violently). If tech optimists are correct, new ICT holds the promise of greatly reducing, or even eliminating, this information gap between parties. As a result, a shuttle diplomacy approach focused on passing factual information becomes less pertinent, as groups are less likely to be isolated and can easily access information online. On the other hand, for tech pessimists, the misinformation/disinformation crisis accompanying social media may enhance the value of the mediator role in establishing confidence in the veracity of information.

- **Mediation as an effort to forge a shared narrative between parties**. Narratives represent the ways parties understand situations and make sense of reality.[25] Accordingly, parties to a conflict often have widely diverging narratives with little common ground. In this case, the role of the mediator is to guide the parties toward a shared narrative. The new possibilities offered by ICT create the potential for narrative-shaping in mediation processes. The battle over narratives is taking place increasingly in cyberspace. It is where mediators, conflict parties and, indeed, citizens make sense of what is happening, who is responsible and what should be done about it. Efforts by mediation teams to help forge shared narratives will need to engage with debates driven on social media.

- **Mediation as an activity which generates solutions through the transfer of knowledge**. Many mediators offer a third-party perspective on the conflict to stakeholders as a way of opening up new possibilities for thinking about a way forward. Online databases of peace agreements, interactive tools and open-access information that can be tailored and packaged by intelligent search engines may mean that this 'peacemaking professional' model of a mediator will lose its pertinence. It is, however, likely that mediators will continue to play an essential role as knowledge interpreters, able to contextualise lessons from other peacemaking efforts.

- **Mediation as a dialogue or problem-solving process**. The archetypal model of a mediation as two political leaders from opposing groups sitting across a table with the mediator in the middle has never quite matched the complicated reality of multi-level and multi-track dialogue and consensus-building processes. It is likely to be even less reflective of reality in a future where power is diffused and ICT offers new platforms for engagement.[26] While initial optimism that online engagement would have an inherently positive impact on the quality of dialogue has proved to be misplaced, online platforms nevertheless offer significant potential for engaging citizens in brainstorming and problem-solving. Attention is now focused on how ICT can be engineered in a way that promotes constructive consensus-building over destructive polarisation. The application of mediation expertise in this area may have as much, or more, peacebuilding potential than focusing on the work of senior envoys.

- **Mediation as an effort to reframe problems and satisfy underlying psychological needs**. One of the (controversial) insights of marketing based on big data is that people may be susceptible to messages due to psychological reflexes and that these reflexes can be exploited by new technologies. Paying attention to the psychological dimension of interactions is not new to the mediation field. It is often discussed in terms of the 'art' of peacemaking with its focus on setting (table layout, food and drink, creating venues for informal encounters), body language (the symbolic handshake) and personal chemistry (charm, asking

about family, gender dynamics). A mediator has an obvious interest in any technology that could provide detailed psychographic profiles of individual conflict protagonists. New techniques for tapping into underlying psychological needs in order to change group behaviour have been made available through the analysis of big data. Algorithms may be used to privilege win/win and conciliatory outcomes in social media exchanges and 'peace bots' might accentuate conciliatory, and diminish polarising, messaging.[27] Peace mediators now have new possibilities for operating at a societal level to shift narratives as well as reformulate problems to address underlying interests and less rational (and more emotional) needs. How insights into human behaviour which have been deployed on social media platforms for commercial, and political, marketing purposes may be deployed in the field of conflict resolution remains underexplored. At the least, mediators need to consider how new technology may be used for positive or negative manipulation of the social contract and its ethical implications.

Given the uncertainties around the impact of ICT on conflict dynamics and mediation processes, there are no simple recommendations for how mediation teams should navigate this new environment most effectively. The mediation community needs to debate these issues and develop a shared language that captures the issues at stake. Simply continuing to operate with the tools and approaches of the pre-social media era means missing out on the opportunities that new technologies offer and mediators risk becoming instrumentalised in processes they do not understand. The contours of this debate and some potential ways forward are offered in the following sections.

Use of digital sources for analysis

Online conversations have become increasingly important to mediators for understanding the context and monitoring opinion about ongoing processes as parties to conflict compete for public opinion in the online space.

This means social media tools such as Twitter and Facebook are increasingly relevant to conflict analysis. Nearly every mediator interviewed for this publication used social media to understand local attitudes towards conflict, the positions of the conflict parties, and to get real-time updates on current events. Most mediators who reported using Twitter and Facebook indicated they only did so as consumers of media rather than as producers, although some have used them for identity verification.

Social media can help mediation teams during the preparation of mediation processes as well as during them. One mediator recounts "in Bosnia our graphs showed that both communities cared about education for kids, which wasn't being done well through quality or messaging to their kids. Through this tool we found that both sides had a problem with that, so we used it as a starting point. But these tools are only as good as questions you asked and data you generated."

Even with a relatively passive use of social media, mediation teams are able to glean information about where tensions are high, what the population needs, and the key actors in the conflict.

As digital identities become more important these will provide new possibilities for data analytics and enable mediators to achieve more up-to-date awareness of the situation than previously possible. By following a broad range of people on Twitter a mediator can quickly assess opinions and perspectives on a situation. One mediator says, "I'm a consumer of Twitter. If one curates lists on Twitter, see when news breaks from groups, hostage videos, etc., one starts to get a sense of which individuals have a basic level of access and knowledge."

Social media is increasingly helping mediators pre-empt conflicts before they turn to violence. Even with a relatively passive use of social media, mediation teams are able to glean information about where tensions are high, what the population needs, and the key actors in the conflict. More active use of specific applications and

technologies can provide greater depth, granularity and reliability. There is an increasing range of technologies to support this analysis such as early warning tools, situational awareness tools, geo-localisation (GPS), crisis mapping carried out by crowdsourcing, and web scraping applications that aggregate across multiple webpages. Crowdsourcing and early warning software can provide real-time crisis analysis: for example, Ushahidi is a non-profit tech platform that enables organisations to crowdsource to-the-minute data on conflict environments. Local populations can contribute reports through SMS, email, Twitter and the Ushahidi web application. Visualisation tools and custom dashboards enable users to receive up-to-date reports on a situation.

One mediator mentioned using the Ushahidi platform in Kenya to monitor ongoing ceasefire agreements through crowdsourced reporting, to communicate with local populations about violations in certain areas, and to warn against travel.

Early warning software such as Ethnographic Edge, Recorded Future and Secureaxis combine publicly available data from numerous sources to draw conclusions about trends in conflict and safety. Monitoring food prices and the tone of global news media in order to compare trends can provide insight into possible political and economic upheavals.[28] Private companies offer surveys, media monitoring, network mapping, and data visualisations to help organisations understand the environment in which they work. Mediation teams may be interested in using these technologies for early warning and/or crisis mapping to increase their awareness of both their own safety and the conflict environment.

One mediator reported using Livemaps and Geographic Information Systems (GIS) to map shifting spheres of influence, discuss areas of access in real-time, and track developments on moving frontlines. Another mediator used satellite imagery for monitoring potential violations of an ongoing agreement. However the majority of mediators interviewed for this publication made relatively little use of data analytics platforms.

These technologies have the potential to enhance the ability of mediators to identify changes in public opinion, verify information from interlocutors, monitor ceasefires and other elements of agreements, and even identify potential crises before they arise.

Mediation teams can monitor multiple social media platforms at once across multiple accounts, identifying trends in particular hashtags, mapping networks to identify key influencers and group composition, and monitoring the use of certain keywords in real-time. Such tools may also help mediators identify shared interests across lines of conflict. This, however, relies on the local population having access to technology.

Risks

Technology is not ubiquitous. One of the concerns of mediation teams is the possibility that relying on social media and big data analytics may exacerbate inequalities in relation to access to technology, and thus disempower certain groups. One of the interviewees highlighted, "one risk is that there's an over-focus on those who are connected as opposed to those who are off-grid or who have chosen to be off-grid because the conflict is not enabling connectivity. So it's only a partial picture and cannot be seen as a full read of the environment."

What are bots and why should mediators care about them?

A bot is a software application that performs automated tasks over the internet. Google assistant, the Siri application in an iPhone and the chatbots now used by many airlines to field initial customer complaints are examples of bots.

It has been estimated that two-thirds of tweeted links to popular websites are posted by automated accounts – not human beings.[29]

The @PeaceTechBot available on Twitter is one example. It is a Twitterbot that 'crawls' the Twittersphere automatically retweeting material related to conflict data and peacebuilding grants.

Whilst not inherently good or inherently bad, the power of bots to aggregate information and internet-based news content can be deployed with malicious, as well as benevolent, intent. They can also have unintended consequences. A feature of the current internet landscape is that it is not easy to determine who has created a bot and for what purpose. Self-reproducing bot-farms may be the product of AI, a state-backed conspiracy or a bored teenager. There are three primary distorting effects mediation teams need to be aware of:

- **Artificial amplification of a voice or strength of popular sentiment**. Bots can be a cheap and easy way to boost the appearance of the strength of support for an individual politician or a particular view point. A citizen journalism website in Sri Lanka, for example, has investigated this possibility and highlighted the situation in Sri Lanka where a "new appetite for social media strategies specifically engineered for electoral gain amongst all politicians... involving human trolls as well as automated bots. The intent is clear – to influence voter perceptions and public discourse, over and beyond social media."[30]

- **Deliberate manipulation of a discourse**. Bots are an effective way to multiply content in an effort to shape a debate. For example, in the context of the online conversation about the White Helmets in Syria it is alleged that social media algorithms have been gamed through "a flood of content, boosted by bots, sock puppet accounts and a network of agitators, to create a "manufactured consensus" that gives legitimacy to fringe views."[31]

- **Automatic posting of inflammatory or erroneous material**. Bots that tweet or retweet inflammatory or erroneous material may not even have been programmed with malign intent but they generate additional 'noise' and potentially incite violence.

Botnets are formed by linking individual devices that are running bots. These kinds of networks are behind some of the malicious attacks on ICT networks, the generation of spam and the dissemination of viruses, but may also play a role in the types of distortions listed above.

Mediation teams that have traditionally been skilled at carrying out political analysis and conflict assessments in polarised and fragmented societies may, in the future, need to be able to counter additional uncertainties and risks to the process that are generated by bots.

A number of techniques are used by parties to manipulate online conversations in ways that risk skewing the analysis of conflict mediators. Astroturfing, or using fake accounts to create the appearance of popular sentiment where it doesn't exist, is one such example. In the words of one of the mediators interviewed for this publication, one does not really know if influential voices on social media are "thirteen-year-old kids in their parents' basement or a top commander."

The fast pace of communications causes misinformation and disinformation to spread rapidly without proper fact-checking. Social media can quickly distort the conversation. In addition, biases in algorithms and search terms can have a significant negative impact on the accuracy of the analysis. Without knowing how these algorithms operate, those drawing conclusions from search results are likely to make errors in their understanding. Interpreting big data is, in itself, a complex task. As Emmanuel Letouzé, Patrick Meier and Patrick Vinck write, "an inconvenient truth is that big data (and fine-grained measurement techniques) are terrific material for statistical shenanigans, biased fact-finding excursions that may lead to false discoveries, spurious correlations, confusion between correlation and causation, and more econometric and logical ills."[32]

Even when used properly, analysing and acting on the gathered data remains a challenge.[33] ICRC Director of Communication and Information Management Charlotte Lindsey-Curtet says, "we are skilled at building likely scenarios but not the worst-case scenarios. Are we taking the signals for really bad events or are we hoping things will happen to minimise the risk? We might be looking afterwards, saying signals were there, but no one wants to interpret them." Other drawbacks include 'noise'. As Lindsey-Curtet says, "in a context we were analysing we completely discounted all of the English-language communication. It was just noise but of such a significant volume that it would have distorted the analysis."

Financial constraints and the time needed to develop adequate analytical resources can be another impediment. Some of the

experts in the field of data analytics said there is currently no tool they would fully recommend for analysing social media due to the difficulty of programming them for specific needs.

Other considerations of growing importance are data privacy and data ownership. Recognising the impact of ICTs on these issues, the European Union has issued the General Data Protection Regulation (GDPR), which establishes a right of disposal of one's own private data.[34] The ICRC and other organisations have increasingly called for these rights to be adequately protected for populations affected by humanitarian crises. While mediators are less likely than aid workers to collect personal data, it is important they take data privacy into consideration in their work.

The Harvard Humanitarian Initiative has developed the "Signal Code", an ethical framework for the use of ICT by communities of practice in humanitarian emergencies.[35] Based on the Universal Declaration of Human Rights and other international agreements on human rights, privacy and technology, the Signal Code designates five basic rights with regard to how information is handled during crises:

- The right to information;
- The right to protection;
- The right to data privacy and security;
- The right to data agency; and
- The right to redress and rectification.

From this we suggest that mediators must have the necessary technical capacity and digital expertise to protect their interlocutors. Moreover, from the guidance on data agency, which includes a mandate for informed consent, we infer that mediators must ensure they are sufficiently able to inform parties about the use of information and the risks they are undertaking by communicating through technology. From a rights-based perspective, mediators have an ethical obligation to ensure that parties in conflict are aware of the risks they are taking by choosing a particular technology for communication during a mediation process.

Use of digital sources for analysis
Benefits and risks summary

Benefits	Risks
• Social media analysis and big data analytics can enable real-time context analysis and assist in preparedness.	• Use of social media analytics and big data analytics for crisis mapping may exacerbate pre-existing inequalities in access to technology and may not be fully representative.
• Social media analysis and big data analytics may enable mediation teams to identify growing areas of conflict at faster rates than previously possible.	• Social media can only provide a partial picture of public opinion, which may be more extreme than actual opinion.
• Social media analysis may help mediation teams identify networks, key influencers, and important locations, enabling them to target and focus their work.	• Social media may be manipulated by parties to influence the mediation process and analytics may not identify this.
• Social media may enable broader inclusion of opinions and populations in mediation processes.	• Bias in algorithms, search terms, and software may go unseen, offering a false sense of certainty.
• Social media analysis and big data analytics can help mediation teams monitor ongoing agreements.	• Predictive capacity is limited and depends on interpretation.

3 | Communicating with apps: what mediators need to know

Private communication: mediation team to parties

The ease and speed with which one can communicate via ICT has increased the capacity of mediation teams and the rate at which they can work. One mediator interviewed for this publication said, "with Skype and WhatsApp, you can do a face-to-face without 10,000 miles and 3 weeks. You would use Skype or WhatsApp for discussions, making arrangements and communications between formal meetings. Or to furnish information that has been requested in formal meetings."

In addition, mediators working on multiple conflicts or with multiple parties within a conflict can maintain communication with all parties at once. Most mediation teams using ICT emphasised the speed and ease of these forms of communication as the primary benefit.

Almost all mediators we interviewed reported using person-to-person messaging apps including WhatsApp, Telegram, Signal and Skype. Text messages and voice calls were the most popular methods used. Interviewees also mentioned other apps such as Line, Viber, and Blackberry Messenger as ones they had used in particular locations, depending on local popularity and the preference of conflict parties.

These technologies are typically used for checking in about logistics and setting up future in-person meetings, maintaining trust-building contact between meetings, and discussing small elements of content. These are also used for fact-checking and floating messages for approval.

Messaging apps help protect the security both of individuals and of networks. The use of ICT also increases the sophistication needed to eavesdrop on these processes, which may benefit mediation teams in less technologically sophisticated areas.

A mediator's story: the death of a trusted agent, a rebel leader, and the shift to WhatsApp

One mediator interviewed for this publication recounted how direct communication technology had increased safety for parties he dealt with directly. He used to rely on 'shepherds' – "trusted intermediaries who can go from key interlocutors of a group to you." These people had to be able to cross the divides within a country and their success was reliant on being trusted by all sides to a conflict. The mediator had used many of these shepherds throughout the years, and vividly remembered when one was killed while crossing a contested border on behalf of the mediator. "This was obviously hard to absorb," he said. Today, his use of trusted intermediaries has decreased by over ninety percent as he relies on WhatsApp to contact the conflict parties.

The need to meet face-to-face can be dangerous for the participants as well as the shepherds. In one case, a shepherd had to meet with a rebel leader, take possession of a letter, and then circuitously bring it to the government. Before the letter was delivered, the rebel leader who had written the letter was killed in a drone strike. It emerged that the shepherd, known and trusted by both sides, had been used by the government to locate and target the rebels.

This mediator now relies heavily on voice communication via WhatsApp. He believes the encrypted application increases the safety of the participants and removes the need to meet in person as often. There are challenges associated with using this tool, though: language differences may make verbal and text communication more opaque than communication in person. Confirming the identity of participants through technology can also be challenging. This can be circumvented by combining both old and new technology, for example using the unique serial number on a bank note to serve as a signifier that the courier delivering a mobile phone is trusted by the leadership of the organisation. Ultimately, this mediator believes that the use of ICTs has made mediation safer and more effective for the participants.

Several mediators mentioned ways that ICT helped with trust-building. In one situation, a mediator reported using technology that analyses facial expressions to determine if video chat interlocutors could be trusted, a practice that may present real benefits but also carries with it ethical dilemmas as surveillance technology continues to develop. One mediator mentioned that during video conversations with parties, the latter would often be at home with children in the background. This offered a view of the interlocutor in a completely different role and later enabled interpersonal connections in more formal meeting spaces, when the mediator asked about the man's children.[36] Frequent informal contact might also increase trust and co-operative behaviour.[37] Thus, technology might add a new dimension to communication which has the benefit of increasing trust when used in specific situations.

Mediators emphasised the need to have a personal relationship with the parties before shifting to digital technologies.

Direct communication technologies make it increasingly easy to reduce the level of formality in a process. A mediation team can have continuous informal contact with parties via messaging. This informal access can create a more personal bond between the mediation team and the parties.

Risks

Mediators emphasised the need to have a personal relationship with the parties before shifting to digital technologies: as described by one of the interviewees, "it's about sleeping at houses and getting to know families. Putting your life in their hands and getting to know them so they see you're not a bad bloke. It starts with a human interaction and direct relationship."

Research has shown that over-the-phone communications foster **less trust** and more competitiveness than in-person conversations.[38] Text-based communication is correlated with even higher levels of mistrust and contentious behaviour in entirely online negotiation

processes, as well as lower satisfaction with outcomes.[39] The loss of information from verbal and non-verbal cues, and the limited amount of context, can create a natural scepticism towards the motives of others. Asynchronous messaging and interruption of streams of thought through multiple simultaneous messages cause conversations over text applications to be more disrupted.[40]

Transmitting a message through digital means may also **change the non-verbal content** of the communication. In comparison with in-person communication, text-based communication over the internet is correlated with lower feelings of accountability for the message, a heightened sense of anonymity, and less care in composition of messages. Studies of email communications have shown that when writing, people are much more likely to focus on logical argument and task-orientation over personal experience-sharing.[41]

There is also a risk of **miscommunication**, particularly when leaving details out due to the sensitivity of exchanging information over insecure channels. One mediator says, "you could be having a conversation around 'friends' but you are speaking of a direct interlocutor, while the other side believes you are talking about an armed group. It's striking how often there is a misunderstanding." In addition, information provided via text may be limited and understanding may go unchecked due to a lack of confirmation or questioning as well as body language cues and responses that are available in person.[42]

Direct communication through ICT also carries the risk of **'spoofing'** or pretending to be someone else. Parties may spoof one another or the mediator, or factions within one party may spoof one another. As the composition of the group shifts over time, the points of contact for the mediator can become less relevant or less representative of the group. In addition, it is almost impossible to initiate contact with a new interlocutor through messaging apps. There is always a risk of not knowing who is on the other end of a communication.

The mediation team's use of a particular platform may also **affect perceptions** of their impartiality. For example, in situations where the state is a party to conflict and may be surveilling certain messaging applications in the local context, the mediation team's endorsement of surveilled vs encrypted applications will affect their perceived impartiality. This factor must be considered in the process. If, for example, a mediation team chooses to use a fully encrypted technology and the government is aware of this, they may be perceived as less trustworthy by the state. However, if they choose platforms which are unencrypted or are shared with third parties, parties who are adversaries to the government may lose trust.

Social media at the negotiating table

A mediator interviewed for this publication was involved in a negotiation between a central government and a province in rebellion. During a formal round of talks between the two parties, supervised by a mediation team, he noticed that the parties were not focused on what was happening in the room. Instead, all eyes were on their mobile phones, with a rapid sequence of duelling Tweets and other messages on social media going out to their various constituencies. These messages were about the content of the negotiation, a particularly contentious issue. While both parties had agreed that, in theory, the negotiation would be discreet and confidential, in practice it was seen as more effective for them to rally their populations to support their positions as opposed to negotiating with the other side.

The impact of this use of social media was similar to having a real-time press conference in the room with a host of uninvited, unaccountable voices involved in the negotiation. The mediator compared this to having agents in the room with no idea who they represented but who were influencing the process. The mediation team had to respond to the dialogue on social media because it was so influential on the negotiation representatives.

In an attempt to manage the impact of social media on the process, the mediators imposed a ban on mobile phone use during face-to-face negotiations. It was however impossible to limit the use of social media by the parties; the mediator recounting this story believed that the exploitation of social media by conflict parties during negotiations will be a major force in conflict resolution moving forward.

Impartiality may also be compromised by the information that is shared. Due to the fast-paced, less thoughtful nature of communication via messaging applications, technology may bring the mediation team in contact with ethically challenging information. For example, one mediator had a rebel group brag to him via WhatsApp about successful attacks on civilian targets. In addition, direct communications technology may put mediators in contact with private information they do not wish to encounter.

Many mediation teams are concerned about the **leakage of strategic information** from the mediation to broader constituencies, third parties or spoilers. This includes the leakage of information about the concessions which are being considered by conflict parties which could, for example, lead to criticisms towards negotiators and destabilise a fragile process.

The greatest concern mentioned by the mediators in interviews was for the physical safety of the parties, particularly in relation to their location.

Safety and security is a major concern for mediators using technology with parties to conflict. Some entities may have the ability to capture metadata of conversations[43], track mobile phones and other networked devices, use technology to eavesdrop on conversations, follow them as they travel, and/or map out who they call and physically meet.[44] Many parties are reluctant to conduct any dialogue of substance over digital means.

The greatest concern mentioned by the mediators in interviews was for the physical safety of the parties, particularly in relation to their location. One mediator mentioned not travelling to the funeral of a friend because he knew his presence might put other attendees at risk. The use of ICT in conflict settings heightens these concerns, notably due to the ability to track and locate any mobile device.

"In Yemen, we have interviews of negotiators who mention that the meeting could only last a certain number of minutes, which

was a calculation of how long it would take for the drones to come to attack the site… It's the counterterrorism environment, the technology used in the field, that sets the context of the negotiation. Sometimes you do it without a physical meeting because of that. Without that technology you can take your time with tea, etc., because there's no possible attempt to attack the site."

Concerns about physical safety may render other forms of surveillance mitigation useless; for example, putting multiple phones in one Faraday bag cuts off all their signals in the same location, indicating that a secret meeting is taking place among those particular users, which may put individuals in greater danger.

In addition to the risks to physical safety, risks to the **confidentiality of information** may be increased with the use of ICTs and the prevalence of hacking technologies. Privileged information can be leaked to the public. ICT leaves all parties involved in the conflict vulnerable to hacking, the loss of confidentiality, and the capability for rapid dissemination of confidential information, potentially for nefarious goals.

Sophisticated hacking technology has become more easily accessible, enabling governments and other actors to access documents, call records, text messages, browsing history and photos on remote hacked devices. Examples of hacking techniques include creating fake versions of WhatsApp, Telegram and Signal, luring users into a false sense of security.[45]

The bottom line is that the use of ICT should never be assumed to be fully secure. As the use of technology becomes more widespread, a back-and-forth has ensued between hackers, intelligence services and security experts.[46] As soon as a new type of encryption is discovered, hackers attempt to crack it. As soon as it is cracked, coders and those wishing to protect privacy attempt to find alternatives. At the moment, it seems that coders are ahead: there are forms of encryption that seem impossible to crack by

currently available technology. Yet evolution in these technologies needs to be constantly monitored.

Public communications: mediators to general public

Using social media platforms for public communication may expand the scope of participation in conflict mediation by allowing historically underrepresented groups, such as women or young people, to participate in the dialogue through direct communication, as well as through public messaging services such as Twitter or Facebook.[47] Representation in mediation processes tends to be skewed towards the powerful groups within society, who are often older, wealthy, and male. The use of ICTs may enable mediators to widen participation and enable members of the public to reach out with contributions or concerns.

By communicating about mediation processes via social media, mediators are able to mitigate the risk of seeming opaque and secretive.

Greater inclusion may also enable broader support for agreements reached through the process.[48] One mediator interviewed likened it to the democratisation of peacemaking and believed radical inclusivity may be a viable alternative to traditional elite-focused approaches. By communicating about mediation processes via social media, mediators are able to mitigate the risk of seeming opaque and secretive. They may communicate to manage expectations around the limitations of the process and to inform the public about the progress of talks. By communicating directly with the target audience, mediators also offset the risk of the parties controlling the narrative about the mediation process. Despite this, at this point few of these potential benefits seem to be actively pursued by mediators.

In addition, mediators may use online platforms to facilitate interaction between parties. In this controlled online space, parties who are not be able to be in the same room may be able to connect with one another.[49] Mediators may use curated online space

to build trust and counteract dehumanisation at times where other forms of communication are logistically impossible or unwise. Online space provides new channels through which mediators can engage in interacting with, and shaping the narratives of, conflict parties.

Risks

The mediation team's network and personal communications may pose risks in terms of **reputation and confidentiality**. There is a fine line between personal and professional views which must be maintained by the mediation team in their use of social media. 'Friending' parties on Facebook or other platforms may create the appearance of partiality, yet refusing such invitations may be ill-received. Posting personal information may also affect the safety of a mediator or his/her family.

One mediator recounted feeling very uneasy when, a week after having posted a photo of his son on his Facebook account, a conflict party asked him about his son. Another mediator recalled a situation in which a team member who was very active on social media had become "marginalised and bullied to the point that he is ostracised and can no longer present himself as neutral in the social media world. . . Anything you say can be used against you and stays there."

In addition, mediators attempting to steer public messaging about mediation processes through public communications join the fray of online voices vying for claims to truth. As a result, they may become subject to online attacks by parties intent on controlling the narrative.

As with direct communication, there is a risk of reaching only a subset of the population who have access to technology. If mediators assume that by publicly posting on social media they are reaching all stakeholders, they may be misperceiving their reach. This false sense of wider inclusion may actually cause the process to become skewed.

The use of digital platforms during the Libyan National Conference Process (NCP)

One notable innovation during the NCP, which took place between April and July 2018, was the possibility for Libyans to contribute to the process online. It was important that the NCP provided different platforms and ways for people to contribute, rather than restricting participation to physical meetings. While the internet penetration rate in Libya is rather low compared to other countries in the region, it appeared evident that the internet could be a powerful tool for informing Libyans about the process and allowing them to contribute electronically to the consultations. To achieve this, a website was specifically designed by the Centre for Humanitarian Dialogue (HD) with parameters set to allow people with low-bandwidth to access and navigate it easily. The website included information about the NCP as well as the dates and locations of upcoming meetings, visual content from past events, meeting reports, and information about how Libyans could organise their own NCP events. Most importantly, the website included an online questionnaire through which Libyans could provide their insights on the questions included in the agenda for the consultations.

Libyans were also able to express themselves on a Facebook page. However, general comments that tended to be nonspecific were not considered as part of the formal consultation process, although messages received through the Facebook inbox were considered. Most comments concerned the location of meetings, how people could contribute to the consultations, questions about the NCP and the outcomes of consultations. In addition, some used these social media tools to voice their aspirations for the future of their country and to suggest solutions to the main challenges Libya was facing, such as ending the transitional phase.

In total, the National Conference Process had 138,000 Facebook followers, while the Twitter account had about 1,800 followers. Half a million comments were generated in the course of 14 weeks. In addition, 1,700 questionnaires were completed on the Conference's website, which made up 30% of the overall contributions to the consultative phase of the NCP. These contributions were included in the final report of the NCP submitted in November 2018 to the Special Representative of the Secretary-General in Libya.

Communication between mediators and parties
Benefits and risks summary

Benefits	Risks
• Communication technologies increase security due to decreased travel. • Communication technologies allow for more frequent, discreet, real-time communication between the mediator and parties. • Communication technologies may increase trust in certain situations and can supplement in-person communications to maintain trust. • Digital communication may allow the mediator to bypass spoilers within the bureaucracy of an organisation by allowing them to communicate directly with the principle decision-makers.	• Lack of verbal intonation and non-verbal cues may decrease trust. • The use of technology tends to decrease careful composition of messages and can affect the way content is interpreted. • Mediators may not be able to confirm who is at the other end of a communication or who is communicating with whom. • Technological communication may compromise impartiality. • Digital communication tools can create physical risks through increased potential for surveillance. • Digital communication is inherently at risk of copying, distribution or manipulation.

Public communication
Benefits and risks summary

Benefits	Risks
• Public communication by the mediator may enable broader participation in the process. • Public communication about the process by the mediator may help control the narrative and increase public trust around the process. • Public communication about the process may increase support for agreements due to increased trust and transparency. • Public communication and a transparent online presence may increase trust in the mediator and improve their reputation.	• The mediator becomes part of the battle for the narrative about the mediation and risks alienating populations. • Posting, 'friending' and maintaining a visible network on social media may pose risks to the mediator's impartiality. • Sharing personal information may put the mediator at risk, and having an online presence may open the mediator up to reputational attack. • There is a risk of assuming a false sense of inclusion and exacerbating existing inequalities in access to technology.

Forms of digital threats

There are several specific forms of threat in the world of digital security:

- **Information disclosure** involves sharing information with actors for whom it was not intended. Examples include government surveillance of content or metadata, the seizing of physical devices for examination of content, the reading of unencrypted messages by any party, or the publishing of information by the interlocutor. Information disclosure is the primary concern for mediators as it relates to both physical safety and the confidentiality of their networks.

- **Spoofing** is when someone pretends to be something or someone they are not. For example, an actor may pretend to be a different party or impersonate the mediator. Software might also be spoofed, with someone creating a false version of a particular application, then using it to pull data. The Electronic Frontier Foundation and Lookout found a malware espionage campaign in the form of fake WhatsApp applications released by a nation-state that allowed it to collect photos, messages, locations and more.

- **Tampering** is when an actor changes a data point they are not supposed to modify. For example, this could include Photoshopping a picture, faking a video or editing a message from a party or the mediator, or adding/deleting data in storage.

- **Repudiation** means that the author of an act (for example, writing and sending an email) is able to deny their action. "Non-repudiation" is a legal concept which prevents successful claims of not being the source of a given piece of data through things like digital signatures.[50]

- **Denial of Service** involves preventing a software from working, such as when a state shuts down a particular service or attackers crash a site. This is often done through a Distributed Denial of Service (DDoS) attack, where multiple compromised systems, often infected with a Trojan, are used to crash a site.

- **Elevation of Privilege** is when an actor is able to do things within a system that they should not be able to do.

Considering each of these types will enable mediators to create threat models which are applicable to their specific context.

4 | Threat model and digital risk management

There are four potential responses to threats: accept, transfer, mitigate and eliminate.

Several mediators report that accepting the threat of information disclosure through network monitoring by governments is often the only realistic approach. One mediator said she regularly communicates over unencrypted channels to let state surveillance know that she is not a threat: "sometimes it is also good to let 'them' (people who monitor us) know what we do and that we are good willing people, who want to solve problems." Another said he just accepts the risk, "I just have to be careful that there's nothing you're saying that would put the mediation at risk. I operate on the basis that whether it's tape-recorded or not makes no difference."

Most mediators interviewed for this publication take the approach of transferring the risk, by following the parties' preferences in relation to what channels of communication to use. This choice places the risk of breaches of confidentiality, revelations around identities and networks, or invalid data with the parties. One mediator said, "if they're not dead, they understand their security context better than we will ever." Another mediator said "I don't know enough about the technology, so I have to assume that the other side will take into account the risks involved. There are clearly risks – people have been identified and subject to drone attacks. But I'm not in a position to provide assurances or advice on it."

Agreeing to use a party's recommended channel can potentially be seen as an endorsement of the technology. Parties "may perceive that our use of it means that it's OK. Because of the perception and our reputation, we may give credence to a particular channel and people may think the security is good. We have to do no digital harm."

Consequently, while transferring risk to parties is an acceptable choice for a mediator, it should be done from a position of some

familiarity with both the technology and potential threats rather than as a default choice. Experts in the field of Online Dispute Resolution – in which mediators work solely online in largely corporate settings – write that, "it is incumbent upon every mediator who wants to use online tools to educate himself or herself about the realistic risks that parties take when they work online. As a matter of ethics, mediators should understand how the technology works on at least a basic level and should make choices about what technology they recommend for use on the basis of that knowledge." Mediators also have a responsibility to be aware of the risks of using a particular technology.

Using a threat model can help identify the digital attacks that might be made on a system, by whom, and for what purpose.[51] Threat modelling enables risks to be analysed which can inform decision-making around responses to digital risk. Assessments need to be continuously updated to take account of the fact that technological capabilities are rapidly evolving. Something which might not be a threat today may become a threat tomorrow; and which risks can be mitigated, or transferred, may not remain constant in the future. It is therefore important to continually revisit and revise threat models.

The Electronic Frontier Foundation recommends asking the following five questions:[52]

- What do I want to protect?
- Who do I want to protect it from?
- How bad are the consequences if I fail?
- How likely is it that I will need to protect it?
- How much trouble am I willing to go through to try to prevent potential consequences?

Categories of risk – what do I want to protect?

Physical security – For parties in conflict with the state, physical safety can be a major concern which lies beneath decisions around digital security. Protecting information related to geographic location,

therefore, becomes a priority. This information may be part of the content of a message (such as in communications about the location of a planned meeting) but, more importantly, physical location is always embedded in the metadata of mobile phone communications. It can also be embedded in messaging from particular applications. To eliminate the risk of physical harm being caused by the use of technology, mediation team members might take care not to take phones with them when travelling to meet parties who are particularly sensitive to this risk (though this only mitigates the risks generated by non-digital factors). One mediator said that "my projects are a special case because they are so sensitive. We are almost anti-technology. We use only laptops, phones, and printers that have never been connected to the internet and never will be. . . . The people we are dealing with are very demanding. If they see you have a phone, they won't meet you again. They don't meet around cars, you can't wear a watch, etc. . . . I travel electronic-less." To fully eliminate threats, mediation team members must not carry or communicate with digitally-enabled devices, due to the metadata that is generated through their use.

Reputation and trust – A mediator's reputation and trusted relationships could be threatened by a number of factors including another actor generating fake online content and pretending to speak on behalf of the mediator or disclosing information or tampering with data.

Confidential information – To maintain trust and informed consent, mediation teams must maintain the confidentiality of information disclosed by the parties, ensuring no unauthorised party has access to private information. They may need to protect location and other identifying information about themselves for political reasons as well as reasons related to physical security. To mitigate the risk of information disclosure, mediation teams might encrypt or password-protect their physical devices and/or only use applications with end-to-end encryption. Mediators may want to ensure that data is stored solely on encrypted servers.

Any use of cloud computing for back-up or file-sharing purposes should be carefully assessed against hacking risks. End-to-end encryption encrypts data as it is being transmitted over the internet. While this deals with the risk of 'man-in-the-middle' attacks by preventing people from eavesdropping on the information as it is being transferred, it does not automatically mean that the data is stored securely. Data centres usually decrypt user's information once it has arrived. Many online services do not re-encrypt data before storage. Emails sent using encrypted technology are only protected if the person receiving the email also uses encryption.

Computers containing highly sensitive data should never be connected to the internet. Several mediators reported avoiding normal email accounts and instead using draft folders on cloud-based file-sharing platforms such as Google Drive or Dropbox for slightly more secure communications. Some prefer to avoid all use of electronic communication and solely using face-to-face communication in the most sensitive cases.

Confidentiality around networks and identity – Protecting the identity of interlocutors, the relationship of a specific individual to a specific phone number, and the relationship of multiple phone numbers with one another are important elements of network confidentiality. These threats relate largely to physical safety, particularly for political dissidents. One mediator said, "I'm exercised by this concern that a call from me if I'm not careful can get someone arrested. Or targeted for killing. I've lost quite a few interlocutors to targeted killings. I don't think any because of their relationship with me but it's a constant reminder of the danger of their lives." Confidentiality around networks is a longer-term concern than simply physical safety. Conflict parties who find their identities compromised may experience negative impacts on more than their physical safety, such as blacklisting, exclusion from the political process, or financial effects. To mitigate the risk of breaches to network confidentiality, mediators may refrain from contacting an individual over the same channels or on the same device as

another. Another aspect of this risk is the need to confirm the identity of participants through technology, so the mediator knows they are interacting with the appropriate people. Secure messaging applications such as Signal offer a degree of assurance, providing they have not been subject to tampering. Mediators must confirm the identity of their online interlocutors while also protecting this information.

Most email providers increasingly embed Artificial Intelligence (AI) features into their service. While this can enhance productivity, for example by automatically offering to add to one's diary an appointment proposed in the body of an email, it also presents confidentiality concerns. Some email providers are notorious for the extent of the data they collect, including contact information. More generally, all internet activity, such as Google searches, are stored and are thus searchable to reveal information about an individual's online profile.[53]

Computers containing highly sensitive data should never be connected to the internet. Several mediators reported avoiding normal email accounts and instead using draft folders on cloud-based file-sharing platforms.

The integrity of data – Mediators need to ensure that no communications have been tampered with, edited or falsified. They may also be concerned about the veracity and authenticity of messages received over social media. For example, data collected from social media may include 'bots' or automated accounts serving to amplify particular messages. Similarly, one mediator interviewed for this publication mentioned the use of a Photoshopped image to incite anger.

In all instances, it is critical that mediators apply all basic cyber-hygiene rules, including installing robust anti-virus programmes, applying updates and security patches as soon as they are released by the vendors, and not opening suspicious attachments or hyperlinks.

Summary of steps to manage ICT risk

The following framework offers an approach for considering the technology a mediator may be using and to understand its impacts on the process or outcome.[54] Mediation teams may or may not have the ability to influence which applications are used by conflict parties, but mediators do have a duty to understand and address risks that may be created.

1. Prioritise aspects of the mediation process

The mediation teams must be clear on the most important aspects of the mediation process in the particular context in which they are working. This prioritisation will inform the tools used, and the choices made, regarding the use of technology.

2. Conduct a threat assessment

The mediation team should assess the threats to the process. Understanding any malicious actors with the intent and capability to threaten the use of information and communication technologies will inform the mediator's selection of these and their use.

3. Analyse tool selection

The mediator should choose the technological tools which best support their priorities while being aware of any limitations imposed by threats to the process. Consider whether the tool will enable those involved to mitigate, transfer, eliminate, or accept risks around digital security.

4. Analyse process choices

New technologies can be used in a way which maximises the benefits while minimising the threats. For example, establishing guidelines with the conflict parties on when they can post information about the content of a draft agreement may help manage some of the risks resulting from the use of social media. Process decisions can help offset risks introduced by technology and harness the benefits.

5. Mitigate disinformation risks

The parties, factions and interested third parties may deploy social media and online news for strategic gain. This may be through legitimate means, or it may be through malicious manipulation of events and perceptions. The mediator should identify the stakeholders that are active in the space and what their influencing goals may be, including who they are targeting and to what end. Specific communications can be evaluated using the following test:[55]

- **Currency**: How recently was this information published/posted? Can you find a publication date?

- **Reliability**: Is the information supported by evidence? Can it be confirmed by other sources?

- **Authority**: Who wrote the information – are they an expert or knowledgeable in their field?

- **Accuracy**: Is the information supported by evidence? Can you verify any of the information in another source or from personal knowledge? Does the language or tone seem unbiased and free of emotion?

- **Purpose / point of view**: Why was it written? To sway opinion? Is it biased toward a particular point of view?

There are also examples of online educational and fact-checking tools designed to combat misinformation (see Annex B). Many of these tools remain regionally-focused, but the tools are rapidly expanding.

Mediators may need to consider action to mitigate disinformation and misinformation. This may be done by seeking to influence parties privately regarding their information campaigns, or for the mediator to consider using their own social media strategy to counter disinformation and misinformation by communicating directly with the public. This approach is, however, contentious, and is an area of mediation practice which deserves further debate.

5 | Conclusion

Using information and communication technologies for direct and public communications has an impact on the substance, process and security of conflict mediation and on the reputation of a mediation team. Simply using such technology for analysis may affect the process by providing a much wider scope of knowledge than was previously available, while also raising the risks around biased information. As technology continues to advance, mediation teams will have new options and tools for improving communication, conflict analysis, and situational awareness.

The UN Guidance for Effective Mediation identifies a number of key fundamentals that should be considered in a mediation effort. The way ICT is used is relevant to each of these fundamentals:

- **Preparedness**: as technology is increasingly used for preparation and communication in mediation processes, risks associated with the use of technology, such as false information, surveillance and hacking, must be considered in the preparation phase.
- **Consent, impartiality and confidentiality**: to achieve informed consent, maintain impartiality and adhere to confidentiality guidance, parties must be aware of how their communications will be protected, how data will be stored and/or deleted, and what risks they are incurring through their agreement to participate. For example, if a mediation team is aware that parties are unwittingly using a non-encrypted technology that could easily give an adversary access to the content and metadata of communications, the mediation team must be able to make an informed decision about the implications of this for the relevant parties' consent and confidentiality. The team must also decide whether or not to alert parties, which creates challenges in terms of impartiality.
- **Inclusivity**: to achieve inclusivity, mediation teams should be aware of the risks and benefits associated with ICTs, particularly social media. Using these tools for analysis may widen the process to a broader range of contributors. However, they can

also give the mediation team a false sense of balance and inclusion if some groups are underrepresented on social media or if others use bots and similar techniques to dominate and shape narratives.

• **National ownership**: mediation teams must find an appropriate balance between the protection of parties to the conflict, who may be adversaries to the state, and respecting a state's right to surveil its people within the confines of the law. Many mediation teams already implicitly apply these guidelines to their use of ICT and we argue that it is incumbent upon all to do so explicitly.

Mediation teams have a responsibility to be generally literate about the technologies present in the mediation environment and their effect on the mediation process, and to make informed choices about their use. Even in situations where mediation teams follow the parties' advice in relation to platforms and security, teams should be aware of the implications of these choices and the possible effects on the mediation process.

This publication has presented a range of considerations and a framework for mediation teams to use to evaluate the ICT tools they either choose to, or are asked to, use. No single platform or technology is recommended for all mediations but mediation teams should identify the effects, risks and benefits of ICTs and consider using a framework for informed decision-making with regards to technology. The following annexes include a list of current technologies and their impact on peacemaking as well as a list of counter-disinformation tools. As technologies continue to develop, new privacy laws come into place, and new forms of 'hacking' and 'cracking' come into being, the considerations outlined in this publication may help mediation teams make informed choices about future technologies and their effect on the design and content of a mediation process.

Although the choice of medium has significant effects on mediation processes, the use of technology has not, so far, fundamentally

altered the practice of mediation itself. This practice still relies first and foremost on the trust built between a mediator, or a mediation team, and the parties, and the ability to generate and maintain buy-in to the process. Technological tools may provide benefits and risks in addition to this, but they do not change its essence. The approaches presented in this publication can increase the ability of mediation teams to assess trade-offs in terms of benefits and risks in relation to the use of communication technology and enable them to make more informed choices during the mediation process. These trade-offs would benefit from further research and debate, particularly given the rapid pace of technological innovation. Answers that apply today may not apply tomorrow.

Annexes

Annex A – New technologies

There is a variety of new tools that are being developed to aid conflict resolution and peacemaking, and which offer increasing capacity for data analysis. The following table gives some examples which may have an impact on conflict resolution in the near future.

Technology	Description	Impact on peacemaking
Machine learning	A subset of artificial intelligence. "The machine's ability to keep improving its performance without humans having to explain exactly how to accomplish all the tasks it's given."[56] Machine learning is driving everything from voice recognition to cancer diagnosis.	Machine learning will allow the automation of some elements of peacemaking. Some areas such as early warning systems and conflict analysis are ripe for automation as they use standard data inputs.
Blockchain	"Blockchain is an open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way."[57] Blockchain can prevent data deletion, tampering, and revision, and allow for complete transparency of all data interactions.	Blockchain is seen as a foundational technology that will be transformative across a range of industries. Applications to peacemaking are still being defined.

Technology	Description	Impact on peacemaking
Remote sensing, satellite and in-situ observation	A variety of tools which are used to collect information remotely. This includes satellite imagery, seismic sensors used to track the movement of people and vehicles, and auditory sensors to detect gunfire and aircraft.	Satellite imagery is being used to analyse humanitarian disasters, both those caused by nature and those caused by conflict. They may also be used to identify military hardware and the destruction of civilian infrastructure.
Early warning systems	Holistic early warning systems are utilising social media, remote sensing, and artificial intelligence to predict impending conflict or rapidly warn those affected by conflict.	Early warning systems can be used to support organisations operating in conflict zones by collecting a cross-section of data from many sources to warn of threats. In addition, there is a growing range of ICT based tools that monitor crowd behaviours and can be used to defuse tensions before they erupt in violence.

Annex B – Examples of counter-disinformation tools

Tool	Description	Website
First Draft	This is a project set up by the Shorenstein Center on Media, Politics and Public Policy at Harvard University's John F. Kennedy School of Government. It uses research-based methods to fight misinformation and disinformation online. It also provides practical and ethical guidance on how to find, verify and publish content sourced from the social web.	https://firstdraftnews.org/
CrowdTangle	A tool (recently bought by Facebook) which is used to monitor the spread of information through social media. Custom alerts can be set up.	http://www.crowdtangle.com/
Verification Handbook	This handbook has been authored by journalists from the BBC, Storyful, ABC, Digital First Media and other verification experts. It is a resource for journalists and aid providers and provides tools, techniques and step-by-step guidelines for dealing with user-generated content (UGC) during emergencies.	http://verificationhandbook.com/
FotoForensics	A free tool that does compression analysis on pictures to flag potential manipulation.	http://fotoforensics.com/
Who Tweeted it First	A tool that identifies the very first use of a hashtag on Twitter, to identify the origin of a story or movement.	http://ctrlq.org/first
Snopes	One of the first online fact-checking websites. Focuses on the USA.	https:// www.snopes.com

Endnotes

1   Statista, *Global Digital Population 2018*. Retrieved from https://www.statista.com/statistics/617136/digital-population-worldwide/

2   Meier, Patrick. (31 December 2011). *New Information Technologies and Their Impact on the Humanitarian Sector*. International Review of the Red Cross 93, No.884.

3   See, for example, United Kingdom Department for International Development. (January 2018). *DFID Digital Strategy 2018 to 2020: Doing Development in a Digital World." Department for International Development. Retrieved from* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/701443/DFID-Digital-Strategy-23-01-18a.pdf

4   Katsh, Ethan, and Rifkin, Janet. (2001). *Online Dispute Resolution: Resolving Conflicts in Cyberspace*. 1st edition. Jossey-Bass; Larson, David. (2004). *Online Dispute Resolution: Technology Takes a Place at the Table.* Negotiation Journal 20, No.1, pp.129-135.

5   Big data refers to large amounts of data produced very quickly by a high number of diverse sources. Data can either be created by people or generated by machines, such as sensors gathering climate information, satellite imagery, digital pictures and videos, purchase transaction records, GPS signals, etc. A number of projects and initiatives have already been launched exploring how the analysis of this data impacts the peacemaking field. For example, the authors would like to mention the development of the #CyberMediation Initiative which was launched in March 2018 by a consortium composed of the United Nations Department for Political Affairs, the Centre for Humanitarian Dialogue, DiploFoundation and swisspeace. The consortium organises regular workshops on topics related to the impact of new technologies on armed conflict mediation. These workshops can be accessed remotely.

6   Eurostat. (20 September 2016). *Eurostat Glossary. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Information_and_communication_technology_(ICT)*

7   Merriam Webster online dictionary, Accessed on 26 November 2018. Retrieved from https://www.merriam-webster.com/dictionary/social%20media

8   Walter, Barbara. (2017). *The New New Civil Wars*. Annual Review of Political Science 20, p.482.

9   Mor, Yifat, Ron, Yiftach and Maoz, Ilfat. (2016). *'Likes' for Peace: Can Facebook Promote Dialogue in the Israeli–Palestinian Conflict?* Media and Communication 4, No.1, pp.15–26.

10 See ICT4peace.org: https://peaceinnovation.stanford.edu/; howtobuildpeace.org; and www.peacetechlab.org. Similar initiatives self-identify using #peacetech and #ICT4Peace.

11 See Oslo Forum Interview. (2018). *Exiting Chaos: Ghassan Salamé reflects on peacemaking*. Centre for Humanitarian Dialogue.

12 United Nations; World Bank. (2018). *Pathways for Peace: Inclusive Approaches to Preventing Violent Conflict*. World Bank, Retrieved from https://openknowledge. worldbank.org/handle/10986/28337

13 Walter, Barbara. (2017). Pp.469-486.

14 Disinformation corresponds to giving false information deliberately, while misinformation is giving false information without malice.

15 See jm/tj (Reuters, AFP). *Nigeria to launch probe into 2007, 2015 elections over SCL-Cambridge Analytica. Deutsche Welle:* www.dw.com/en/nigeria-to-launch-probe-into-2007-2015-elections-over-scl-cambridge-analytica/a-43228067; Epstein, Helen. (30 August 2017). *Kenya: The Election & the Cover-Up*. New York Review of Books: https://www.nybooks.com/daily/2017/08/30/kenya-the-election-and-the-cover-up/; Segal, David. (3 February 2018). *How Bell Pottinger, PR firm for despots and rogues, met its end in South Africa*. New York Times: https://www. nytimes.com/2018/02/04/business/bell-pottinger-guptas-zuma-south-africa.html

16 Khalaf, Roula. (25 July 2018). *If you thought fake news was a problem, wait for 'deep-fakes'*. Financial Times. Retrieved from https://www.ft.com/content/ 8e63b372-8f19-11e8-b639-7680cedcc421

17 Williams, James. (2018). *Stand Out of Our Light: Freedom and Resistance in the Attention Economy.* Cambridge University Press.

18 United Nations; World Bank. (2018).

19 Smith, Claire, Burke, Heather, de Leiuen, Cherrie and Jackson, Gary. (June 2016). *The Islamic State's Symbolic War: Da'esh's Socially Mediated Terrorism as a Threat to Cultural Heritage*. Journal of Social Archaeology 16, No.2, 1, pp.164–88. Retrieved from https://doi.org/10.1177/1469605315617048

20 See: https://www.vientianepost.com/2018/03/26/is-facebook-contributing-to-genocide-in-myanmar

21 See: https://www.reuters.com/article/us-myanmar-rohingya-facebook/u-n-investigators-cite-facebook-role-in-myanmar-crisis-idUSKCN1GO2PN

22 Idem.

23 See: https://www.irrawaddy.com/opinion/editorial/facebook-slow-react-violence-hate-speech-myanmar.html. The contribution of social media to hate speech was also highlighted by the Declaration of a State of Emergency by the Government of Sri Lanka in March 2018 after attacks against the minority Muslim population in Sri Lanka that led to at least one death (see https://www.nytimes.com/ 2018/03/06/world/asia/sri-lanka-anti-muslim-violence.html).

24 Ebner, Noam. (2012). "ODR and Interpersonal Trust" in Wahab, Mohamed S. Abdel, Katsh, Ethan and Rainey, Daniel (Eds.). Online Dispute Resolution: Theory and Practice. A Treatise on Technology and Dispute Resolution. Eleven International Publishing, pp. 215-248.

25 Winslade, John and Monk, Gerald. (2000). *Narrative mediation: a new approach to conflict resolution*. Jossey-Bass.

26 Examples of online platforms for dialogue and intercultural exchange include Erasmus+ Virtual Exchange (https://europa.eu/youth/erasmusvirtual_en); Soliya's Engagement Programs (https://www.soliya.net/programs/engagement-programs);

http://pol.is, an interactive, crowd-sourced survey tool which can generate maps of public opinion to discover the nuances of agreement and disagreement on contentious issues; www.kialo.com, an online tool for collaborative decision-making; and www.micgoat.com, a video application where users engage in one-to-one discussions.

27 Build Up and MISTI. (2018). *The Commons: A pilot methodology for addressing polarization online*. Build Up and MISTI.

28 See Mancini, Francesco. (April 2013). *New Technology and the Prevention of Violence and Conflict*. International Peace Institute; Godoy, Maria. (2 October 2012). *Can Riots Be Predicted? Experts Watch Food Prices*. NPR; and Leetaru, Kalev. (2011). *Culturomics 2.0: Forecasting Large-Scale Human Behavior Using Global News Media Tone in Time and Space*. First Monday Vol.16, No.9.

29 Wojcik, Stefan, Messing, Solomon, Smith, Aaron, Rainie, Lee, and Hitlin, Paul. (9 April 2018). *Bots in the Twittersphere.* Pew Research Center.

30 See http://groundviews.org/2018/01/24/namal-rajapaksa-bots-and-trolls-new-contours-of-digital-propaganda-and-online-discourse-in-sri-lanka/

31 Solon, Olivia. (18 December 2017). *How Syria's White Helmets became victims of an online propaganda campaign.* The Guardian. Sock puppet accounts is an online identity specifically created for the purpose of deception.

32 Mancini, Francesco. (April 2013).

33 Mancini, Francesco. (April 2013).

34 Perugini, Maria Roberta. (21 April 2016). *Why GDPR. Europrivacy.*

35 Greenwood, Faine, Howarth, Caitlin, Escudero, Poole, Danielle, Raymond, Nathaniel A. and Scarnecchia, Daniel P. (January 2017). *The Signal Code: A Human Rights Approach to Information during Crisis. Harvard Humanitarian Initiative.*

36 However, in cases like these, privacy and do-no-harm concerns may need to be assessed, depending on the likelihood that the conversation may be under surveillance.

37 Heide, Jan B. and Miner, Anne S. (1 June 1992). *The Shadow of the Future: Effects of Anticipated Interaction and Frequency of Contact on Buyer-Seller Cooperation*. Academy of Management Journal 35, No.2, pp.265–291.

38 Ebner, Noam. (2012).

39 Ebner, Noam. (2012).

40 Ebner, Noam, Bhappu, Anita, Brown, Jennifer, Kovach, Kimberlee and Schneider, Andrea. (2009). *You've Got Agreement: Negoti@ting Via Email*. Marquette Law School Legal Studies Paper No. 09-16.

41 Ebner, Noam. (2012).

42 Ebner, Noam. (2012).

43 See International Committee of the Red Cross and Privacy International. (October 2018). *The humanitarian metadata problem: "Doing no harm" in the digital era*. International Committee of the Red Cross and Privacy International.

44 Schneier, Bruce. (2016). *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. W. W. Norton & Company.

45 For example, in 2018 The Electronic Frontier Foundation (EFF) and the mobile security company Lookout uncovered a malware espionage campaign infecting the devices of thousands of people in more than 20 countries, which they attributed to Lebanon's intelligence agency. See EFF press release from 18 January 2018: *EFF and Lookout Uncover New Malware Espionage Campaign Infecting Thousands Around the World*.

46 Schneier, Bruce. (2016).

47 Paffenholz, Thania. (2014). *Broadening Participation in Peace Processes*. Mediation Practice Series. Centre for Humanitarian Dialogue.

48 Paffenholz, Thania. (2014).

49 Mor, Yifat, Ron, Yiftach and Maoz, Ilfat. (2016). Pp.15–26.

50 "A digital signature is a mathematical scheme for verifying the authenticity of digital messages or documents." See https://en.wikipedia.org/wiki/Digital_signature

51 Cobb, Camille, Sudar, Samuel, Reiter, Nicholas, Anderson, Richard, Roesner, Franziska, and Kohno, Tadayoshi. (June 2016). *Computer Security for Data Collection Technologies*. ICTD Lab and Computer Security & Privacy Research Lab, University of Washington.

52 For information on digital security, see the Electronic Frontier Foundation *Guide to Surveillance Self-Defense* available at https://ssd.eff.org/en

53 See, for example, Dylan Curran. (30 March 2017). *Are you ready? Here is all the data Facebook and Google have on you. The Guardian.*

54 For another framework for analysing the context-specific use of technology in a given area, see Social Impact Lab's *Context Analysis Framework*.

55 Renirie, Rebecca. *Research Guides: Website Research: CRAAP Test*. Central Michigan University.

56 See Brynjolfsson, Erik, and Mcafee, Andrew. (2017). *The business of Artificial Intelligence: What it can–and cannot–do for your organization*. Harvard Business Review Digital Articles 7: 3-11.

57 See Iansiti, Marco, and Karim R. Lakhani. (2017). *The Truth About Blockchain*. Harvard Business Review 95.1, pp.118-127.

Further reading

Brown, Rachen, Heinzelman, Jessica and Meier, Patrick. (2011). "Mobile Technology, Crowdsourcing and Peace Mapping: New Theory and Applications for Conflict Management" in *Mobile Technologies for Conflict Management*. Law, Governance and Technology Series, Vol.2 (edited by Maria Poblet), pp.39–53. Springer.

Dorn, A. Walter. (July 2016). *Smart Peacekeeping: Toward Tech-Enabled UN Operations.* Providing for Peacekeeping no.13. International Peace Institute.

Ebner, Noam. (2012). "ODR and Interpersonal Trust*"* in Wahab, Mohamed S. Abdel, Katsh, Ethan and Rainey, Daniel (Eds.). *Online Dispute Resolution: Theory and Practice. A Treatise on Technology and Dispute Resolution*. Eleven International Publishing, pp. 215-248. Front Line Defenders. *Digital Security Resources*. See https://www.frontlinedefenders.org/en/digital-security-resources (last accessed on 26 November 2018)

Electronic Frontier Foundation. *Surveillance Self-Defense. Tips, Tools and How-tos for Safer Online Communications.* See https://ssd.eff.org/en (last accessed on 26 November 2018)

Gienger, Viola. (13 April 2013). *'Big Data,' Text Messages Can Aid, Not Drive Conflict Prevention*. United States Institute of Peace.

Harvard Humanitarian Initiative, Signal Human Security + Technology, *The Signal Code: A Human Rights Approach to Information During Crisis.* See https://signal code.org/ (last accessed on 26 November 2018)

Himelfarb, Sheldon. (25 April 2014). *Can Big Data Stop Wars Before They Happen?* United States Institute of Peace.

ICRC, The Engine Room, Block Party. (January 2017). *Humanitarian Futures for Messaging Apps.* ICRC, The Engine Room, Block Party.

Kenyon, Miles. (9 November 2017). *Secure Your Chats: Why Encrypted Messaging Matters.* The Citizen Lab.

Mancini, Francesco. (April 2013). *New Technology and the Prevention of Violence and Conflict.* International Peace Institute.

Meier, Patrick. (31 December 2011). *New Information Technologies and Their Impact on the Humanitarian Sector.* International Review of the Red Cross 93, No.884. OCHA. (2012). *Humanitarianism in the Network Age: Including World Humanitarian Data and Trends 2012.* OCHA Policy and Studies Series.

Patrikarakos, David. (2017). *War in 140 characters: how social media is reshaping conflict in the twenty-first century.* Hachette UK.

About the authors

Joelle Jenny was, when she conducted this research, an Associate at the Weatherhead Centre for International Affairs at Harvard, and a member of the World Economic Forum's Global Council on the Future of International Security. Her research at the time focused on disruptive technologies, international peace and security, and conflict prevention. She now heads the United Kingdom's Joint Funds Unit, which oversees the UK's cross-government Prosperity Fund (PF) and Conflict, Security and Stability Fund (CSSF). From 2011 to 2016 she was Director for Security Policy and Conflict Prevention at the EU's External Action Service in Brussels.

Rosi Greenberg is a mediator, facilitator, and leadership coach based in Cambridge, MA. Rosi holds a Master's Degree in Public Policy from the Harvard Kennedy School of Government, where she studied conflict transformation and leadership development. Prior to coming to Harvard, Rosi lived in Amman, Jordan and worked at Questscope, an NGO focusing on social development for youth and their communities.

Vincent Lowney became interested in conflict mediation while serving in Afghanistan. Following his military service, Vincent studied negotiation theory and alternative dispute resolution at the Harvard Kennedy School of Government. He has mediated in Boston courts and is interested in the transformative impact of new technologies on mediation and diplomacy.

Guy Banim is an expert on preventive diplomacy with a range of field (eg Afghanistan, Nepal, Zanzibar, Mozambique, Myanmar) and HQ experience. Currently adviser to the European Institute of Peace, he has previously worked for the EEAS Mediation Support Team (2011-2014), EU Institute of Security Studies/Council for Security Cooperation in Asia Pacific (2015-2017), AU Peace and Security Department/African Peace & Security Architecture programme (2009-2011), Crisis Management Initiative (2009 & 2006) and the EC conflict prevention unit (2001-2004). He began his career as a Northern Ireland civil servant implementing the Belfast Agreement. He tweets @guybanim