

Geneva, 30 April 2023

# Global Digital Compact: Submission from the International Committee of the Red Cross

### Introduction

The International Committee of the Red Cross (ICRC) is an independent, impartial, and neutral organization that provides humanitarian assistance and protection to victims of armed conflicts and other situations of violence and is the guardian of international humanitarian law (IHL). Digital technologies impact communities and dynamics in these environments; are today used as a means or method of warfare in the conduct of hostilities; and play a part in humanitarian space and operations. Thus, understanding and responding to the role that technologies play in the modern world is an important part of the ICRC's efforts to protect civilians and minimize the impact that armed conflict and other situations of violence have on vulnerable communities.

It is critical that armed conflict and other situations of violence are part of the conversation around digital governance. For communities affected by these situations, the stakes for effectively governing the effects of cyber operations that disrupt essential services, digital surveillance, mis- and disinformation and hate speech, access to lifesaving information could not be higher - but debates around how to protect people from potential harms often omit considerations for how potential governance approaches might fail or be circumvented in times of crisis. Further, these discussions must encompass the ways in which technologies can create or compound specific harms and vulnerabilities in these situations. For example, losing access to digital connectivity can prevent people from keeping in touch with loved ones and gathering information essential to their survival, protection, and well-being so that they can make informed decisions. Hostile cyber operations risk disrupting essential services, such as electricity, water, or medical care. Digital data can be used to identify and target specific groups of people, and mis- and disinformation can drive armed conflict and other situations of violence or put at risk humanitarian actors and the communities accessing their services. Yet States and communities affected by these situations are less likely to have the resources to engage in global digital governance debates. As we consider how to best govern our digital world, it is imperative that we bring the voices and experiences of communities affected by conflict and violence into these discussions and keep their needs at the forefront of our minds.

The potential human costs of the use of new digital tools during armed conflict and violence must also be considered as we grapple with digital governance issues. Using new technologies, it is now easier than ever to involve civilians in military cyber operations, for example by engaging them in offensive cyber operations against enemy targets; repurposing civilian mobile apps for military intelligence gathering; and through the provision of digital infrastructure and services by private civilian companies to belligerents. Such scenarios bring civilians much closer to the conduct of military operations, can easily do so at scale and risk exposing civilians to serious harm, for example where the collection of information is regarded as a participation in hostilities.

IHL applies and must be complied with in all situations of armed conflict and to all types of operations, including digital ones. It imposes long-standing limits in warfare which are essential to protect civilians and their infrastructure against harm. To uphold the fundamental principle of distinction, it is particularly important that states prevent the osmosis between civilian and military digital tools. Since 2019, the ICRC has consistently voiced concern about several of these issues in the UN 'Open-ended working group on developments in the field of information and telecommunications in the context of international security', focusing in particular on the potential human cost of cyber and other digital operations against civilian infrastructure, medical facilities, and humanitarian organizations. The ICRC continues to call on States to

interpret – and apply – existing rules of IHL in a manner that ensures adequate and sufficient protection for civilians and civilian infrastructure, ICT systems and data in our ever-increasingly digitalized societies.

Digital technologies, including artificial intelligence, are also increasingly being integrated into the functioning of physical weapon systems, including autonomous weapon systems. The ICRC has made it clear that, while existing IHL rules already set constraints, new international rules with specific prohibitions and restrictions on autonomous weapons are needed due to the particular humanitarian, legal and ethical concerns. Such rules should specifically prohibit unpredictable autonomous weapons, which may include AI-controlled autonomous weapons and autonomous weapons designed or used to target humans directly. New international rules should include strict constraints on the development and use of other autonomous weapons.

Humanitarian organizations must harness the opportunities offered by digital transformation ethically and responsibly to deliver positive outcomes for communities affected by conflict and disasters. This includes the promotion and application of robust data protection standards in the humanitarian sector. For neutral, impartial, and independent actors such as the ICRC, the types and providers of technology and internet services they use risk to influence how they are perceived by the parties to an armed conflict, thus potentially jeopardizing their reputation and acceptance, and ultimately their safe and continued access to populations in need. Private sector services are increasingly the backbone of modern operations for international humanitarian organizations. They face the challenge of meeting organizations' complex needs and requirements, such as the exclusive control over their humanitarian data and communication in line with their diplomatic privileges and immunities. Care must be taken to preserve common digital spaces and the humanitarian organizations' ability to operate safely in the digital sphere. This requires sustained dialogue and engagement between humanitarian organizations, States, technology providers and other relevant stakeholders. In parallel, humanitarian actors must find new ways to meet the needs of their organizations and the people they serve in the digitally interconnected world, while preserving the confidentiality of the information they collect and their ability to act according to the humanitarian principles.

The ICRC welcomes the opportunity to contribute to the Global Digital Compact and wishes to offer the following thoughts and recommendations on some of the proposed thematic areas. The text below matches the submissions made to the thematic survey.

### Theme 3: Data Protection

## Core principles

Humanitarian data is integral to the assistance and protection of people affected by armed conflict and other situations of violence. But new digital risks, such as cybercrime or targeting through digital means, can result from humanitarian data collection or from unauthorized access, manipulation, deletion or misuse. People's privacy, dignity and safety can be put at risk, and their fundamental rights violated. Trust in humanitarian organizations and their neutrality can be undermined, constraining their access and activities, and putting them at risk.

To protect the activities of impartial humanitarian organizations, IHL places obligations on parties to armed conflict that provide protections for relief operations. Impartial humanitarian organizations have a right to offer their humanitarian services. Once impartial humanitarian relief operations have been agreed to, the parties must allow and facilitate rapid and unimpeded passage of these relief operations. Such operations must not be unduly interfered with, including via digital means. Humanitarian operations and personnel must be respected and protected: they must not be harmed in any way and must be protected against harm by private actors.

These obligations apply equally to the digital activities and data of impartial humanitarian organizations. The ICRC regularly emphasizes the obligations parties to armed conflict have under IHL, including cyber operations that have harmful effects on data belonging to specifically protected entities (e.g. hospitals) and other data belonging to humanitarian actors. Alongside the obligations of IHL, humanitarians should

effectively implement cyber security measures and data protection practices and collect and use data from communities only for humanitarian purposes.

In 2022, a resolution of the Red Cross and Red Crescent Movement reaffirmed these points and called on States to recognize the protected status of humanitarian data online as they do offline, strengthen legal and other protections for such data, safeguard such organizations from harmful cyber and information operations, and refrain from requesting or using such data in a way that might undermine humanitarian principles and trust. The Resolution also committed the ICRC to continue to explore the feasibility of a 'digital emblem' to identify data and digital infrastructure entitled to display the distinctive emblems recognized under IHL and to indicate, where applicable, their legal protection.

### Key Commitment/Pledges/Actions

States should recognize the protected status of humanitarian data online as they do offline, strengthen legal and other protections for humanitarian data, safeguard humanitarian organizations from harmful cyber and information operations, and refrain from requesting or using humanitarian data in a way that might undermine humanitarian principles and trust.

Under IHL, parties to armed conflict must place limits on their cyber operations, including those operations that the data belonging to specifically protected entities that belongs to specifically protected entities. The obligations to respect and protect medical facilities and humanitarian relief operations extends to medical data belonging to those facilities and data of humanitarian organizations that are essential for their operations. Similarly, deleting or otherwise tampering with data in a manner that renders useless objects indispensable to the survival of the civilian population, such as drinking water installations and irrigation systems, is prohibited. Beyond these categories, it is the ICRC's position that it would be difficult to reconcile with the object and purpose of IHL legal interpretations that operations against other essential civilian data would not be prohibited by IHL in today's ever more cyber-reliant world. Additional measures that States could take to protect civilian infrastructure from the harmful effects of attacks might include segregating military from civilian cyber infrastructure and networks, segregating military data from civilian data, and segregating computer systems on which essential civilian infrastructure depends from the internet.

The Red Cross and Red Crescent Movement has committed in its 2022 Council of Delegates resolution on Safeguarding Humanitarian Data to strengthening its own humanitarian data protection and cyber-security, including through the ICRC's own Data Protection Framework, and continue researching the technical feasibility of a digital emblem.

More broadly, States should continue to study and work with partners like the ICRC to understand the risks associated with cyber and information operations, including escalation of conflict and inciting violence, and of new technologies such as AI and IoT, and the quantum-enabled increase in the speed and scale of cyber and other operations, in order to continue to develop our understanding of how civilians are to be protected and how States are to act in this space.

# Theme 5: Introduce accountability criteria for discrimination and misleading content

### Core Principles

The use of digital technologies to spread harmful information in places affected by armed conflict and other situations of violence is a growing concern for the ICRC. While harmful information has always existed, digital social media have transformed the speed and scale at which mis- and disinformation and hate speech (MDH) travel. In situations of conflict and violence, MDH can exacerbate risks of harm for vulnerable people. For example, online hate speech can incite violence against a minority group or can contribute to psychological and social harm through harassment, defamation and intimidation, and lead to persecution, discrimination or displacement. MDH, polarization and violence can drive one another in a vicious cycle, and legal, policy and institutional safeguards are often dysfunctional or lacking. Social media platforms can also be vectors for

inciting IHL violations when used to, for example, encourage attacks against civilians, hospitals, or to threaten violence with the primary purpose of spreading terror among the civilian population, or subject prisoners of war to public curiosity.

MDH can also constrain humanitarian access and operations. False and manipulated information can erode trust, undermine communities' acceptance of humanitarian organizations, and cause reputational damage. This can reduce operational capacities and restrict the space for humanitarian action, potentially leaving affected people's needs unattended. MDH can also lead to threats or attacks against humanitarian staff. Under such circumstances, information that harms humanitarian organizations would violate IHL.

The use of propaganda, mis- or disinformation during armed conflict is not, as such, prohibited under IHL. However, as illustrated above, certain activities may violate specific IHL as well as other rules of international law, in particular human rights law. It is also essential that, in their efforts to prevent and mitigate the risks of MDH, States ensure respect for the safety, dignity and fundamental rights of people affected by conflict, and protect their access to humanitarian information and essential services necessary to their survival and wellbeing.

### Key Commitment/Pledges/Actions

This critical challenge requires multi-stakeholder collaboration and dialogue which considers conflict-specific challenges, needs and circumstances.

Opportunities to prevent and mitigate the risks of MDH include enhanced due diligence and prevention measures, collaboration on research, and policy development. Existing State practice demonstrates that certain policy and practical measures can have a positive effect on the prevention and mitigation of MDH. Respecting the principles of transparency, accountability and personal data protection can help support tech and social media companies to improve their policies and practices, due diligence mechanisms and corporate social responsibility. Finally, encouraging awareness raising, digital literacy, and the implementation of effective safeguards and accountability mechanisms are essential ingredients for progress.

The ICRC urges States and other relevant local authorities, including non-state armed groups, to take all necessary measures to ensure that information or influence operations carried out in the context of armed conflict settings comply with international law. It is important that measures that States may take to 'counter' MDH, are compliant with IHL and other applicable rules of international law, in particular human rights law. Such activities should not trigger or contribute to negative consequences for civilian populations and local communities, for instance by making family contact impossible. They should also not target or negatively impact the humanitarian organizations mandated to protect and assist them, or their ability to operate in accordance with the principles of neutrality, impartiality, and independence.

The ICRC urges States and other relevant authorities to take all necessary measures to preserve and protect the space for neutral, independent, and impartial humanitarian action, information, and communication, and to protect humanitarian actors from MDH-related political instrumentalization that would endanger their access and security.

Tech and social media companies can and must do more to help prevent and mitigate MDH and related consequences in contexts affected by conflict. The ICRC calls on them to adopt a conflict-sensitive approach to their activities and to develop and maintain enhanced due diligence policies and practices in the development of, prior to, and after launch of their products in those contexts.

# Theme 6: Promote regulation of artificial intelligence

# Core principles

Any new technology of warfare must be capable of being used, and must only be used, in compliance with existing rules of IHL. There is already a vast array of ways in which parties to armed conflict – whether States or non-State armed groups – might use AI, including ML, in warfare. The ICRC has identified at least three overlapping areas that are relevant from a humanitarian and IHL perspective.

One area of concern is the use of AI to increase autonomy in weapon systems: The ICRC has consistently raised concerns around autonomous weapons systems (AWS), notably regarding the potential loss of human control over the use of force. This challenge is likely to be exacerbated when these systems incorporate AI, particularly ML. To clarify the specific design and use limitations deriving from existing IHL and required by ethical perspectives, the ICRC has recommended that States negotiate and adopt new, binding international law: (1) prohibiting AWS with unpredictable effects, and AWS designed or used to target humans; and (2) imposing strict constraints on the design and use of all other AWS.

The other two areas of concern are with (1) the application of AI to the development of cyber capabilities – including information operations –, and (2) its use to support decision-making in armed conflict, enabling widespread collection and analysis of data sources, for example to identify people or objects, assess patterns of life or behavior, make recommendations for military strategy or operations, or make predictions about future actions or situations.

The ways in which AI and ML might be used for humanitarian action, including by the ICRC, are likely to be very broad. They include tools for environment scanning, monitoring and analysis of data sources and applications that help inform needs assessments, such as the type and location of assistance needed (food, water, shelter, economic, health). These applications also bring potential risks, as well as legal and ethical questions with respect to impartiality, data protection, human rights, transparency, accountability and ensuring human involvement in decisions with significant consequences for people's lives and livelihoods. Any applications for humanitarian action must be designed and used under the principle of "do no harm" in the digital environment and preserve and support the ability of humanitarian actors to operate in line with the fundamental principles of humanity, impartiality, neutrality and independence.

### Key commitments/Pledges/Actions

On AWS, for legal, humanitarian, and ethical reasons, States should adopt new, binding international law rules that:

- 1. prohibit AWS that are designed or used:
  - such that their effects are unpredictable; and/or
  - to target humans; and
- 2. regulate all other AWS through a combination of:
  - limits on the types of targets, such as constraining them to objects that are military objectives by nature;
  - limits on the duration, geographical scope, and scale of use, including to enable human judgement and control in relation to a specific attack;
  - limits on situations of use, such as constraining them to situations where civilians or civilian objects are not present; and
  - requirements for human-machine interaction, notably to ensure effective human supervision, and timely intervention and deactivation.

Where AI is used to support decision making in armed conflict, human judgement must be preserved, especially for decisions that pose risks to the life and dignity of persons affected by armed conflict and other

situations of violence. Accordingly, responsible military applications of AI must support and facilitate human decision-making, rather than replacing it. This will require ensuring that the system:

- provides decision makers with the type, quality, and quantity of information that, in practice, facilitates and improves situational awareness;
- afford the decision-maker the time required for deliberation and the practical possibility to exercise human agency in a contextual manner; and
- enables them to adequately account for the limitations and drawbacks inherent in AI-enabled systems and the challenges of human interaction with them.

END