HUMANITARIAN LAW & POLICY

What we don't understand about digital risks in armed conflict and what to do about it

July 26, 2023, Analysis / Artificial Intelligence and Armed Conflict / Cyber / Humanitarian Action / New Technologies / Technology in Humanitarian Action

15 mins read

Joelle Rizk
Digital Risks Adviser, ICRC

Sean Cordey
Sean Cordey, Digital Risk Researcher, ICRC

The deployment and use of new digital technologies in modern conflicts – from information to cyber operations – creates new risks and enables actual risks of harm to civilians' rights, lives, safety, dignity, and resilience. Understanding these risks is at the core of protection work in the digital age.

In this post, ICRC Digital Risk Adviser Joelle Rizk and Digital Risk Researcher Sean Cordey reflect on some key protection concerns in the digital age and lay out the way forward for protection actors to improve their preparedness to address these.

Humanitarian protection is defined by the efforts of humanitarian actors in times of armed conflict and other situations of violence to safeguard the lives, safety, and dignity of civilians. To do so, humanitarians and other actors engage in protection activities that aim to ensure that authorities and other actors respect their obligations and the rights of individuals *in accordance with the letter and spirit of the relevant bodies of law*. Such activities aim to prevent or put a *stop to actual or potential violations* and address their consequences. Protection activities inherently seek to reduce exposure to risks and reduce vulnerabilities through technical and humanitarian assistance, by supporting self-protection measures, risk education, and providing adequate accurate information, etc. Protection work requires a continuous analysis of risks people face in such situations.

Protection activities constantly adjust to the changing realities of conflict, including the deployment of new technologies that shape warfare. The use of new and digital technologies by different actors in armed conflict settings – be them states, non-state armed actors, criminal groups, or private companies (hereafter referred to as conflict actors) – to conduct cyber and digital operations is one of the most important contemporary evolutions in

armed conflict. While cyber and digital operations *rarely exist in a vacuum*, the intended and expected *circumstances* of the use of digital technologies may represent an array of risks – we call them digital risks – to civilian populations' lives, safety, dignity, and resilience. These are harmful consequences that often come in addition to the suffering they face as a result of kinetic operations.

Humanitarian response is yet to understand the weight of digital risks to civilians in conflict. Documenting, assessing, and further understanding the uses and harms of new digital technologies in both the physical and digital environments is thus critical. Currently, digital risks can be separated into three broad categories, of which this article will focus on the latter two:

- First, those related to the use by humanitarian actors of digital technology to support humanitarian and protection activities, such as the use of *biometrics*;
- Second, those related to the use by armed actors in support or independently of kinetic operations, such as information operations, cyber operations against civilians or *civilian infrastructures*, or the *misuse of personal or humanitarian data*;
- And third, those related to repurposing or dual-use technologies and infrastructure and allow for the *involvement of* civilians in conflict-related actions, such as for *surveillance*, intelligence collection or cyber and information operations.

These specific behaviors and digital technologies may transversally affect and restrict specific rights, such as *freedom of expression*, *assembly*, or *movement*, *liberty and security, personal identity*, and *privacy*. However, of particular concern to the ICRC in conflict settings are those with harmful consequences to the life, safety, and physical and psychological integrity of affected populations – their dignity; their ability to protect themselves and resilience; their economic livelihood; and their access to essential and humanitarian services.

As protection and humanitarian actors increasingly look towards new digital technologies to support their humanitarian and protection activities and reinforce the agency of people affected by conflict, they aspire to *do no harm*. This is especially so when the deployment of digital humanitarian technologies in conflict settings can foster and exacerbate *risks* that can, in turn, undermine civilians' fundamental rights, and their trust in humanitarians, as well as lead to *various types of harms*. This article focuses on the behavior and uses by conflict actors, not the behavior of protection actors and their use of digital technologies.

'Protection' in the digital age

Without undermining the positive impact technology can bring in conflict, including enhancing access to life-saving information and potentially minimizing collateral damage, protection work must consider the risks in the digital age. In other terms, it must encompass the protection of the rights of people when their lives intersect with the digital sphere. For instance, under international humanitarian law (IHL), civilians and civilian objects must not be the target of attack during armed conflict – an obligation equally applicable to cyber and digital operations.

Digital risks in a protection context may thus relate to the protection of data or other digital assets but aren't limited to these issues. They revolve around the use of digital technologies in contexts of armed conflicts and the way their application exposes civilians to harm, affects their rights, safety, and dignity (such as the use of *spyware against civilians*), including when the abuse or violation occurs exclusively online (such as *hate speech*). That is, any risk mediated or enhanced by digital technologies, whether it is physical (incl. supporting infrastructures like satellites), logical, or informational. In other terms, the scope of protection work should encompass behaviors and violations that are committed through actions between humans, between humans and machines, and between machines (such as cyberattacks *targeting civilian or dual-use infrastructure*).

Risk exposure and protection concerns

Protection in the digital age does not necessarily translate to fundamentally novel protection concerns. However, an important distinction is that digitally related protection concerns may be less visible, tangible, understood (especially by the people affected), and reported. In addition, due to the wide attack potential and the prevalence of vulnerabilities, digital threats may *actually scale up fast* and have a wide reach. They may also evolve as digital technologies and practice progress, potentially fostering new, unforeseen risks that protection actors will have to monitor.

Harmful information online

The spread of harmful information such as misinformation, disinformation and hate speech (MDH) can *fuel conflict* and *compromise people's* safety and dignity. Online information and media platforms have amplified the scale, reach and speed of the spread of MDH. Information communication systems are leveraged by *states* and non-state actors to exert influence, change behavior or achieve operational objectives. In that space, information narratives can contribute to or incite acts of violence against *people*, cause targeted distress and lasting *psychological harm*, further increase vulnerabilities due to discrimination, stigmatization, and denial of access to essential services, compromise situational awareness and self-protection measures, and disrupt or undermine *protection actors* and their *operations*. This risk only becomes more exacerbated as AI-generated content becomes more accessible. At the same time, digital communication tools have been used in ways that may violate specific rights and obligations – such as their use to spread harmful information in violations of the *prohibition on child recruitment* or the *prohibition of exposing prisoners of war* to *public curiosity*.

Cyber activities targeting civilians

Civilians are also the direct targets of cyber activities that risk being detrimental to their well-being and undermining their rights. The deployment of *spyware* targeting civilians, for instance, can enable the misuse of personal data to the *detriment of individuals* and, potentially, affect the *broader conflict*. Meanwhile, populations already vulnerable due to conflicts, such as conflict refugees and others *displaced*, might be targeted online by criminals and other malicious actors, leading to concerns around identity theft, fraud, or *scams*.

Cyber operations against civilian infrastructure

Conflict actors leverage cyber means, such as *ransomware, DDoS, or wipers*, to impact and disable civilian infrastructure and essential services, such as electricity, water or medical, e-governance, and financial services. Such operations can have concerning *human costs* and potentially *devastating* humanitarian consequences, affecting the effective delivery of essential services to crisis-affected populations and therefore potentially causing socio-economic, societal, and *psychological* harms or even death. Civilians might also be incidentally harmed when cyber operations affect dual use infrastructure, such as satellites.

Data misuse and mishandling

The deployment and use of data-driven technologies such as sensors, predictive analytics, or *biometric data processing*, raises an array of concerns about the rights, safety and dignity of crisis-affected populations. For instance, intercepted humanitarian data, such as via access requests to third-party providers, *hacking*, or *leaks*, can be misused for non-humanitarian purposes, such as in *law enforcement*, arrest operations, and border screenings. Meanwhile, affected populations' private, personal, and identifiable data, including those linked to their own use of digital technologies (e.g., social media), can be leveraged to identify and target them directly (e.g., disinformation, scams, or violence).

Data, AI, and decision making

Conflict actors are integrating AI-enabled automated "decision-support systems" in their conduct of warfare. These are software tools that provide *analyses*, *recommendations*, and even *predictions* for military decision-makers. These could be used in a wide range of military decisions at all levels of command, such as in "threat assessment" and target recognition, decisions on how to conduct a specific military operation or other decisions that impact people's rights, such as detention. Their use raises concerns not only regarding how to ensure human legal judgment and intervention but also that users are able to explain, challenge, and not overly rely on these AI-based systems. Other concerns exist as per these systems' transparency, potential bias and errors, their indiscriminate targeting, but also harm due to disproportionate attacks with potential consequences to the life and dignity of people and their rights.

Disruption of humanitarian operations

Humanitarian operations are increasingly *disrupted* by digital means, whether through information campaigns that target their integrity and neutrality or through *cyber operations* and *data breaches*. This can impact the capacity of humanitarian actors to operate, access affected populations, coordinate with other actors, assess needs, and provide aid to affected populations. It can also have a detrimental impact on people's safety and their *trust* in humanitarian actors and operations. Furthermore, it *endangers humanitarian and aid workers*.

Disruption of people's connectivity

The disruption of access to internet and communication infrastructure is an increasingly used *practice* by conflict actors to control the information environments and/or support political or military objectives. Such shutdowns can create or exacerbate humanitarian consequences for those on the ground, with potentially life-threatening consequences. For instance, they not only limit crisis-affected populations' access to life-saving information (e.g., humanitarians, food, shelter, *healthcare*) but might also increase the *risk of separation* based on the importance of connectivity to maintain and restore family connection. Meanwhile, they may also impede on civilian's resilience and risk awareness in situations of conflict, as well as their ability to protect themselves, to leverage economic opportunities, and speak and assemble freely.

Civilian involvement

The growing involvement of civilians and *private companies* in activities on the digital battlefield puts individuals at risk of harm and further *diminishes the distinction* between civilians and combatants. Indeed, civilians could actively support conflict actors, whether by getting involved in military *intelligence collection* (e.g. via repurposed apps), supporting the cyber defense of one belligerent, or engaging in *cyber operations* against enemy targets including against other civilian targets. Such involvement can expose civilians to serious harm, such as being targeted by militaries, their *property destroyed, detained*, or even *killed*. It can also be the cause of false accusation and suspicions that lead to further harm.

Preparedness and protection work in the digital age

As protection concerns in the digital age continue to arise, humanitarian actors still have a long way to go to unpack the limits and risks of digital technologies. The interplay of the online and offline aspects of conflicts and the resulting humanitarian consequences will require humanitarians to adapt their skills, methods, and approaches in several ways:

1. Enable protective frameworks and dialogue

Wars have limits, including in the digital sphere. Protection work in the digital age must therefore work with and expand – where possible – these protective frameworks to protect the rights, safety, and dignity of conflict-affected individuals: whether by furthering their development, raising awareness to them, or engaging and advocating with states for their implementation.

Non-state actors such as tech companies and cybergroups have emerged as stakeholders in armed conflict and operations which further exacerbate threats to civilians and other protected persons. Dialogue with relevant actors in that space should be encouraged. Issues to be addressed include governance, prevention of harm and incidental damage, application of IHL and the distinction between civilian and non-civilian targets, collaboration in delivering or enhancing protection activities, principled and human-centric technologies, etc. These could rest on and leverage existing frameworks, such as the *UN's Guiding Principles on Business and Human Rights*. At the same time, dialogues with states could recall their legal obligations to ensure that *private companies respect the relevant rules of IHL* and international human rights law.

Finally, and in tandem, as humanitarians collect personal and sensitive data, they need to integrate data-protection practices and frameworks into their work. Such practices include data minimization, data protection impact assessments, data protection by design and consideration of data subject's

rights. Commendably, considerable work has been done in the past years on responsible use of technology and data in humanitarian contexts, such as the *ICRC's Handbook on Data Protection*, the *Protection Information Management* (PIM), or the *Professional Standards for Protection Work*.

2. Foster resilience

There is an opportunity in infusing existing protection work with digital literacy and risk awareness programs and trainings on digital risks for both affected populations and humanitarian practitioners. These efforts must not, however, come at the cost of pushing the responsibility onto affected populations. As more organizations and private actors adopt tech-based solutions and create opportunities for digital literacy and awareness, it is important that humanitarian and protection actors adopt a careful approach that recognizes the risks of the exclusion of communities or groups, creating a *false sense of security*, and more importantly shifting the responsibility onto the most vulnerable.

3. Build capacity

Risk awareness across the humanitarian sector remains fragmented. There is a considerable *gap* in the understanding and documentation of the digital threat landscape and associated risks for affected populations and behaviors of various actors in conflict environments. As such, digital risk assessment instruments should be developed and integrated into protection work. Accordingly, humanitarians will have to *continue* to strengthen their collaboration with the academic, military and tech experts to produce timely and comprehensive evidence-based protection analysis and response.

Moreover, to better detect, assess, and mitigate digital risks, protection workers will have to be upskilled to rely on hybrid approaches that merge traditional approaches with new means. This notably includes further leveraging and mainstreaming open-source information and social media analysis, which can both provide greater visibility and evidence to inform protection work, from incident monitoring that can inform protection dialogues to tailored community-based protection and engagement. Protection workers should also be trained and supported to be able to be aware of and document the consequences of cyber and digital operations in conflict. This includes engaging with communities whose lives intersect with digital technologies.

***

The rise of *digital risks* stemming from or exacerbated by harmful information online, cyber defense operations, the automation of military systems, misuse of personal and humanitarian data, connectivity shutdowns, or the increased involvement of civilians in conflict through digital means is a reality of conflict in the digital age. The related effect on the rights, safety, dignity, and resilience of conflict-affected populations is a concern that cannot be ignored.

While important challenges arise, such as digital literacy, risk awareness and the capacity to generate evidence of harms, protection actors should work towards expanding on the existing protective legal and policy frameworks (including data protection ones); engaging in protective dialogue on digital risks; fostering resilience of affected populations, such as via risk education and awareness; and building their own expertise and capacity to detect risks and prevent or address resulting harms.

In this rapidly evolving digital environment, the preservation of the humanitarian space and a protection-centered approach is key. While humanitarians continue to unpack what this means for their respective actions, it is important not to reinvent the humanitarian wheel but adapt existing programs. Engaging on digital risks is not 'nice to have' but is instead an ethical and professional imperative for humanitarians.

See also:

- Pierrick Devidal, "*'Back to basics' with a digital twist: humanitarian principles and dilemmas in the digital age*", February 2, 2023
- Tilman Rodenhäuser, Mauro Vignati, "*Towards a 'digital emblem'? Five questions on law, tech, and policy*", November 3, 2022
- Tilman Rodenhäuser,Balthasar Staehelin, Massimo Marelli, "*Safeguarding humanitarian organizations from digital threats*", October 13, 2022

Tags: AI, armed conflict, artificial intelligence, cyber, cyberwarfare, digital, digitalrisk, humanitarian action, IHL, protection

You may also be interested in:

Persons with disabilities in armed conflict

Towards a disability-inclusive IHL: ICRC views and recommendations

27 mins read

Analysis / Artificial Intelligence and Armed Conflict / Cyber / Humanitarian Action / New Technologies / Technology in Humanitarian Action

Elizabeth Rushing, Nawaf Kabbara, Veronica Ngum Ndi & NG'AA Michael Mwendwa

Persons with disabilities constitute approximately fifteen percent of the global population – a figure that …

10 mins read

Analysis / Artificial Intelligence and Armed Conflict / Cyber / Humanitarian Action / New Technologies / Technology in Humanitarian Action

Alexander Breitegger

As a 2022 global report by the World Health Organization estimates, about 1.3 billion people …