Symposium on Cybersecurity & Data Protection in Humanitarian Action

23 – 25 January, 2024
Luxembourg

Symposium briefing papers

DigitHarium

Symposium on Cybersecurity & Data Protection in Humanitarian Action
23 – 25 January, 2024
Luxembourg

TABLE OF CONTENTS:

Working Session #1

NEUTRALITY, IMPARTIALITY, AND INDEPENDENCE IN SOFTWARE TOOLS: FREE AND OPEN-SOURCE SOFTWARE IN HUMANITARIAN ACTION

Objectives of the Working Session:

- Assess the technical, infrastructural and organizational capacities that humanitarian organizations need to have to positively leverage Free and open-source solutions.
- Explore challenges and concrete recommendations on how humanitarian organizations can involve communities of experts; as well as the risks coming along the reliance on these communities.
- Map the stakeholders that humanitarian organizations need to engage in the context of Open-source software, and the strategies to do so in a continuous and sustainable way.

Background information:

Humanitarian action and software products

Humanitarian organizations increasingly rely on digital tools, not only to communicate and to coordinate their daily activities but also to offer digital services directly to affected communities. They depend on a range of software products often developed and commercialized by tech companies. This growing dependence can make them vulnerable to the politicizing and polarizing trends of cyberspace. These trends include allegations and perceptions of participation – intentional or unintentional – in implementing national security priorities and surveillance; business model based on data exploration or profiling; explicit support for some parties to conflict and/or boycotting of others; and participation in global competition over digital infrastructure and standard setting.

The geopolitical tensions that have infiltrated the tech sector may affect humanitarian organizations that use the tools provided by tech companies. More specifically, by using such tools to manage their growing digital footprint, humanitarian organizations may be endangering the perception – and the acceptance based on that perception – that they are neutral, impartial and independence organizations capable of acting on exclusively humanitarian grounds. In this context, Free and open-source software (FOSS) is seen as a possible alternative to address these issues.

Open-source software: What is it?

Open-source software (OSS) refers to computer software for which the source code is publicly available. This means that the underlying code of the software can be freely accessed, modified, and redistributed by the community. A related concept is that of Free

and open-source software (FOSS). Users have certain freedoms with these solutions, including using, modifying and distributing the software, or part of it, in its original or derived form. Software can be considered FOSS if it adheres to a specific set of licenses (such as those of the Open-Source Initiative), which ensure certain rights and freedoms to users.

Open-source played a crucial role in the early days of computing when developers needed to collaborate and share information. It also provided an alternative to proprietary software, which was often expensive and restricted in terms of customization and distribution. The use of Open-source saw a decline in the '90s, as commercial software companies were gaining more and more dominance. These business-driven companies often saw Open-source as a threat to their business model, so they promoted proprietary tools. At the same time, Open-source solutions lacked user-friendly interfaces, which made them less appealing to non-technical users. These tools emerged again around the 2010s, as some projects (e.g. Linux) matured and improved, offering high-quality alternatives to proprietary software. These solutions were seen as attractive alternatives in terms of cost efficiency. In addition, with the growing concerns about cybersecurity, the transparency of OSS became an important advantage.

Today, more than 90% of all code bases incorporate OSS. Around the world, there are millions of experts in Open-source tools contributing, at times on a voluntary basis, to developing and maintaining these solutions. Several governments have started looking into Open-source, and some have even classified it as critical economic and security infrastructure. Similarly, several organizations, such as NASA, Amazon, Google, and Spotify, have also played a role in developing and adopting them.

Open-source solutions for humanitarian organizations

Numerous features of Open-source software resonate with the core values and principles driving humanitarian action. These solutions effectively diminish the reliance on specific vendors, promoting openness and eliminating the risk of getting locked into a single provider/vendor. Furthermore, by embracing transparent and community-driven technology, humanitarian organizations can demonstrate their dedication to maintaining their neutrality and transparency. Moreover, humanitarian organizations operate on a global scale, making it essential for them to have access to software solutions that cater to their unique and specific needs. However, off-the-shelf proprietary software often falls short in meeting these requirements. OSS, on the other hand, offers the flexibility necessary to effectively address these specific needs.

Furthermore, OSS grants humanitarian organizations greater control and ownership over their software systems, which decreases their reliance on specific vendors and providers. This is particularly beneficial as it eliminates the risk of being subject to the control of vendors in terms of updates and support. By opting for OSS, humanitarian organizations can avoid exclusive partnerships with vendors, ensuring their ability to remain independent and act in the best interest of those they aim to help.

Besides that, another interesting aspect of OSS is its transparency. OSS tools are open for examination by a vast community of experts who can detect and address potential security vulnerabilities. Additionally, these experts can collaborate to enhance and develop new features for these tools. By opting for OSS solutions, humanitarian organizations can demonstrate to stakeholders, partners and beneficiaries their commitment to use technology solely for humanitarian purposes.

Open-source solutions also stand out in terms of security and privacy. They undergo global scrutiny for security vulnerabilities and adherence to privacy standards. This transparency could foster trust in the reliability and integrity of the humanitarian digital tools that may be developed.

Not all that glitters is gold

Several are the challenges associated with effectively implementing OSS for humanitarian purposes. For this reason, it is important to thoroughly examine these concerns and gather recommendations on how to move forward. Firstly, OSS relies on Open-source communities, which often consist of technical individuals whose identities may not always be verifiable. Additionally, these communities are not formal entities but rather flexible and dynamic voluntary groups of experts. There may be instances where an Open-source community is not reliable or available to provide technical support and maintenance for a solution. Therefore, sustaining an Open-source project may involve significant costs associated with the community supporting the solutions.

Moreover, the openness and transparency of OSS can be a double-edged sword in terms of security. On one hand, conducting source code audits can increase the overall security of the too. On the other hand, it is possible that vulnerabilities may – willingly or unwillingly – be introduced into the code. Some of the other challenges relate to the costs coming along OSS, as well as the interoperability of the solutions developed. Another issue that humanitarian organizations have to consider is the potential misuse of an Open-source tool developed for humanitarian use. This raises questions of accountability and responsibility, and also challenges a neutral and independent perception of the humanitarian organization and of its humanitarian action.

Guiding questions:

- How can Free and open-source software (FOSS) benefit humanitarian organizations, and what advantages do Open-source communities offer to them?
- What specific challenges hinder the practical implementation of FOSS for humanitarian organizations, and how can these obstacles be effectively addressed? These include the reliability, motivation and trustworthiness of Open-source communities, as well as the costs associated with the retention of Open-source communities. Additionally, how can accountability, responsibility and concerns about potential third-party misuse of humanitarian FOSS be mitigated?

- What are the essential requirements for collaboration between humanitarian organizations and Open-source communities?
- What are the necessary measures to operationalize FOSS for humanitarian action? Should specific policies be put in place, and should legal or non-legal framework be used? What communication strategies should be adopted to facilitate the practical application of FOSS in humanitarian action?

Additional material:

- Ackermann R. (2023). "The future of open source is still very much in flux", in *MIT Technology Review*.
- Parra E., Haiduc S. & James R. (2016). "Making a difference : an overview of humanitarian free and open source systems", in *ISCE '16: Proceedings of the 38th International Conference on Software Engineering Companion*.
- Foss2Serve (2022). *HFOSS Projects*.
- Nyakundi H. & De Souza C. H. (2023). "Fostering FOSS Communities: A guide for Newcomers", in *Business Models and Strategies for Open Source Projects.*

Working Session #2

NEUTRALITY, IMPARTIALITY, AND INDEPENDENCE AND THE CLOUD: ACHIEVING EXCLUSIVE HUMANITARIAN USE OF HUMANITARIAN DATA

Objectives of the Working Session:

- Increase awareness among the stakeholders on the difficulties that humanitarian organizations encounter in ensuring that humanitarian data is used exclusively for humanitarian purposes.
- Assess the advantages and drawbacks of storing data on the cloud.
- Collect practical inputs on the solutions that humanitarian organizations ought to explore to achieve exclusive humanitarian use of humanitarian data.

Background information:

Humanitarian data

As humanitarian organizations' use of digital technology grows, so too does the amount and variety of the data they have to handle. This data encompasses diverse content, sensitivity levels, and sources, including internal information, correspondence with states and other actors, and confidential personal information on members of affected populations. Some of this data, crucial for fulfilling their humanitarian mandate, can be of great sensitivity.

Data collection in humanitarian settings, for organizations carrying out neutral, impartial and independent humanitarian assistance, requires a steadfast commitment to affected communities that their data will be used only for humanitarian purposes. Consequently, regardless of its source, humanitarian organizations must ensure the preservation of the confidentiality, integrity and availability of this data.

Humanitarian organizations in the "humanitarian cloud"?

Organizations have the option to store data either in their own on-premises services or on the cloud, where it is stored in remote servers owned by external entities. Traditionally, data has been stored on-premises, but there is a trend towards storing data on servers that are managed externally. Cloud storage has gained popularity for businesses of all sizes in recent years. As a result, on-premises solutions are becoming less common, and organizations are exploring cloud solutions to complement or even replace their traditional on-premisis tools.

There are several compelling reasons that prompt humanitarian organizations to adopt cloud solutions. Firstly, cloud solutions can reduce operational cost by eliminating the

need for most hardware, software, and in-house experts. Additionally, software updates are often handled by third-party entities managing the cloud, which reduces the need for infrastructural support. Cloud solutions are also seen as more resilient against cybersecurity risks due to automated patching and the resources provided by external entities such as Google, Apple, and Amazon. Another notable advantage of cloud is its scalability and elasticity, which can be invaluable when responding to adverse cyber-attacks. Lastly, the faster time-to-market element of cloud-based tools renders them especially attractive.

However, there are some risks associated with cloud technologies. From a strategic standpoint, switching from one provider to another can be complex, and result in customers being locked-in to working with a specific cloud supplier. Additionally, in many cases, only the external provider has access to the cloud infrastructure, which means that customers are unable to verify and control how their data is effectively managed. This is further exacerbated by the fact that cloud providers can unilaterally make changes to features and costs, taking advantage of the difficulty in switching providers. In addition, the use of cloud-based tools involves third-party tech companies, many of which then become effective intermediaries in humanitarian action. These companies have no humanitarian mandate and do not enjoy the privileges and immunities that would enable them to fulfil such a mandate in a neutral, impartial and independent manner. They are also vulnerable to exercise of jurisdictional authority by the countries in which they operate and cannot guarantee that data will not be accessed by governments for purposes that are not exclusively humanitarian. This is an important consideration, because a number of different actors have an interest in gaining access to humanitarian data for reasons unrelated to humanitarian action – for instance, to gain a strategic and operational advantage over adversaries.

Humanitarian organizations are not the only actors facing this situation: challenges related to ensuring exclusive jurisdictional authority over data are also of concern to states, in connection with data related to their functions and their citizens. Initiatives are therefore being proposed at state and regional levels to regulate the technology and services used in a digital infrastructure: these initiatives often refer to the notion of "digital sovereignty." In this context, this session focuses on what international humanitarian organizations can do – when employing the services of third-party providers – to ensure that data under their control are used for exclusively humanitarian purposes. Technical solutions – like "trusted execution environments" and homomorphic encryption – need to be closely examined, as well as organizational measures and legal ones, such as ensuring the reliability of contractual clauses and the precise wording of the privileges and immunities of international organizations concerned.

Guiding questions:

- What are the benefits that humanitarian organizations receive from storing sensitive data on remote servers owned by external entities? What are the potential risks and

vulnerabilities associated with this approach? Is there a turning back from cloud solutions?

- When using "the cloud," how can humanitarian organizations ensure the selection of a provider in a way that the principles of neutrality, independence and impartiality are upheld?
- How can exclusive humanitarian control of humanitarian data be legally safeguarded?
- Should humanitarian organizations segregate humanitarian and non-humanitarian data when relying on cloud solutions? If so, how can it be achieved and maintained?
- What specific criteria should guide humanitarian organizations in making decisions about their storing systems? How to balance factors such as reliability, stability, resilience, security, and the capacity to ensure exclusive humanitarian use of humanitarian data?

Additional material:

- Kuner C. & Marelli M. (eds.) (2020). "Cloud Services", in *Handbook on Data Protection in Humanitarian Action*. ICRC.
- CERN. (2023). *Policy for Cloud Usage at CERN*.
- Sabt M., Achemlal M. & Bouabdallah A. (2015). "Trusted Execution Environment: What it is and What it is not", in *2015 IEEE TrustCom, Big Data*.
- Peng Y., Zhou W., Zhu X., Wu Y. & Wen S. (2022). "On the security of fully homomorphic encryption for data privacy in Internet of Things", in *Concurrency and Computation: Practice and Experience*. Vol. 35, n° 19.

Working Session #3

NAVIGATING TRUST AND SAFETY AS TECH COMPANIES DURING ARMED CONFLICT

Objectives of the Working Session:

- Provoke honest, open conversations about the ways that ICT companies' business activities can have an impact and be impacted in the context of armed conflict. This session is intended to help ICT companies, humanitarian organizations and other stakeholders improve their understanding of conflict-related risks and improve their ability to make responsible decisions to improve the protection of people in conflict settings.
- Identify common risk scenarios and/or pain points; reach agreement on potential next steps, (such as best practice for designing for conflict affected people; creation of a toolkit for safety measures for ICT companies and civil society organizations working in armed conflict settings)

Background information:

The acceleration of the global digital transformation is profoundly affecting society, business, and governance. This transformation generates unique dynamics and effects in armed conflicts. Digital technologies have vital military and civilian functions in such contexts, at the same time they can also be misused to harm civilians.

Industry, civil society, academic and international organizations have taken specific actions, conducted research, and provided recommendations, tools, and frameworks, such as the Office of the High Commissioner for Human Rights B-Tech Project, and Access Now's Declaration of Principles for Content and Platform Governance in Times of Crisis. The private sector, likewise, has developed frameworks and initiatives, including developing trust and safety (also known as integrity) teams and policies, as well as Corporate Social Responsibility (CSR) and Environment, Social, and Governance (ESG).

While conflict is on the map of digital technology companies – with many taking important conflict decisions in the wake of the US withdrawal from Afghanistan and the Russia-Ukraine armed conflict – it is still on the margins. Many companies tend to be very familiar with International Human Rights Law (IHRL) and the UN's Guiding Principles for Business and Human Rights (UNGPs) – and the UNGPs make explicit reference to conflict-affected areas and International Humanitarian Law (IHL, also known as the Law of Armed Conflict) – but tend to be much less familiar with IHL and the risks of operating in armed conflict settings.

There are many reasons why IHL and conflict sensitivity are important for ICT companies (and the private sector more broadly). Among these is the fact that the number of conflict-affected contexts – and beyond that, fragile states where risk of conflict and widespread violence is elevated – is high: more than 100 globally. In addition, armed conflict can be sudden, large in scale, and cause significant suffering and destruction – resulting many difficult decisions, often very quickly. Such decisions may range from whether to continue, modify, or stop operations; how to manage ***risks to people***, populations, and society (e.g., clients, users of products/services, communities of users, and broader society); and how to manage risks to the company itself (e.g., its plants, equipment, and personnel that are in the territory of an armed conflict, as well as its global reputation).

IHL provides a legal framework that regulates how parties to armed conflict conduct warfare and provides protections to civilians and others not taking part in hostilities. As the UNGPs reference, IHL reflects the unique realities of warfare and the vulnerabilities of people living in conflict. Unlike some other areas of international law, IHL can directly apply to private individuals, including managers or personnel of private companies. IHL can help inform tech companies operating in, or that may in the future operate in, situations of armed conflict.

As noted in the UNGPs, risks and impacts (to people in general, and especially to vulnerable groups) are heightened during conflict. IHL protects civilians generally and contains specific protections for certain categories of persons – such as the war wounded, prisoners of war, and medical as well as humanitarian workers. The history of why these rules came into being can be as important as the specific rule and can help inform practices that are not only compliant with the law, but help achieve the aim of the law.

The number and complexity of life and death decisions people living in armed conflict contexts make, often based on limited, imperfect, and contradictory information, can be staggering. Moreover, the nature of unstable situations can lead people to rationally prioritize short-term over long-term security, and survival instinct may lead to decision-making that may look irrational from the outside (and/or in the long-run), but makes sense in the moment to people in that context.

Companies have experience in developing personas (or avatars) and red-teaming to address identified issues – and it could be that using these approaches on the basis of a conflict affected person persona could yield important insights. For instance, how ICT technologies – such as hardware (e.g. devices), infrastructure services (e.g. connectivity and cloud services), and/or social media - are designed and deployed in such contexts ***could help civilians in armed conflicts navigate these decisions***, to insulate them from noise/scams, and to help keep them safe.

As an example relevant to some ICT companies, IHL provides relevant guidance with regard to the spread of harmful information online. While IHL does not expressly prohibit propaganda/disinformation, it does prohibit certain harmful activities by belligerents regardless of whether they occur on- or offline. Such prohibitions include encouragement of IHL violations (for example, encouraging attacks on civilians or torturing prisoners); threatening violence to spread terror among the civilian population; undermining the work of medical or humanitarian actors; and exposing prisoners of war to insults and public curiosity. Where ICT products and services are used to engage in such conduct, those products and and services are, in effect, being used to commit IHL violations.

The implications that IHL may have for risks to the company itself, in particular in regards to its workers and properties, has typically garnered less attention. When considering these risks, a critical consideration for companies should be the legal paradigm shift from peacetime to wartime. During situations of armed conflict, tech company personnel and assets generally qualify as "civilian." As such, they are legally protected from being the object of an attack by a belligerent. In exceptional circumstances, normal operations may suddenly expose those personnel and assets to risk of attack (above and beyond the risk that any entity faces in war by being in the wrong place at the wrong time). This may happen in such circumstances as when a tech company employee "directly participates in hostilities" (DPH) or when a piece of company property qualifies as a "military objective."

As an example that is relevant to the ICT sector, in recent years, public-private partnerships in cybersecurity have become increasingly common. In peacetime, the overall benefits of such arrangements may be high, and the risks relatively low (though still present, including in relation to issues such as privacy and data protection). Whereas in an armed conflict, depending on the type of partnership, a company may want to assess whether such partnerships might cause its plants, equipment, or personnel to lose that protection.

Finally, companies also should be aware that while IHL has strict obligations aimed at avoiding and minimizing incidental civilian harm, it does not provide a legal rule that outright prohibits those harms unless they are excessive. This legal rule (also referred to as the principle of proportionality) reflects the popularized concept of "acceptable" collateral damage. This is an important consideration to keep in mind, especially when companies think about how they might minimize exposing themselves and others to "digital crossfire." It is also why input and advice from agencies and actors well-versed in these issues can be valuable for ICT companies.

As digital technologies becoming increasingly ubiquitous in situations of armed conflict, there are a number of humanitarian (as well as legal) considerations ICT companies should explore. People living in armed conflict are having to make frequent life and death decisions in a poor information quality context (lack of good information, a glut of

misinformation and disinformation, lack of good cyber security), where the "right" thing to do from a digital literacy point of view may be the "wrong" thing in the short-term for the safety of you and your family. Understanding these contexts, and the situation of such people, and using this knowledge to make design improvements for people using these products, could help keep more people safe.

Guiding questions:

- How do companies evaluate the risks associated with operating in situations of armed conflict, especially those affecting people and communities in such situations, including clients/users?
    - What policies, tools, processes etc are in place?
- How do companies navigate the current frameworks, best practices or codes of conduct that already exist? How do these (if they do) differ from due diligence?
- How can companies and civil society organizations form a community of practice to tackle these issues effectively?

Additional material:

- Office of the United Nations High Commissioner for Human Rights (OHCHR). "Guiding Principles on Business and Human Rights." Geneva, Switzerland: OHCHR, January 1, 2012. https://www.ohchr.org/en/publications/reference-publications/guiding-principles-business-and-human-rights.
- Office of the United Nations High Commissioner for Human Rights (OHCHR). "B-Tech Project." OHCHR. https://www.ohchr.org/en/business-and-human-rights/b-tech-project.
- Demeyere, Bruno, ed. *Digital Technologies and War*. Vol. 102. International Review of the Red Cross 913. Geneva, Switzerland, 2021. https://international-review.icrc.org/reviews/irrc-no-913-digital-technologies-and-war.
- JustPeace Labs, and BSR. "Conflict-Sensitive Human Rights Due Diligence for ICT Companies." BSR, December 7, 2022. https://www.bsr.org/en/reports/conflict-sensitive-human-rights-due-diligence-for-ict-companies-guidelines-and-toolkit-for-corporate-human-rights-practitioners.
- Carrillo, Arturo J. "Between a Rock and a Hard Place? ICT Companies, Armed Conflict, and International Law." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, August 30, 2022. https://papers.ssrn.com/abstract=4205028.
- ICRC. "The Use of ICTs in Armed Conflicts Poses a Real Risk of Harm to Civilians and Civilian Infrastructure." Statement. International Committee of the Red Cross, July 28, 2022. https://www.icrc.org/en/document/icts-armed-conflicts-risk-harm-civilians-civilian-infrastructure.
- Investor Alliance for Human Rights. "Sector Wide Risk Assessment: ICT; Salient Issue Briefing: Conflict & Security." Investor Alliance for Human Rights and Heartland Initiative, 2020.

https://investorsforhumanrights.org/sites/default/files/attachments/2020-03/Investor%20Alliance_Salient%20Issue_Conflict%20Security.pdf.

- Stauffacher, Daniel, William Drake, Paul Currion, and Julia Steinberger. *Information and Communication Technology for Peace: The Role of ICT in Preventing, Responding to and Recovering from Conflict*. ICT4Peace. New York, USA: United Nations Information and Communication Technologies Task Force, 2005. https://ict4peace.org/activities/information-and-communication-technology-for-peace-the-role-of-ict-in-preventing-responding-to-and-recovering-from-conflict-stauffacher-drake-currion-steinberger-2005/.

Working Session #4

CIVILIANIZATION OF CONFLICT THROUGH CYBER MEANS

Objectives of the Working Session:

- This session will explore the how the principles of international humanitarian law (IHL) – a body of law that sets important limits on the conduct of parties to armed conflict – can address the worrying trend of "civilianization of conflict," in which civilians become involved in armed conflicts through digital means.
- Likewise, the session will discuss strategies for international humanitarian organizations, states, and civil society to mitigate the potentially harmful consequences of this trend.

Background information:

As digital technology changes how militaries conduct war, a worrying trend has emerged in which a growing number of civilians become involved in armed conflicts through digital means. This development is putting renewed pressure on established international legal principles that constrain what is permissible during war. Such principles – which, according to a consensus in the international community, include the principles of humanity, necessity, distinction, and proportionality are part of IHL.

The principle of distinction requires that parties to an armed conflict must at all times distinguish between civilians and combatants and between civilian objects and military objectives. While in the physical world the difference is normally readily apparent, the cyber and digital environment has added new complexities to the issue.

The digitalization of society has fundamentally shifted the role of civilian involvement in conflicts in both quality and quantity. The main qualitative shift is that these activities are now much closer to the actual conduct of military operations: civilian involvement has moved from the production or provision of food, shelter, or equipment at some distance from the physical battlefield to the direct contribution of operations on the digital battlefield as support to kinetic operations.

The main qualitative shift is that, in the digital space, it is much easier to scale civilian activity in conflicts, as groups comprising thousands or even tens of thousands of individuals may be formed and coordinated online in an matter of hours. Similarly, the attack surface of societies has vastly increased. Digital devices, apps, and networks exist almost everywhere, which means that in times of armed conflict, there are exponentially more vulnerabilities than in the wars of the past. These developments have major implications that can be broken down into roughly three main areas:

Civilians as digital warriors. Civilians may replicate everyday activities (downloading and installing an application or other kinds of software, sending messages, clicking buttons on web interface, and so on) with a purpose to actively contribute to the offensive

capabilities of an armed force, for example, by installing and using an application to carry out cyberattacks. It is important to emphasize that computer interfaces (such as screens and keyboards) continue to perpetuate an illusory and artificial separation between the digital and physical worlds, as if the former did not affect the latter.

Civilians as military sensors. Civilians with digital capacities and access can provide actionable information to armed forces, a form of information gathering known as crowdsourcing intelligence. For example, individuals with smartphones can take pictures to report and track the movements of enemy troops. Recent reports indicate that providing this sort of civilian gathered information often has an immediate, real impact, as the information civilians send to their government could be followed by destructive military action.

Civilians as digital victims. Civilian use of digital applications provided by the government can expose civilians to serious harm. For example, there have already been report of militaries targeting civilians and their property for being suspected of using their mobile phones to report the enemy's location. If parties to armed conflict encourage civilians to engage in this type of conduct, such incidents may become much more common, including situations where the civilian in question was using the phone for another reason – for instance to warn their families to leave or seek shelter.

These concerns underscore the need to understand the legal constraints applicable to such forms of civilian involvement on the digital battlefield.

Guiding questions:

- How has the distribution of offensive methods and digital tools for civilians already changed the behavior of civilians during conflict?
- What are the IHL, or other legal implications, of this kind of civilian participation during conflict? Once the boundaries between civilian and combatant have been blurred, how can they then be restored?
- What measures should states, international organizations, and/or corporations take to address this trend?

Additional material:

- ICRC Global Advisory Board on Digital Threats During Armed Conflicts. "Protecting Civilians Against Digital Threats During Armed Conflict." Report, October 18, 2023. https://www.icrc.org/en/document/protecting-civilians-against-digital-threats-during-armed-conflict.
- Wenger, Andreas, and Simon J. A. Mason. "The Civilianization of Armed Conflict: Trends and Implications." *International Review of the Red Cross* 90, no. 872 (December 2008): 835–52. doi:10.1017/S1816383109000277.

Working Session #5

USING OPEN SOURCE INFORMATION ANALYSIS TO DOCUMENT DIGITAL HARM

Objectives of the Working Session:

- Identify information needed to properly document and understand digital risks – how can humanitarian actors effectively detect, observe, and analyze the ways that civilian populations are negatively impacted by digital technologies during armed conflict (including both trends and specific instances)?
- Discuss the challenges of using Open Source Information Analysis in Humanitarian Action, particularly those related to the public disclosure of sensitive information

Background information:

With the increasing digitalization of conflict has come the rise in the importance of digital open source information. Content posted to social media is used by a variety of actors to provide evidence of potential war crimes and to support the accountability process, as well as serving as a crucial source of information for civilians affected by conflict. However, as humanitarian actors have begun to recognize, while digital technologies can play a positive, even life-saving role in armed conflict, it equally presents an array of risks to the civilian populations' lives, safety, dignity, and resilience (see Working Session: Mapping and digital risks and digital harms). This includes exposure to harmful information online, cyber activities targeting civilians, internet shutdowns, and cyber operations against civilian infrastructure.

Despite this recognition, humanitarian organizations are yet to fully understand the weight, scale, or complexity 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. However, in order to better detect, assess, and mitigate digital risks, humanitarian workers, particularly those working in civilian protection, will have to be upskilled in digital methods of information gathering and documentation and will need 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.

As with traditional forms of harm to the civilian population addressed by humanitarian actors, it is clear that an effective approach to documentation and analysis is essential to developing a strong, evidence-based response to digital risks. This, however, requires capacities and methodologies not currently possessed by the humanitarian sector. Additionally, the documentation of identified digital harms presents many challenges – cyber operations and information operations are notoriously difficult to track and

attribute, which is partially why states use them. Also, causality between online events and offline harms can be difficult to establish with confidence.

Guiding questions:

- What open data sources are available to support documentation of digital risks?
- What tools, techniques and methodologies can be applied to the detection and assessment of digital risks? Consider also the roles of different actors, both public and private.
- What are the limitations on humanitarian use of open source data to detect and assess digital risks? Consider technical, ethical and resource limitations.
    - What specific challenges related to Neutral, Impartial, and Independent Humanitarian Action arise when Humanitarian Organizations use Open Source Analysis in their operations?

Additional material:

- Böhm, Isabelle, and Samuel Lolagar. "Open Source Intelligence." *International Cybersecurity Law Review* 2, no. 2 (December 1, 2021): 317–37. doi:10.1365/s43439-021-00042-7.
- Millett, Ed. "Deploying OSINT in Armed Conflict Settings: Law, Ethics, and the Need for a New Theory of Harm." *Humanitarian Law & Policy Blog*, December 5, 2023. https://blogs.icrc.org/law-and-policy/2023/12/05/deploying-osint-in-armed-conflict-settings-law-ethics-theory-of-harm/.
- United Nations, and University of California, Berkeley, eds. *Berkeley Protocol on Digital Open Source Investigations: A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law*. New York; Geneva: [Berkeley, California]: United Nations Human Rights, Office of the High Commissioner; Human Rights Center, UC Berkeley School of Law, 2022. https://www.ohchr.org/en/publications/policy-and-methodological-publications/berkeley-protocol-digital-open-source.

Working Session #6

OPEN SOURCE INFORMATION AND PERSONAL DATA PROTECTION

Objectives of the Working Session:

This working session will delve into the application of personal data protection legal frameworks to the growing utilization of Open Source Information (OSI) in humanitarian action. As humanitarian organizations begin to leverage OSI analysis across various use cases, often updating or implementing new data management processes to do so, it becomes increasingly important to establish a common understanding and concrete approach to protecting personal data that is collected as OSI or that is generated by inference or assumption from OSI techniques.

Concretely, the conversation will start by exploring two data protection principles, ***lawfulness of processing*** and data subject rights, weighing in the fact that OSI is considered *public* information. In a second step, the audience will examine the role of the principle of ***purpose limitation*** and whether a technical approach to this principle could help to maintain data protection standards in the context of OSI.

During the discussion, we will use a humanitarian scenario to evaluate potential solutions: missing persons investigations in natural and man-made disasters.

Background information:

The multi-faceted nature of OSI, defined as information which 'encompasses publicly available information that any member of the public can observe, purchase or request without requiring special legal status or unauthorized access,'[1] often requires organizations to adapt their existing data management processes or even create new ones.

As discussed in Session #5 Using Open Source Information Analysis to Document Digital Harm, humanitarian organizations have begun to leverage OSI analysis to detect, measure and mitigate digital harm. This practice present and array of risks to the lives, safety, dignity, and resilience of civilians. This session will dig into another, and little-examined challenge related to the use of OSI in humanitarian action: the application of personal data protection frameworks.

Publicly available protocols and guidance on OSI mention the importance of adhering to data protection frameworks, but they do not explain those obligations in an actionable level of detail. The Berkeley Protocol on Digital Open-Source Investigations, for example, notes that 'open source investigators should be aware of [the GDPR] and its approach to individual data protection, because this law has set a high standard and other States are considering adopting similar legislation. However, data protection regulations differ from

[1] The Human Rights Center at the University of California B, School of Law, *Berkeley Protocol on Digital Open Source Investigations* (Office of the United Nations High Commissioner for Human Rights (OHCHR), and the Human Rights Center at the University of California, Berkeley, School of Law 2022), p6

country to country, with significant variations and even sometimes directly conflicting rules.'[2] The Protocol then advises investigators to consult a legal expert in their jurisdiction, and personal data protection is not explicitly mentioned again in the document. Similarly, the European Open Source Intelligence Organizations Observatory's (ObSINT) Guidelines for Public Interest OSINT Investigations specify that 'researchers will need to adhere to data protection laws, for example the [GDPR], as applicable in their country. Data collection needs to strike a balance between the public interest purpose of the research and the individuals' fundamental rights.'[3] While the Guidelines do emphasize the importance of personal data protection obligations such as data minimization, data security, and conducting Data Protection Impact Assessments (DPIAs), any unique challenges or considerations related to OSI use cases are not discussed.

As illustrated by these examples, the GDPR is often named as the reference point when it comes to data protection frameworks, but the privilege and immunities of international organizations (IOs) could mean that IOs do not apply domestic legal obligations concerning data protection. Accordingly, multiple IOs have established their own data protection regimes. Since these regimes often mirror the principles and rights set forth in the GDPR, the same general questions remain open for discussion. The diversification of legal frameworks also makes it increasingly important for humanitarian organizations to establish a common understanding and concrete approach to protecting personal data that is also OSI.

The use of OSI does not just present ethical dilemmas or protection risks – depending on the information collected, it also implicates the personal data protection framework(s) to which the organization may be subject. While not all OSI may be or may include personal data, the meta-data about this consent will almost always include personal data. Furthermore, the operations relating to OSI often create new or linked inferences relating to individuals. The below table maps some of the potential considerations and challenges related to data protection frameworks.

Data Protection Consideration	***Discussion***
Legal Basis	What to do when consent is impractical, or not possible (e.g. in investigation cases) what can be don? Certain humanitarian organization have public interest as a possible legal basis.
Transparency and the Right to Information	Who are your data subjects (i.e. not just the 'subject' of a piece of content)? What

[2] Ibid, p28
[3] The European Open Source Intelligence Organisations Observatory, *Guidelines for Public Interest OSINT Investigations*, 2023), p8

	is a reasonable timeframe for provision of information? Exceptions?
Purpose Specification	How to prevent function creep? How to adhere to narrow purpose when you do not know exactly what data you are collecting?
Rights to Rectification and Deletion	Challenges associated with archiving processes? Of linkages between data?
Right to Objection	Exceptions?
Right to Access	How to verify identity of requester? Exceptions?

Guiding questions:

- What is the legal basis for a humanitarian organization to collect OSI data? And how does it relate to the concept of "public" data?
- Which rights of the data subject can be enforced and how? Is there any exception?
- Which are the humanitarian purposes that are compatible with the processing of OSI?
- Would a technical purpose limitation help to maintain data protection standard in the case of a missing person investigation?

Additional material:

- European Open Source Intelligence Organisations Observatory (ObSINT). "Guidelines for Public Interest OSINT Investigations." *ObSINT*. https://obsint.eu/the-european-open-source-intelligence-organisations-observatory/.
- United Nations, and University of California, Berkeley, eds. *Berkeley Protocol on Digital Open Source Investigations: A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law*. New York; Geneva: [Berkeley, California]: United Nations Human Rights, Office of the High Commissioner; Human Rights Center, UC Berkeley School of Law, 2022. https://www.ohchr.org/en/publications/policy-and-methodological-publications/berkeley-protocol-digital-open-source.

Working Session #7

NEUTRALITY, IMPARTIALITY, AND INDEPENDENCE AT THE NETWORK LAYER

Objectives of the Working Session:

- Sensitize the various stakeholders on the challenges that humanitarian action face at the network layer.
- Gather practical orientations on how humanitarian organizations can assert their neutrality, impartiality and independence in their routing of data flows.
- Identify and assess existing and alternative internet infrastructures that humanitarian actors should explore to safeguard their operational capacity in the digital age.

Background information:

The network layer: understanding this dimension

Humanitarian organizations have become more reliant on digital technologies as their use of digital means increases, and as digitalization becomes more pervasive among actors around them. One consequence of this is an increasing reliance on the network layer of the technological infrastructure. This dimension encompasses the information transmitted from one digital system to another, including communication between resources within the same organizations and communications between other services and the end-users of a digital tool. This dimension is essential for humanitarian organizations, as well as for states and other actors, for a variety of reasons. It is related to a transmission of information that is confidential and secured from eavesdropping and interception by malicious actors. Apart from that, the network layer also ensures the data integrity during the transmission, which is crucial for maintaining the information's accuracy and reliability. As humanitarian organizations operate in challenging environments that are spread across the globe, the network layer is responsible for ensuring reliable connectivity between these locations. Along the same line, reliable network connectivity enables the continuity of humanitarian action.

Challenges to a neutral, impartial and independent humanitarian action at the network layer

The reliance on the network layer leads humanitarian organizations to be increasingly affected by dynamics of their digital infrastructure which include global trends such as the emergence of splinternets and the imposition of connectivity restrictions, which pose a significant challenge to the seamless flow of data. This fragments the global network into "splinternets" and further the risk of connectivity denials as well as for censorship – which could halt the operational capacity of humanitarian organizations. In addition, the

pervasive nature of digital surveillance and the looming risks of hostile cyber activities raise serious concerns regarding the confidentiality and integrity of data transmitted by humanitarian organizations. These threats jeopardize the security of sensitive information and may compromise the trust placed in these organizations.

The current infrastructure for internet routing, from a technical standpoint, is de facto the Border Gateway protocol (BGP). This protocol is not transparent regarding the path taken by the data packets. In other words, it remains unclear as to what paths the packets will take at the time of sending the traffic. Thus, it is possible that an unexpected entity on a path can eavesdrop on the communication, unbeknown to the sender and receiver. Although most traffic is encrypted, website fingerprinting or VPN tunnel attacks may nonetheless cause a leak of information. Within the current infrastructure, there is also the problem related to "data flow bottlenecks". The cause for this is due to particular nodes and countries holding pivotal positions in facilitating global internet connectivity, and a certain portion of traffic takes international detours, thus traversing the digital infrastructure of other countries even when this is not intended for them. This element, adding to the potential threat of surveillance and connectivity denial, causes a direct dependence on those particular nodes and their internet infrastructure. This means that the data traffic could traverse undesirable nodes without controlling the path taken. A small number of nodes and countries are central for the global internet reachability, thus meaning that most of the actors – including humanitarian organizations – depend on them for internet communication. From a technical standpoint, given the peculiar operational environments of humanitarian organizations, the availability of the infrastructures for connectivity – especially when it comes to the "last-mile connection" – is an element that is not to be taken for granted as it may be for other actors that are active in areas with well-developed digital infrastructure. As such, for humanitarian organizations, this element directly infringes on their operational capacity to reach a particular humanitarian setting.

What "path" forward?

This session focuses on how humanitarian action is impacted by the dynamics of the digital infrastructure, particularly the network and routing elements, on which it relies. The session seeks to develop an understanding of the threats to the humanitarian organization's ability to safeguard the neutrality, impartiality, independence as well as security of their routing of data flows.

The session pays particularly close attention to international humanitarian organizations, which enjoy privileges and immunities that enable them to fulfil their humanitarian mandate – in full respect of their neutrality, impartiality, and independence – and safeguard the exclusively humanitarian nature of their work. These humanitarian organizations may need to determine how to route their data flows in a way that ensures respect, at all times, for the legal inviolability of their data. This session also aims to discuss concrete means of addressing these challenges. This may include technical solutions available within current internet infrastructure (e.g. peering) and alternatives to

this infrastructure that can ensure a reliable, stable and resilient infrastructure and that incorporate deliberate routing protocols (e.g. Scalability, Control and Isolation on Next-Generation Networks, or SCION). It may also include policy solutions, such as recommendations for states and humanitarian organizations, particularly those having the status of "international organization."

Guiding questions:

- What are the challenges and risks emerging, at the network layer, *vis-à-vis* the delivery of neutral, impartial, and independent humanitarian action?
- How should "neutrality," "impartiality," and "independence" be conceptualized in the context of data flows?
- How should humanitarian organization "choose" its choice of routing? How to find a balance in the following aspects: reliability, stability, security, and resilience of the infrastructure; diffusion of the infrastructural network; capacity to control the data flows; technical capacity (latency and bandwidth) of the route; technical expertise required *vis-à-vis* resource available?
- What possible solutions should humanitarian organizations explore to use within their operations?

Additional material:

- AccessNow (2023). *KeepItOn. An overview of global internet shutdowns*.
- Ganz A., Camellini M., Hine E., Novelli C., Roberts H. & Floridi L. (2023). *Submarine Cables and the Risks to Digital Sovereignty*, paper under review by SSRN.
- Cox P. (2020). "Data in transit", in *ITNOW*. Vol. 61, n° 1: 60-61.
- Kohler K. (2022). *One, Two, or Two Hundred Internets? The Politics of Future Internet Architectures*. Cyberdefence Report. Center for Security Studies, ETH Zurich.
- Samans R. & Lee-Makiyama (2020). *Can governments agree on global data flows? Here are 6 recommendations*. Blog published on June 10, by the World Economic Forum.

Working Session #8

SAFEGUARDING HUMANITARIAN CONNECTIVITY: THE ISSUE OF DUAL-USE SATELLITES

Objectives of the Working Session:

- Explore the role that the different actors, relevant in space systems, can have in enabling a neutral, impartial, independent and meaningful humanitarian action.
- Gather practical inputs on the role that states, international organizations, the international community and the space industry can have in enabling humanitarian responders with uninterrupted access to space systems, especially during emergencies.

Background information:

What are humanitarian organizations doing in the sky?

The frequency and risk of natural hazards is on the rise, as is the number of people who are being displaced because of conflict and persecution. In these – and many other – times, where infrastructure and connectivity may be inadequate, space systems are indispensable for essential civilian services. They support people affected by crises as well as humanitarian organizations delivering humanitarian action. Indeed, satellite communication can be an effective mean to exchange logistic, medical, and situational information. More generally, satellite services contribute to every phase of humanitarian operations for humanitarian organizations, from needs assessment to emergency relief delivery, from disaster risk reduction to resilience building in protracted conflicts. Navigation satellites provide low-cost and accurate real-time tracking for personnel and equipment daily. At the same time, earth observation satellites offer unique information and imagery for emergency mapping, risk assessment, and planning and implementation of humanitarian operations. In addition, satellites have global coverage, making it possible to monitor vast and remote regions across countries and continents consistently. Besides that, satellite data is increasingly available for use soon after it is acquired, enabling humanitarian organizations to receive and transmit the information quickly, and thus act promptly – which is fundamental in humanitarian settings. Another interesting element is that, along with the increase in commercial satellites, there is also an increase in satellites that allow free and open access to data, such as the Copernicus Sentinel missions.

Humanitarian organizations are not alone: the 'dual-use' nature of satellites

However, humanitarian organizations are not alone in making use of satellite systems; governmental and military entities do so as well, to the extent that, as recent events highlight, the military application of satellites is an integral component of strategic operations. This situation has been furthered by the increase of cheaper and smaller commercial satellites, which has expanded the possibilities that satellites provide. This

has created a fresh set of challenges for humanitarian organizations. Indeed, if a satellite system used for military and humanitarian purposes is targeted by an actor that wants to hinder the military activities of an enemy, the humanitarian action would also be impacted. For example, a cyber operation against a satellite system on which humanitarian services rely could disable such system temporarily or permanently, resulting in widespread adverse consequences for the humanitarian action and the people benefiting from it. Besides that, the satellite infrastructure is managed mainly by private companies. States are becoming more and more dependent on the private sector for satellite connectivity – particularly on Low/Medium Earth Orbit satellites; and the strategic importance of these satellites is growing by the day. This dependence is attributable mainly to states' wish to amortize the costs of setting up and operating the infrastructure. The result is a dual use of these satellite systems, a situation in which the same infrastructure and potentially the same communication frequencies are used for governmental and military purposes on the one hand, and humanitarian purposes on the other. This dual use of the same tools – for military purposes by one party to a conflict against another, and by humanitarian organizations for humanitarian purposes related to the same conflict – might also call into question the neutrality, impartiality, and independence of the humanitarian organizations concerned.

The way forward: how to ensure the continuity of humanitarian action in satellite systems

Given the negative impact on humanitarian operations of the disruption of satellite services, there is a need to understand how to ensure to humanitarian relief personnel uninterrupted multi-system access to satellite services. In this context, there is a set of possibilities that can be explored to address these concerns. For instance, the space industry can contribute to the resilience-building of space-based services critical to civilians, by separating military from civilian services and enabling humanitarian responders with uninterrupted access to space systems in emergencies. At the same time, looking at the future, other technical, and legal solutions – and policies – can be taken into consideration, such as segregating the infrastructure logically and physically or marking it as protected object; as well as setting standards, drafting regulations or protecting a "humanitarian frequency" for satellite communication.

Guiding questions:

- What potential obstacles may arise for humanitarian organizations as a result of their use and dependence on satellites? What specific risks are associated with the 'dual-use' nature of satellites?
- What technical options can humanitarian organizations explore to separate the use of satellite systems for humanitarian and non-humanitarian purposes?
- What role can have the space industry in enabling humanitarian responders with uninterrupted access to space systems in emergencies?
- Are there any lessons and strategies that humanitarian organizations can draw from other dual-use infrastructure, such as telephone antennas?

Additional material:

- International Committee of the Red Cross (2023). *Preliminary recommendations on possible norms, rules and principles of responsible behaviours relating to threats by states to space systems*. Working paper submitted to the open-ended working group on reducing space threats to norms, rules and principles of responsible behavior.
- Zhou W. (2023). "War, law and outer space: pathways to reduce the human cost of military space operations", in *Humanitarian Law & Policy Blog*.
- Caribou Space (2022). *Beyond Borders: Satellite Applications for Humanitarian Emergencies*.
- Guida E. (2021). *The use of satellites in humanitarian contexts*. Norwegian Centre for Humanitarian Studies Paper.
- Luxembourgish Directorate for Development and Humanitarian Affairs (2023). *Humanitarian Aid and Emergency.lu: Real hep in case of a crises or disaster*.

Working Session #9

MARKING PROTECTED OBJECTS IN A DIGITAL SPACE: DIGITALIZING THE RED CROSS AND RED CRESCENT AND RED CRYSTAL EMBLEMS?

Objectives of the Working Session:

- Inform stakeholders on use and function of "distinctive emblems" and of "distinctive signals".
- Assess the solutions currently explored to mark protected objects in a digital space as well as the challenges in their operationalization, to gather practical inputs towards their further development.
- Examine the risks associated to the use of these tools and identify appropriate mitigation measures that can be put in place.

Background information:

The distinctive emblems

International Humanitarian Law (IHL) asserts that during armed conflict the wounded and sick, and those who tend them, as well as their facilities, units and transports, must be respected and protected at all times. The distinctive emblems were created as signs of this protection. The legal protection provided to those displaying the emblem lawfully is significant: entities displaying a distinctive emblem must be respected and protected. Moreover, as stated in the Statute of the International Criminal Court, it is a war crime to intentionally direct attacks against buildings, material, medical units and personnel lawfully using the distinctive emblems of the Geneva Conventions of 1949. Under International Humanitarian Law, the distinctive emblems take the form of a red cross, red crescent or red crystal. International Humanitarian Law also foresees the possibility of using "distinctive signals", including lights, radio or electron signals to indicate that an entity is protected.

Importantly, no single authority is entrusted to regulate, monitor and enforce the use or misuse of the emblems throughout the world. It is each state's responsibility to do so, preventing and suppressing misuse, and enforcing the protection due to the emblems, both in peacetime and in armed conflict.

The digital age: the need for a new emblem?

As societies digitalize, cyber operations are becoming a reality of armed conflict. A growing number of states are developing military cyber capabilities, and their use in armed conflict is likely to increase. The ICRC has issued public warnings about the potential human cost of cyber operations, and in particular, about the vulnerability of the medical sector and humanitarian organizations to harmful cyber operations, both having been targeted in recent years. With all this in mind, the ICRC decided to investigate the

possibility of incorporating the red cross, red crescent and red crystal emblems in information and communication technology, for instance by means of a “digital emblem”.

The idea of adapting protective emblems or signals to technological progress is not new: International Humanitarian Law foresees, and states have made us of, possibilities for introducing new means of identification in the form of “distinctive emblems” or “distinctive signals” – namely light, radio or electronic signals – to indicate that an entity enjoys specific legal protection. A “digital emblem” could – in pursuit of the same protective purposes as the display of the distinctive emblems in their physical form – become an additional component in the identification and protection of medical and certain humanitarian actors during armed conflict.

The “digital emblem”

At a glance, the objective of a “digital emblem” is simple: identify, and thereby protect, the assets, services and data of authorized medical and humanitarian actors in times of armed conflict. The “digital emblem” signals protection and does not contribute to defending the marked entities. In addition, it could make it easier for cyber operators to avoid harming protected infrastructure. However, there is the risk that such “digital emblem” increases the exposure of medical and humanitarian assets, services, and data, due to their increased visibility. Another aspect that needs to be considered is that a “digital emblem” could create a false sense of safety and protection, or give a false impression that unmarked entities are not protected. With this idea, the ICRC has – since 2020 and in partnership with a number of research institutions – been exploring the technological feasibility of developing a “digital emblem”. It convened a group of international experts to assess the potential risks and benefits associated with such an emblem. At the same time, the ICRC – with the support of the Australian Red Cross – has consulted different National Red Cross and Red Crescent Societies about this initiative.

Besides that, other initiatives for identifying objects to be protected have gained ground in recent years; and proposals for the use of new technologies to expand the capabilities of the physical emblem have been suggested by various researchers. In this context, these different proposals need to be discussed, assessing their potential shortcomings and possibilities for further improvement, as well as to suggest means of adoption and implementation. Along the “digital emblem”, this Session will also discuss an initiative called WhiteFlag Protocol, that allows parties involved in conflicts and disasters to globally and securely communicate using free, decentralized and open source technology.

Guiding questions:

- From an operational and technical standpoint, how should a “digital emblem” and “distinctive signals” be developed? What requirements should these solutions have to be meaningful?

- From a technical perspective, what are the challenges in operationalizing these initiatives?
- What are the risks related to the use of a "digital emblem" and, in that context, what mitigation measures could be explored?
- What role can have states, the international community, and the actors active in conflict settings in promoting the development and adoption of these solutions?

Additional material:

- International Committee of the Red Cross (2022). *Digitalizing the Red Cross, Red Crescent and Red Crystal Emblems*.
- International Committee of the Red Cross (2015). *The Emblems*.
- WhiteFlag Foundation (2018). *WhiteFlag Protocol*.

Working Session #10

MAPPING DIGITAL RISKS AND DIGITAL HARMS

Objectives of the Working Session:

- Explore the various angles different stakeholders use to identify, understand and address digital risks in humanitarian settings.
- Further the understanding and awareness regarding the types of digital risks, threats, and harms in humanitarian settings.
- Gather consensus on framing digital risks as a source of harm in conflict settings.

Background information:

The rise of digital risks to civilians

In humanitarian crises and conflict settings, access to digital technologies and infrastructure can save lives: they can help provide critical information to those that seek refuge, allow them to maintain or find contact with loved ones, and even effectively support the delivery of humanitarian assistance. However, the digitalization and datafication of the conduct of armed conflicts and humanitarian action has also fostered and amplified an array of tangible risks for civilians.

As malicious uses of digital technologies increasingly destabilize broader societies, they create compound concerns for civilians in conflict zones. These uses cannot only aggravate vulnerabilities of civilians, but also affect, restrict and violate their fundamental rights, undermining their safety, dignity, and resilience, as well as impede access to essential and humanitarian services, or lead to physical and psychological harm. Certain vulnerable populations can be disproportionately or uniquely affected including due to intersectional factors.

Categorizing digital risks

As humanitarian organizations start grappling with, understanding, and responding to digital risks and their effects, there is a need for common language, concepts, and frameworks. More fundamentally, there remains a gap in terms of cataloguing, documenting, and contextualizing different categories of digital risks as well as devising inclusive and community-based approaches to them.

While recognizing the centrality of protection as well as the varied issues related to upholding the 'do no harm' principle in the digital sphere, it is important to focus on digital risks and protection concerns stemming from the use of specific digital technologies and behaviors of specific actors that risk harming affected populations as well as disrupting impartial humanitarian relief efforts.

Indeed, over the past decades, conflict parties and malicious actors have increasingly relied on a panoply of digital technologies to conduct digital operations in the digital domain, or in support of kinetic operations or integrated with them in a hybrid manner. They have resorted to offensive and disruptive cyber operations, intrusive surveillance, misinformation-disinformation or hate speech campaigns, or connectivity denials, to target and disrupt civilian and humanitarian infrastructure and services, to incite violence against civilian populations, to prevent communication lines, or enable discrimination.

Documenting digital risks

Despite the increasing number of anectodical cases (e.g. Myanmar, Armenia, Syria, Ukraine), there remain a noticeable gap in the documentation of digital risks and their considerations into humanitarian activities. Coherent, actionable, and systematic documentation remains however essential for informing risk awareness and how different stakeholders are making use of and are affected by digital technologies. Incomplete documentation hampers well-informed, rational, and effective development, planning, and implementation of protective frameworks and actions for those facing these contemporary types of harm.

Documenting digital risks, however, present significant challenges throughout its constituting elements (i.e. event, attribution, victim and harm). For instance, they may be less visible, tangible, understood, while their effects may be delayed or of secondary order. Digital risks may scale up fast and have a wide reach as well as evolve rapidly as digital technologies and digital humanitarian action evolve. Moreover, the causal link between event and harm may be harder to establish, the actual harm and impact harder to measure, and require the use of novel techniques, tools and infrastructure.

Guiding questions:

- How should humanitarian organizations frame and categorize digital risks and harms?
- What specific digital risks should humanitarian organizations pay attention to and why?
- What are the specific challenges linked to the documentation of digital risks and harms?
- Are there specific vulnerabilities that humanitarian organizations should address?
- What are the ways to ensure inclusion of affected people in addressing digital risks?
- What best practices or principles should humanitarian abide by when documenting or assessing digital risks?
- How and where can academia, private sector, and humanitarian organizations cooperate?

Additional material:

- Rizk J. & Cordey S. (2023). "What we don't understand about digital risks in armed conflict and what to do about it", in *Humanitarian Law & Policy Blog*.
- AccessNow (2023). KeepItOn. *An overview of global internet shutdowns*.
- United Nations High Commissioner for Human Rights (2022). *Internet shutdowns: trends, causes, legal implications and impacts on a range of human rights*.
- ICRC (2023). *Misinformation, disinformation and hate speech – Questions and answers*.
- ICRC Global Advisory Board (2023). *Protecting civilians against digital threats during armed conflict: Recommendations to states, belligerents, tech companies, and humanitarian organizations*.

Working Session #11

DATA PROTECTION, DIGITAL RISKS, DATA RESPONSIBILITY AND ETHICS: HOW THESE FRAMEWORKS FOR ANALYSIS INTERACT TO GUIDE RESPONSIBLE USE OF TECHNOLOGY IN HUMANITARIAN ACTION

Objectives of the Working Session:

- Investigate the different frameworks for analysis that can guide the use of technology in humanitarian action.
- Discuss the interaction and overlapping of these frameworks, identifying possible commonalities between them as well as their specificity *vis-à-vis* the others.
- Understand whether and how these frameworks are necessary in the context of humanitarian action.

Background information:

The digital transformation is changing humanitarian action and can facilitate humanitarian work, but it also generates new challenges. Several frameworks for analysis are suggested in connection to these emerging issues. They include personal data protection; mapping and framing digital risks and digital harm; data responsibility; and ethics.

Personal Data Protection

Data protection legislation is rapidly evolving, and now more than 100 countries have data protection laws, and new ones continue to be drafted as awareness of the need to protect data spread throughout the world. These legislations establish a lawful, fair and transparent processing – looking at the consent of the individuals and at the necessity for such processing. These rules also promote purpose limitation and data minimization. In doing so, the data protection framework protects the autonomy of individuals, by providing clear and understandable information on the data collected. Importantly, this framework does not consider individuals as "passive" people on which data protection rules apply. Rather, it promotes an active role of the individuals, thanks to the consideration given to their autonomy and decision-making agency. This means that, safeguarding the personal data of individuals, particularly in testing conditions such as armed conflicts and humanitarian emergencies, is an essential aspect of protecting people's live, physical and mental integrity, and their dignity.

Digital risks and harm

Understanding the intricacies of the digital risks created by the deployment and use of new digital technologies is at the core of protection work, aimed at safeguarding the lives, safety and dignity of civilians. Protection activities seek to reduce the exposure to risks,

reducing vulnerabilities through technical and humanitarian assistance by supporting self-protection measures, risk education, and providing adequate accurate information. A closer focus on these risks allows to appreciate that they relate to the protection of data and digital assets and also to the any risk mediated or enhanced by digital technologies, whether physical, logical, or informational. Harmful information online, cyber operations, the automation of military systems, misuse of personal and humanitarian data, connectivity shutdowns, as well as the increased involvement of civilians in conflict, impact the rights, safety, dignity and resilience of populations affected by conflict. Looking closely at digital risks can allow humanitarian organizations to tailor and adapt their humanitarian action, reinforcing the agency of the affected people as well as their trust in humanitarian organizations.

Data responsibility

In humanitarian action, data responsibility is the safe, ethical and effective management of personal and non-personal data for operational response. In fact, given that data is a critical component of humanitarian response, an irresponsible data management in humanitarian responses can place already vulnerable people and communities at greater risk of harm or exploitation. In particular, in humanitarian contexts, both personal and non-personal data can be sensitive – meaning that, if disclosed or accessed without proper authorization, can cause harm on a person and/or a negative impact on an organization. Indeed, while the sensitivity and importance of personal data is well-understood, this is not the case for non-personal data, such as the location of a medical facility in an armed conflict, which can expose patients and staff to risk. Data responsibility encompasses actions to ensure data protection and data security, as well as strategies to minimize risks while maximizing benefits in operational data management. Data responsibility takes in consideration the whole data lifecycle, underscoring the importance of data management to avoid unintentional harm to vulnerable populations.

Ethical considerations

Ethics ensures that technology is used with integrity, respect for human dignity and in accordance with moral principles. In addition, digital tools developed in accordance with ethical considerations are trustworthy. These ethical considerations are about more than technical specifications, encompassing also risks of harm at every stage of the digital products lifecycle. For example, several codes of ethics invoke principles like fairness and justice in connection to the accountability coming from a misuse of those technologies. Importantly, considering the global scale of the digital transformation, ethical considerations around the use of technology emerge at the individual, societal and state level. Ethical considerations can guide the use, as well as the design, development and deployment of technology in humanitarian action, providing a moral compass that is not affected by the ever-changing technological landscape.

The interplay between these frameworks for analysis

These frameworks are often described as being alternatives to or mutually exclusive from one another. However, a closer examination of the matter can reveal how they interact and how, when used together, can guide responsible use of technology in humanitarian

action. This session therefore aims to discuss how these frameworks interact and overlap, and whether it is possible for them to complement one another. Discussing the commonalities will illustrate their respective necessity, as each one of them addresses at a different level the challenges brought by the digital transformation in humanitarian action. The session will also explore how, taken together, these frameworks for analysis can have a positive impact on the humanitarian sector's approach to tackling digital harm.

Guiding questions:

- How do the identified frameworks for analysis interact and overlap in guiding responsible use of technology in humanitarian action?
- What are the commonalities between the identified frameworks?
- How is each of these frameworks specific *vis-à-vis* the others? In the context of humanitarian action, how are these different frameworks necessary and complementary?

Additional material:

- International Committee of the Red Cross (2020). ICRC Rules on Personal Data Protection.
- Kuner C. & Marelli M. (eds.) (2020). Handbook on Data Protection in Humanitarian Action. ICRC.
- Rizk J. & Cordey S. (2023). "What we don't understand about digital risks in armed conflict and what to do about it", in Humanitarian Law & Policy Blog.
- Inter-Agency Standing Committee (2023). IASC Operational Guidance on Data Responsibility in Humanitarian Action.
- OCHA Centre for Humanitarian Data (2021). OCHA Data Responsibility Guidelines.
- Hardebolle C., Macko V., Ramachandra V., Holzer V. & Jermann P. (2023). Digital Ethics Canvas: A Guide for Ethical Risk Assessment and Mitigation in the Digital Domain.

Working Session #12

MEASURING THE HARM DONE BY CYBER OPERATIONS TO THE AFFECTED PEOPLE

Objectives of the Working Session:

- Sensitize the various stakeholders on the particular impact of cyber operations on communities living in places characterized by armed conflicts and other situations of violence.
- Examine existing methodologies, categories as well as indicators of harm, towards improving our capacity to effectively measure the harm done by cyber operations to the affected people.

Background information:

The components of cyber harm

In recent years there has been an increase in the breadth and scope of cyber operations. Along with this increase, there has also been an increase in the severity of the harms caused by these operations. Efforts to understand and measure the harm caused by these activities have been in progress for some time. Particular attention has been placed towards measuring the financial impact of these operations – with some reports that focused on assessing cyber harm in terms of the "dollar damage". At the same time, other approaches have looked at cyber harm identification and management as a component of overall cyber risk management. Besides that, there has also been interest on classifying and measuring cyber harm as a way to increase the scale and effectiveness of cybersecurity capacity-building. In the meantime, some experts have looked at the reputational impact of cyber harm at an organizational and national metric, using metrics such as victimization surveys, sentiment analysis and specific case studies. Only limited attention has been given to measuring the cyber harm done to the people affected by these operations, especially in humanitarian emergencies. The research dealing with this topic has mainly looked at victimization surveys and at medical and crime statistics. However, understanding the *harm* caused by cyber operations as limited to the physical sphere leads to undermining a true evaluation of the scope and magnitude of such attacks.

The need to measure the cyber harm done to the affected people

In this context, it may be very difficult to assess the true harm done to affected people – important to understand whether the harm caused is lawful in light of the normative framework, and to determine what mitigatory and protection programs to carry out in response. Indeed, a concrete set of indicators of harm as well as tools and standards to measure cyber harm can strengthen the capacity to address the protection concerns of victims of these operations. In addition, a harm assessment can potentially help to increase cyber resilience. This is because understanding the direct and indirect harm

caused by cyber incidents can help individuals and organizations to identify potential risks and vulnerabilities; prioritizing anticipatory and preventive action, and responses; and develop effective mitigatory strategies to minimize harm.

A recent initiative to develop a standardized harms methodology

Given the importance of protecting the affected people by cyber operations, as well as of developing preventive programs and resilience mechanisms, such as accountability processes, there is the need to measure the harm done by cyber operations to the affected people. In this context, a recent initiative is that of the CyberPeace Institute, that initiated a research to develop a standardized methodology of harm. The objective of this project is to identify means to measure and assess the harm of a single incident across multiple indicators and categories of harm.

This Session, looking at this initiative, aims at examining and discussing the identified categories of harm, as well as the way in which these categories inform the broader framework (e.g. should indicators be derived from categories, or should categories be created from indicators?). At the same time, it is important to closely understand what scoring methods need to be considered, and the solidity and robustness of the proposed indicators. Indeed, by further developing the capacity to measure the harm done by cyber operations to the affected people, it will be possible to develop more solid and holistic accountability measures for the affected people.

Guiding questions:

- What are the different types of harm in which this impact can manifest (e.g. physical, psychological, economic)?
- What is the relation between the specific categories and a broader methodology?
- What approaches should be adopted to assess and measure harm? How (and where) to find a balance between measurability and quantifiability *vis-à-vis* a holistic conceptualization of *harm*?
- How to choose the indicators to measure cyber harm and indirect harm? How to identify a set of indicators that are context-specific and relevant to different categories of harm?

Additional material:

- CyberPeace Institute (2023). CyberpeaceWatch. Report of Expert Meeting on the development of a Harms Methodology.
- Agrafiotis I., Bada M., Cornish P., Creese S., Goldsmith M., Ignatuschtschenko E., Roberts T. & Upton D. M. (2016). "Cyber Harm: Concepts, Taxonomy and Measurement", a Working Paper in Saïd Business School WP 2016-23.

- Agrafiotis, I., Nurse J.R.C., Goldsmith M., Creese S. & Upton, D. M. (2018). "A Taxonomy of Cyber-harms: Defining the Impacts of Cyber-attacks and Understanding How They Propagate", in Journal of Cybersecurity. Vol. 4, n° 1: 1-15.

Working Session #13

HUMANITARIAN HEALTH SERVICES AND DIGITAL TOOLS

Objectives of the Session:

- Reach a common understanding of the risks and challenges related to humanitarian provision of digital health services for people on the move.
- Discuss how Data Protection by Design can inform how humanitarian organizations respond to challenges in service delivery.
- Identify the proper mechanisms related to data protection can reduce and mitigate the harms experienced by people on the move.

Background information:

Some of the world's most vulnerable populations, including refugees, internally displaced persons, and migrants, receive medical attention and services exclusively from humanitarian organizations. As such, humanitarian organizations, such as Médecins Sans Frontiers (MSF) and the ICRC, have worked towards developing effective, and often digital modes, to assist the delivery of such care. This session considers the risks related to the protection and management of health data that arise when affected people, such as migrants and refugees, rely on digital tools for healthcare services. Health data, such as digital medical records, is highly sensitive, and its processing falls under the highest data protection standards. This session invites stakeholders to discuss how these standards figure in to the development of such services, and will investigate how to bridge health innovation with data protection by design.

Each year, thousands of people are forced to flee from war, persecution, and poverty and undertake dangerous journeys, often with sporadic access to medical care. These complex journeys, which involve frequently crossing multiple international borders, pose major difficulties to the safe and seamless management of their health data. For instance, different countries may have different processes and policies for the processing of medical data, which can make it difficult for people on the move to maintain agency and clarity over their own medical information. Humanitarian organizations, many of which supply medical aid to people on the move, must learn how to effectively navigate these dynamics. Much of the data collected within programs serving People on the Move is highly sensitive. Yet, some of this data is essential for public health monitoring and response, and is helpful in designing and developing humanitarian action. Protecting this data from surveillance, while also enabling the sharing of information is continuously balanced, requires the identification of appropriate safeguarding and risk mitigation measures.

Guiding questions:

- What are the data risks, harms, and threats for people on the move whose medical data is processed by humanitarian organizations?

- How can humanitarian organizations that manage medical data from people on the move safeguard the data by implementing Data Protection by Design practices?
- How can humanitarian digital services and tools maintain data protection, responsible data, and information security standards to safeguard the medical data of people on the move?

Additional material:

- Beirens, Hanne. "Rebooting the Asylum System? The Role of Digital Tools in International Protection." Migration Policy Institute, October 2022. https://www.migrationpolicy.org/sites/default/files/publications/mpi_digitalization-asylum_final.pdf.
- Alencar, Amanda. "Technology Can Be Transformative for Refugees, but It Can Also Hold Them Back." *Migration Policy Institute*, July 26, 2023. https://www.migrationpolicy.org/article/digital-technology-refugees.
- Bowsher, Gemma, Nassim El Achi, Katrin Augustin, Kristen Meagher, Abdulkarim Ekzayez, Bayard Roberts, and Preeti Patel. "EHealth for Service Delivery in Conflict: A Narrative Review of the Application of EHealth Technologies in Contemporary Conflict Settings." *Health Policy and Planning* 36, no. 6 (June 25, 2021): 974–81. doi:10.1093/heapol/czab042.

Working Session #14

DATA PROTECTION BY DESIGN AND BIOMETRICS IN HUMANITARIAN ACTION

Objectives of the Working Session:

The main objective of this working session is to explore practical measures that can reduce the risks related to the processing of biometrics data in humanitarian action. The goal is to come up with a list of such measures.

To achieve this, it is proposed to explore two orthogonal but complementary dimensions:

1. What are the potential mechanisms to assess the proportionality of using biometrics to provide a mean of identifying? How could we properly balance the theorical benefits with reality of the field?
   (note: underlying to this question is the further question of whether those benefits really exist but this fundamental point should not be the focus of the session).
2. If biometrics is deployed, what technical means can be included in the design of the solution to tackle the problem of being over-purposed by nature? The principle of purpose limitation should guide the reflection which is intended to be quite technical. Properties of irreversibility, revocability and unlinkability will also play a role in the reflection.

Background information:

The purpose of collecting biometrics, as its Greek roots indicate ('bio' meaning 'life', and 'metrics' meaning 'measure'), is to measure a parameter of life. In other words, biometric data relate to who we are and provide the measure of a particular trait of an individual. Whenever an organization collect biometric data, it is with the intention to identify people.

In humanitarian action, the need to identify individuals has always been crucial to deliver aid. Whether to provide adequate health treatment or in forensic science, humanitarian workers rely on the ability to identify affected people. In practice, this means that lists or databases of individuals are created. Those lists contain records of personal data with biographic data (e.g. names, sex, marital status, origin, data of birth), identity numbers referencing other lists (e.g. national id, voter id, tax id) and biometrics data.

The use of biometrics-based identification systems in humanitarian programs is not a novel concept and many humanitarian organizations put an important effort into the development and deployment of biometric solutions as a mean for facilitating individual registration, authentication and aid distribution. Using biometrics to replace paper-based mechanisms come with the promise of strengthening programme accountability toward affected communities and donors, avoiding fraud and aid diversion, and increasing the efficiency and effectiveness of humanitarian programmes.

Despite these promises, the use of biometrics-based identification systems in humanitarian action has been the subject of intense debate, and figures prominently in the discourse on "humanitarian experimentation"[4]. Documented cases of abuse and significant risks induced by the nature of biometrics have been reported, accompanied by recommendations to either limit or totally ban the use of biometrics. Parallelly, some publications have shown projects where biometrics greatly helped to reach objectives otherwise complicated to achieve.

Interestingly, a quick comparison between successful projects and more problematic biometrics deployments often share the same technical components. Majority of biometrics-based identification systems are often based on the same design and the same data flows. In particular, the processing of the raw biometric data to create the template that will be stored in the database use the same techniques no matter what solution is deployed. This observation is also true for the algorithms used to compare two (or more) templates. In other words, whether a biometrics-based system will cause harm or not may not depends so much on the technology itself but rather on the context, the scale and the governance.

The importance of the context does not come as a surprise, as it is often said that risks brought by technology are essentially stemming from the usage made rather than from the technology itself. However, it can be argued that designing a system taking into account a threat model and following Data Protection by Design (DPbD) practices can mitigate several risks related to the processing of biometrics data.

One of the core data protection principles is the principle of purpose limitation. Purpose limitation is essential with biometrics-based system because biometrics is "over-purposed by nature". By that we mean that any biometrics trait (e.g. finger prints, iris, faces, DNA) intrinsically bear more information than needed for a simple identification. Iris can reveal health condition, faces reveal ethnicity, gender and age, while DNA reveals pretty much anything about the individual. The second aspect of being over-purposed by nature is caused by the unicity of biometrics data: while we can have many passwords or IDs, we only have one iris. Therefore, once our iris scan is in one database, it is in all databases!

Guiding questions:

- What are the most feasible mechanisms for humanitarian organizations to take to safely deploy biometrics?
- How can the practices of DPbD be leveraged to help humanitarian organizations overcome such risks?

[4] See "Humanitarian Experimentation", https:// blogs.icrc.org/law-and-policy/2017/11/28/ humanitarian-experimentation/ and Engine Room & Oxfam "Biometrics in Humanitarian Sector", https://www.theengineroom.org/ wp-content/uploads/2018/03/Engine-RoomOxfam-Biometrics-Review.pdf

- What are the current obstacles that may prevent DPbD from being incorporated into the working modalities of organizations that rely on biometrics?

Additional material:

- Kuner, Christopher, and Massimo Marelli, eds. "Chapter 8 Biometrics." In *Handbook on Data Protection in Humanitarian Action*, 2nd ed., 127–42, 2020. https://shop.icrc.org/download/ebook?sku=4305.01/002-ebook%20.
- Sandvik, Kristin Bergtora, Katja Lindskov Jacobsen, and Sean Martin McDonald. "Do No Harm: A Taxonomy of the Challenges of Humanitarian Experimentation." *International Review of the Red Cross* 99, no. 904 (April 2017): 319–44. doi:10.1017/S181638311700042X.
- The Engine Room. "Biometrics in the Humanitarian Sector: A Current Look at Risks, Benefits and Organisational Policies," July 2023. https://www.theengineroom.org/wp-content/uploads/2023/07/TER-Biometrics-Humanitarian-Sector.pdf.

Working Session #15

THE PRINCIPLE OF HUMANITY AND THE USE OF ARTIFICIAL INTELLIGENCE IN HUMANITARIAN ACTION

Objectives of the Working Session:

- Discuss how humanitarian principles, primarily Humanity, can be used to understand and address challenges that arise from autonomous decision-making
- Explain how humanitarian organizations can ensure a "human-in-the-loop" approach and "meaningful human intervention" in the context of their operations
- Translate the technical elements of AI, in a manner that informs the policy and regulatory implications of its deployment

Background information:

Humanitarian organizations are driven by the principle of Humanity, or the will to save lives and alleviate suffering, wherever it may be found, in a manner that respects and restores personal dignity. This element distinguishes humanitarian activities from other ones, for example those of a political, ideological, religious, or military nature.

This aspect is behind the adoption of the four Geneva Conventions, the pillars of the law of armed conflict, and highlights that – even in conflicts – a compromise must be achieved between military necessity and Humanity. This principle has been endorsed, among others, by the UN General Assembly as well as by the largest humanitarian network in the world: the Movement of the Red Cross and Red Crescent.

While the principles guiding humanitarian organizations have – and should – remain unchanged, the tools the sector uses to realize such principles have transformed dramatically. One of the starkest examples of such evolution, has been the humanitarian adoption of Artificial Intelligence (AI). The popularization of so-called "generative AI," or Large Language Models (LLMs) over the past year has catapulted Artificial Intelligence into the global spotlight. During this period of intense media and political coverage, it can be challenging to discern AI hype from merit. However, leveraging AI and autonomous decision-making is crucial to how the humanitarian sector carries out its day-to-day operations, and how it promotes safeguards for individuals during armed conflict and other situations of violence.

For the purposes of this session, the term "artificial intelligence" (AI) refers to the "set of sciences, theories and techniques whose purpose is to reproduce by a machine the cognitive abilities of a human being."1 This delegation of decision-making from human to machine represents a major shift in the way organizations operate. Humanitarian organizations already experience the impact of this change, as many now rely on AI-powered programs to conduct detailed environment scanning, trend analysis, predictive

response planning, and even beneficiary identification. By relying on AI-facilitated analysis of large data sets, humanitarian organizations can make certain elements of their operational modalities more efficient, but can simultaneously open themselves – as well as their beneficiaries – to new sets of possibilities, complexities, and harms.

For example, some humanitarian organizations now rely on AI to support their operations in predictive environmental forecasting and trend analysis. While the speed at which AI systems can "make sense" of such great quantities of data may be attractive, understanding the means through which AI processes information is critically important to ensuring that its analysis stays unbiased and accurate. This "black box problem" of AI – or the inability for humans to assess how and why AI systems reach their conclusions – stresses the need to bring human accountability and human control to the fore.

While this session focuses on the preservation of Humanity throughout AI-mediated humanitarian operations and contexts, it is likewise necessary to consider the human cost of the labor and environmental resources necessary to keep sophisticated AI algorithms up and running. Addressing the wider implications of AI systems' operationalization goes beyond the scope of this session, however remains an important factor in the debate surrounding AI and the principle of Humanity.

Guiding questions:

- As organizations increase their operational reliance on AI, how can we ensure that humanity remains at the center of humanitarian action?
- Is it possible to have "human intervention" or "human control" in autonomous decision making? If so, what does this look like in the context of humanitarian operations?
- What types of regulatory frameworks can be but in place to maintain a "human-in-the-loop" approach. How can humanitarian organizations help ensure they are sufficient?
- How can humanitarian organizations institutionalize digital literacy, so that all employees are prepared to work in environments mediated by AI?

Additional material:

- Stewart, Ruben, and Georgia Hinds. "Algorithms of War: The Use of Artificial Intelligence in Decision Making in Armed Conflict." *Humanitarian Law & Policy Blog*, October 24, 2023. https://blogs.icrc.org/law-and-policy/2023/10/24/algorithms-of-war-use-of-artificial-intelligence-decision-making-armed-conflict/.
- Terry, Fiona, and Fabien Dany. "Harnessing the Power of Artificial Intelligence to Uncover Patterns of Violence." *Humanitarian Law & Policy Blog*, May 25, 2023. https://blogs.icrc.org/law-and-policy/2023/05/25/artificial-intelligence-patterns-violence/.
- Chen, Christopher. "The Future Is Now: Artificial Intelligence and Anticipatory Humanitarian Action." *Humanitarian Law & Policy Blog*, August 19, 2021.

https://blogs.icrc.org/law-and-policy/2021/08/19/artificial-intelligence-anticipatory-humanitarian/.

Working Session #16

UNDERSTANDING DIGITAL RISKS AND OPPORTUNITIES FOR CHILDREN AFFECTED BY ARMED CONFLICT, INCLUDING PROTECTING THEIR PERSONAL DATA

Objectives of the Working Session:

This working session is to encourage discussion of the most pertinent concerns and dilemmas related to the digital environment as experienced by children affected by armed conflict. Through discussion of how learning from different stakeholders and sectors at national, regional and global levels can be adapted to complex environments, including working across borders and in conflict zones, propositions will be made for stronger support to children and their families to safely navigate these spaces and for different actors to better anticipate and respond to the specific risks for these children when crisis hits.

We want to:

1. highlight the known digital risks and opportunities that children affected by armed conflict face
2. discuss contemporary dilemmas related to humanitarian responses for children and brainstorm ethical ways forward
3. discuss practical tools, approaches and good practices from non-conflict affected environments that could be adapted for children affected by conflict or migration to prevent or mitigate risks or make the most of opportunities
4. discuss challenges and good practices in protecting children's data, even in complex operating environments.

Background information:

Children affected by armed conflict, including those who are displaced across borders, face unique challenges, threats and opportunities in the digital environment. Many of the threats to children in complex environments transition between the online and offline worlds, often intersecting when children are at their most vulnerable and their habitual protective mechanisms, including family and community, are weakened or destroyed. Children can also be particularly affected when there is no, limited or controlled access to the internet.

In some of these environments and for vulnerable children, there is a delicate balance between leveraging the opportunities provided by technology for the benefit of children - such as maintaining links with family members or finding out crucial information for their own safety and security - whilst trying to guard against risks of exploitation and abuse. Children are not all the same, and due care and attention should be paid to specific groups of children with particular vulnerabilities and needs, and the risks and opportunities affecting them. In addition, all efforts need to be made to ensure to avoid

certain groups of children being excluded from opportunities. Ensuring children's protection online and offline, particularly in complex environments, remains a critical challenge for governments, humanitarian organisations, and the global community as a whole.

Digital Risks:

Exploitation, violence and abuse: Children affected by conflict and/or displacement, including migration, are at heightened risk of exploitation, including sexual exploitation, and trafficking, as their circumstances and information that either they share online as users of technology or that others share online about them can make them targets.

Mental Health and Psychosocial Harm: Exposure to graphic content and violent imagery online as well as cyberbullying can cause psychological distress. This is not unique to children in conflict zones or child IDPs or migrants, including refugees, but may be particularly harmful to children who are already distressed, traumatised or suffering from mental health conditions, especially children without adequate support systems in place.

Misinformation, Disinformation and Hate Speech: The digital environment can expose children to misinformation, disinformation and hate speech, including about the conflict or their own or other communities, which may create the conditions for other kinds of harm (violence, exploitation, trafficking, recruitment etc).

Recruitment by Armed Actors: digital platforms are used by armed actors to circulate information, create communities and, ultimately, to recruit and traffick children.

Surveillance and monitoring: children and their families may be exposed to increased online surveillance and monitoring through digital means. Children may often be tracked without their knowledge or consent.

Data Protection and Security: Inadequate data protection measures expose children's personal information, putting them at risk of identity theft and other forms of abuse and exploitation, which can also emerge at any point in their future. The diverse range of actors online affect both data protection and security.

Child Safeguarding and Discrimination: the use of new technologies, including AI, may reinforce structural biases against groups of children that may lead to unintentional discrimination.

Opportunities through the digital space:

Education: Digital tools can provide access to education to children affected by conflict or who are displaced, allowing children to continue learning, including new languages, and to take exams even when traditional education provision is disrupted.

Skills Development: Access to the digital world can empower children with valuable digital skills that can improve their future prospects and ability to navigate the world.

Mental Health and Psychosocial Support: Online resources and counselling services can help children cope with trauma and mental health challenges in the absence of adequate in-person support in their own language.

Communication with families and friends: Digital technology allows children separated from their families and friends by conflict and displacement to restore and maintain contact with loved ones which, in turn, provides emotional support, reduces the sense of isolation and creates a sense of connection despite physical separation.

Family tracing: Digital tools and databases help humanitarian organizations and families search for and locate separated family members and quickly share required documentation (birth certificates, ID, school certificates), making proof of identity or the process of requesting family reunification more efficient.

Provision of information and services: Digital resources such as navigation and money transfer apps and translation tools assist migrants, including refugees, in finding safe routes, accessing essential information and services (including healthcare), and communicating locally during their journey and on arrival.

Advocacy and child participation: Digital platforms enable children to share their stories, connect with others and advocate for their rights, raising awareness about their unique struggles on a global scale.

Guiding questions:

In this working session we aim to discuss:

Identifying Digital Risks and Opportunities for Children Affected by Conflict:

- What are the specific digital risks that children affected by armed conflict encounter (which translate into online/offline harms), and how can we better understand and mitigate them? What are the mechanisms in place to identify these risks (such as human rights/child rights due diligence and impact assessments)? Where do we see gaps in tools, understanding or response?
- What are the opportunities that the digital environment can bring to children affected by armed conflict?

Learning from successes elsewhere – what can be adapted, for example on policy enforcement? Do we have good practice from complex environments?

- Can you share examples of successful initiatives or projects that have addressed the unique challenges faced by children affected by conflict in the digital age, and what lessons can be learned from these experiences?
- Are there examples of successful initiatives that have effectively addressed digital risks for children and provided support for children in non-conflict and/or migration settings that could be relevant for responding to the harms we see today? What are the good practices from less complex operating environments that could be replicated?

- What can be adapted or replicated for children affected by armed conflict? What are the barriers to transposing these approaches to other settings?
- What are the elements that are more likely to be factors for success? (Scale; proximity; focussing on different groups of children; digital, media and information literacy; existence of clear legal frameworks and effective complaint and remedial mechanisms etc?)

Safeguarding Children's Personal Data:

- What are the current challenges in safely collecting and managing children's data in complex operating environments? When is the theory difficult to implement? What are the good practices?
- Is the need to safely protect children's data slowing down or changing the possible humanitarian responses for children affected by conflict?

In conclusion: How can governments, humanitarian organizations, and tech companies design/implement protective actions in and/or related to the digital environment for children affected by armed conflict?

Additional material:

- Office of the United Nations High Commissioner for Human Rights (OHCHR). "General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment." *OHCHR*, March 2, 2021. https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation.
- UNICEF. "Policy Guide on Children and Digital Connectivity." New York, USA: Policy Lab; Data, Research and Policy; United Nations Children's Fund, June 2018. https://www.unicef.org/esa/media/3141/file/PolicyLab-Guide-DigitalConnectivity-Nov.6.18-lowres.pdf.
- UNICEF. "Child Protection in Digital Education." Policy Brief and Technical Note. New York, USA: UNICEF, January 2023. https://www.unicef.org/documents/child-protection-digital-education.
- UNICEF. "Legislating for the Digital Age: Global Guide on Improving Legislative Frameworks to Protect Children from Online Sexual Exploitation and Abuse." Report. New York, USA: UNICEF, May 2022. https://www.unicef.org/reports/legislating-digital-age.
- Georgiou, Myria, and Koen Leurs. "Smartphones as Personal Digital Archives? Recentring Migrant Authority as Curating and Storytelling Subjects." *Journalism* 23, no. 3 (March 1, 2022): 668–89. doi:10.1177/14648849211060629.

Initiatives related to children in the digital sphere:

Global Kids Online

EU Kids Online

The Responsible Data for Children (RD4C)

Disrupting Harm

WeProtect Global Alliance - against child sexual exploitation and abuse online

Virtual Safe Spaces | UNICEF