DIGITAL SHIELD: THE PROTECTIVE ROLE AGAINST HUMAN RIGHTS VIOLATIONS IN CYBER INTERVENTIONS

Muhammet Ali Demir[*]
ORCID ID: 0009-0004-0201-8391

Declaration[*]

Abstract

Atrocity crimes represent some of the most severe violations of international order and are primarily addressed within the framework of humanitarian intervention and the Responsibility to Protect (R2P). Traditional military interventions have been widely criticized due to their potential infringement on state sovereignty and the high risk of operational failure, whereas emerging digital technologies have introduced cyber humanitarian intervention as a possible alternative. The aim of this article is to explore the potential of cyber operations in preventing or halting mass atrocity crimes within the context of R2P and to critically assess the legal, ethical, and practical constraints of this approach.

Methodologically, the study adopts a normative analytical framework, drawing on international law, cybersecurity, and humanitarian intervention scholarship to establish a conceptual and legal basis. Existing literature tends to focus predominantly on military or diplomatic means of intervention, with only limited engagement with the notion of cyber humanitarian intervention. This gap highlights the need for a comprehensive assessment of how cyber measures align with international law, their feasibility, and associated risks.

The findings suggest that cyber interventions may support the implementation of R2P by safeguarding access to information, protecting communication infrastructures, and limiting the digital capacities of perpetrators. Nevertheless, the approach also entails significant limitations, particularly concerning state sovereignty, attribution challenges, the lack of international cooperation, and ethical accountability. In conclusion, while cyber humanitarian intervention does not constitute a definitive solution on its own, it can be considered a complementary tool for enhancing the effective realization of the R2P principle.

Keywords: Cyber, Humanitarian Response, Responsibility to Protect, R2P

[*] Lecturer, Turkish National Police Academy, Karaman, Türkiye, malidemir1501@gmail.com
[*] This article has been prepared without the use of any Artificial Intelligence (AI) tools or assistance.

Introduction

Atrocity crimes are crises that threaten individuals right to life, lead to widespread human rights violations and mass victimisation, and necessitate solutions from the international community. Such crimes, particularly in situations such as wars, genocides and internal conflicts, make the protection of civilians a moral and legal responsibility. In this context, the approaches developed by the international community through the principles of *humanitarian intervention* and *R2P* play a significant role in preventing and resolving victimisation. Humanitarian intervention aims to establish a rapid and effective intervention mechanism for crisis areas by balancing the sovereign rights of states with the fundamental rights of individuals. However, the political, legal, and ethical dimensions of these interventions give rise to international debates.

In recent years, alongside technological developments, transformations have been occurring in the dynamics of conflict and crisis, with cyber technologies reshaping the concepts of war and intervention. In this context, cyberspace has emerged as a new arena of struggle for individuals, institutions and states through information and communication technologies. The growing influence of cyberspace has brought the concept of humanitarian intervention into the digital realm. At this point, the concept of *cyber humanitarian intervention* refers to an innovative approach developed to prevent human rights violations and protect civilians through digital technologies. Methods such as information operations, digital surveillance, and the protection or manipulation of communication networks in crisis areas are considered within the scope of cyber humanitarian intervention. However, this new paradigm raises questions about how it will align with the principles of sovereignty and intervention in international law and how it will be ethically grounded.

The role of cyber humanitarian intervention in preventing human rights violations is of critical importance, particularly in conflict zones, in areas such as protecting communication infrastructure, preventing disinformation, and ensuring the safety of victims. However, fundamental challenges encountered in this process include interventions that conflict with states sovereign rights, the risk of misuse of technological tools, and technical and political obstacles that limit the effectiveness of digital interventions. Therefore, cyber humanitarian intervention stands out as a multidimensional phenomenon that presents both opportunities and risks.

This article will first examine the theoretical foundations of the concepts of humanitarian intervention and the responsibility to protect. It will then discuss the characteristics of cyberspace and the concept of cyber humanitarian intervention. Finally, it will explore the role of cyber humanitarian intervention in preventing human rights violations and evaluate the opportunities and limiting factors in this field. In this context, a critical analysis will be presented on how cyber technologies provide advantages in humanitarian intervention processes, as well as how international law and ethical values will be shaped. The article aims to discuss the potential of cyber humanitarian intervention to offer an innovative solution to humanitarian crises.

The Concept of Humanitarian Intervention and the Responsibility to Protect (R2P)

Over the past thirty years, the concept of *humanitarian intervention* has been one of the most frequently debated topics among both academics and practitioners. Questions such as what the challenging role of the concept is in relation to state sovereignty, or what the minimum level of crisis should be for intervention in humanitarian crises caused by the sovereign itself, have formed the basis of this debate. Given that humanitarian intervention is a concept that is controversial in essence, it is important to define it in order to express its scope (Gulati and Khosa, 2013: 398).

Şaban Kardaş (2003: 21) defines humanitarian intervention as coercive action taken by a state or states or international organisations against a target state that seriously and flagrantly violates human rights, with the aim of protecting the target state's citizens, through the use of armed force or the threat of force, regardless of the target state's consent. This definition, on which Kardaş bases his argument, is actually in line with the definition adopted by NATO in November 1999. The fundamental elements of this definition are focused on *sovereignty* and *human rights*. Firstly, for an action to be considered humanitarian intervention, there must be a violation of the sovereignty of the target state. Secondly, the fundamental trigger for the intervention must be the aim of resolving human rights violations (Roberts, 2000: 1).

International law did not consider any intervention on the territory of a state without the consent of that state to be legitimate, even for urgent humanitarian purposes agreed upon by the entire international community, until the Second World War. In 1945, however, the United Nations (UN) prohibited intervention, banning the use of force or the threat of force against the territorial integrity of a state, and also prevented any state from providing military support or intervention to either side in another state's civil war. Serious efforts to develop a

form of collective intervention began under the leadership of the UN Security Council (UNSC) with the end of the Cold War. In 1991 and 1992, interventions took place in Iraq and Somalia, not primarily justified on humanitarian grounds – a term not found in the UN Charter – but fundamentally due to mass human rights violations (Helkin, 1999: 824).

In 1999, NATO bombed Yugoslavia to protect the Albanian population in Kosovo from ethnic cleansing. Although this military operation was considered morally justified, it was criticised for violating international law for the sake of interests, and indeed the UN Security Council did not express a favourable opinion on the military intervention in question. (Gilligan, 2013: 22). The Kosovo intervention and crises such as those in Rwanda, Burundi and Bosnia and Herzegovina, which led to mass killings in political history, have revealed a normative deficiency agreed upon by both states and international organisations (Coady, Dobos and Sanyal, 2018: 18-19).

The Canadian government established an independent commission called the *International Commission on Intervention and State Sovereignty (ICISS)* in 2000 in an effort to overcome the humanitarian intervention crisis. In 2001, the commission published a 90-page report entitled *The Responsibility to Protect (R2P)*, accompanied by a 400-page book detailing the report. The most significant development for the concept came in 2005 when heads of state endorsed R2P in the Outcome Document of the UN World Summit. In subsequent years, the UNSC referred to the R2P concept and published a report entitled *Implementing the Responsibility to Protect* in 2009 (Badescu, 2011: 3).

The ICISS report essentially consists of three main sections. These are *prevention*, *response* and *reconstruction*. *Prevention* is the first section placed at the centre of R2P. This stage involves a shift from the habit of responding after a crisis has occurred to the habit of taking preventive measures before a crisis occurs (ICISS, 2001: 39). The prevention phase is itself divided into three sub-headings. The first part is the *Early Warning and Analysis* section, where data and information on human rights violations are collected, the reality is clearly revealed, and the aim is to take swift political action based on the data collected. The second part is *Efforts to Prevent the Root Causes of Crises*, which aims to transform the main causes of conflict, such as income inequality, underdevelopment or political oppression, through various reforms, effective governance, the protection of fundamental rights and freedoms, the rule of law and the development of welfare. The final section is *Direct Prevention Efforts*. This section includes various sanctions such as providing direct assistance to the violated

community, imposing political sanctions on the violating state, diplomatic isolation, and the threat of force (ICISS: 21-24).

Reaction is the second and most controversial section of the R2P report. There are two main reasons why it is controversial: firstly, reacting poses a threat to state sovereignty; secondly, the question of who has the authority to react. At this point, the ICISS has established six fundamental criteria for the legalisation of a military response to mass human rights violations. These are: *proper authority*, *just cause*, *right intention*, *last resort*, *proportionate means*, and *reasonable expectations*. *Proper authority* lies with the most appropriate international body, the UN Security Council. *Just cause* refers to situations involving large-scale loss of life or ethnic cleansing. *Right intention* is to stop and prevent human suffering. *Last resort* means that all possible diplomatic or non-military means must be considered before resorting to military force. *Proportional means* that the level, duration and intensity of the intervention must be kept to a minimum, taking into account humanitarian safeguards. *Reasonable expectations* should be pursued if there is a likelihood of success in preventing a humanitarian crisis following the intervention (ICISS: 32-37).

Reconstruction is the final section of the ICISS report. This section actually addresses the question of how to emerge better from the state that has been intervened in after the intervention and focuses on the post-intervention period. The fundamental aim is to ensure *lasting peace*. The objective here is not to provide humanitarian aid or achieve development goals, but to create the right conditions for genuine reconciliation that will eliminate the possibility of renewed conflict. The reconstruction process is divided into three sub-sections: security, justice, reconciliation and development. The coordination established between local and international actors facilitates reconstruction efforts (ICISS: 39-45).

The 2009 UN Secretary-General's R2P report examined the extent to which the Responsibility to Protect implementation strategy assisted the organisation's efforts to fulfil its commitment to protect communities from atrocity crimes, highlighted shortcomings, and stated that R2P should be understood as a programme based on three pillars. The first pillar is the *state's responsibility to protect*. A state is required to protect its own society from serious human rights violations such as genocide, ethnic cleansing or crimes against humanity, and this is an international obligation. The second pillar is *international assistance and capacity building*. Unlike the responsibility placed on the state in the first pillar, this pillar places a responsibility on the international community, including supporting states in protecting their populations

from these crimes, including meeting the urgent needs of communities at risk. The third and final pillar is *the international community's timely and resolute response through the UN*. This means that, when peaceful options prove insufficient and the relevant community cannot be clearly protected from atrocity crimes, member states must take collective action through the UN Security Council (United Nations General Assembly, 2009: 2).

Nicholas J. Wheeler (2000, 34-35) argues that four fundamental thresholds must be met for an international military intervention to be considered a legitimate humanitarian intervention. First, the existence of an urgent humanitarian situation, such as mass deaths or ethnic cleansing; second, the exhaustion of all diplomatic and economic avenues, making military force the last resort; third, the violence used must be proportionate and not exceed the humanitarian objective; and finally, the intervention must have a reasonable prospect of producing a positive outcome in improving the humanitarian situation in the region.

Sovereignty in the modern international system functions not merely as a protective shield against external interventions, but rather imposes a positive responsibility upon states to ensure the welfare and security of their own populations (Deng, 2010, 354-355). The R2P doctrine has redefined state sovereignty by shifting it away from Jean Bodin's classical interpretation of absolute and inviolable authority, reconceptualizing it instead as a sphere of responsibility inherently linked to the duty to protect the population. Gareth Evans, the former co-chair of the ICISS and one of the primary architects of the doctrine, articulates this transformation in the following terms:

The issue is not the right of states to intervene, but rather the responsibility of states to protect their own people from crimes of mass atrocity and the responsibility of the international community to assist them in this regard. This shift is a transformation from sovereignty as control to sovereignty as responsibility. (Evans, 2008: 42).

The R2P concept is criticised from many angles. The first criticism concerns the fact that, although it is referred to in UN reports or documents, it is not a binding international legal norm. The absence of an international agreement that explicitly refers to R2P and its conflict with certain customary international law principles, such as sovereign equality among states and non-interference in internal affairs, has been the focus of criticism regarding this lack of legal norms (Borgia, 2015: 228). Another point of contention is that R2P has yet to establish a standard of success or implementation in humanitarian crises. The fact that the UNSC acts within the framework of national interests in international humanitarian crises and does not

grant authorisation to intervene in one crisis while refusing to do so in another crisis of similar severity forms the main argument of such critical studies (Alexander, 2024). Furthermore, the fact that the concept of R2P is discussed more than humanitarian values in situations where humanitarian crises occur is another criticism levelled at the doctrine (Illingworth, 2024: 185). The final point of criticism is that the doctrine legitimises the use of force by citing humanitarian objectives. These criticisms emphasise that the doctrine is used in the same sense as military intervention, which it envisages as a last resort in the response phase (Massingham, 2009: 804).

Cyber, Cyberspace and Cyber Humanitarian Intervention

The concept of cyberspace has been expressed in many different ways in the literature, such as anything related to computers/the internet or a virtual reality, but no common definition has been agreed upon. The concept's limitless and multi-layered structure has led to it being called *cyberspace*. Nezir Akyeşilmen (2018a:54-55) has stated that in order to conceptualise cyberspace, it is necessary to identify all its elements. According to Akyeşilmen, cyberspace essentially consists of four elements. These are: the actor *human* who uses the internet/computer, which is the environment of virtual space, and who creates, destroys or disseminates the information/data found there; *information*, which contains elements such as images, videos or text developed within the virtual framework; the virtual language, i.e. the *logical framework* (software) created with code prepared according to a specific protocol, and the *physical infrastructure* (hardware), from computers to cables or other service providers, which enables the formation of this logical framework.

Cyberspace is frequently discussed in International Relations (IR) literature alongside concepts such as *cyber attack*, *cyber warfare*, or *cyber security*. When examining the main arguments of these studies, the focus is generally on whether reciprocal cyber attacks can be labelled as warfare, and if cyber warfare exists, whether it is similar to or different from traditional warfare. Consequently, within the discipline, one can observe either a reductionist approach or an approach that attaches excessive importance to concepts with the prefix “cyber”. For example, Thomas Rid (2011: 5-7) defines cyber attacks not as warfare but as actions that can be used for destruction, espionage, and sabotage. focusing on the deadly nature of war and its character as a means to political ends, as described in Clausewitz's On War, and considers it unlikely that cyber will acquire the nature of war in the past, present or future. In the literature, there are views that the idea of cyber capabilities being used as an

absolute weapon is pessimistic, that very few cyber attack outcomes translate into political impact, and therefore the use of cyber capabilities will not be widespread (Liff, 2012: 426). John Stone (2012: 106-107), however, responds to Rid's arguments by emphasising that war involves power and violence but does not necessarily result in death, and states that cyber warfare is possible, referring to cyber as an unusual phenomenon.

While the ontological status of cyber warfare as a distinct phenomenon of armed conflict remains a subject of scholarly contention within the discipline, the strategic significance accorded to cyberspace by sovereign states continues to intensify. The fundamental reason for this is that cyberspace essentially encompasses *information*. Andrey Kokoshin, former Deputy Defence Minister of Russia, defined cyberspace as a way to render the opponent's command and control systems ineffective through misinformation, highlighting its strategic and operational aspects (Thomas, 2014: 103). It can be said that today's states are *information-based actors*. They analyse and attempt to solve problems related to their governance by gathering information. Individuals also need information, or data, from states, ranging from social security rights to justice, agriculture to weather data (Balkin, 2012: 4).

Rapid developments in information and communication technologies have integrated the internet, computers, smartphones and social media into every aspect of life. While these developments have significantly facilitated access to information, they have also brought about certain negative consequences. Particularly in digital and chaotic environments where individuals' rational and instinctive thinking abilities weaken in the face of complex situations, and where excessive and diverse information flows prevail, mental shortcuts aimed at reducing cognitive load have begun to be used. This situation makes it easier to change or direct the perceptions of individuals and societies. Regardless of their objectives, various actors can exploit this vulnerability to wage a kind of ‘information war’ through self-serving propaganda or false content (Lin, 2019: 189).

The boundless and largely anarchic nature of cyberspace makes the principles of cyber governance more essential than ever today. Cyber governance emphasises that cyberspace is not merely a technical infrastructure domain; it is also an integral part of the global governance paradigm that encompasses strategic objectives such as respect for human rights, the rule of law, and the establishment of online democracy. In this context, cyber governance serves as an effective safeguard and refuge for the protection of fundamental rights and freedoms (Akyeşilmen, 2018b, 2-5).

However, the digital age has also generated new threat domains that facilitate interference in democratic processes by both state and non-state actors and challenge the fundamental values of democratic societies. Among the most prominent of these threats are the sabotage of democratic electoral processes, the dissemination of violent content, and the manipulation of public opinion. Allegations of Russian interference in the 2016 United States presidential elections, state-sponsored cyber operations such as the Stuxnet and Sony attacks, and the decision of the Australian government to prevent Huawei from participating in the country's 5G infrastructure constitute notable examples of how cyberspace can be exploited by states for malicious purposes (Paterson, 2020: 439–440). In addition to states, hacker groups such as Anonymous—lacking a centralized authority, a coherent ideology, or a fixed objective—also engage in activities within cyberspace that influence states and societies. These groups are particularly known for actions such as releasing leaked materials, gaining unauthorized access to the data of global security firms, and disrupting the websites of multinational corporations (Uitermark, 2017: 403). Moreover, cyberspace is extensively utilized by global terrorist organizations. For instance, the Islamic State of Iraq and Syria (ISIS) has recruited militants from various countries through social media-based propaganda campaigns; in this regard, El-Ravi (2016: 744) notes that the organization increased its global visibility by disseminating positive content in multiple languages that emphasized charitable activities toward the elderly and portrayed everyday life as sustainable in the cities under its control.

Data is as threatening as bullets and bombs (Pellerin, 2011). In an era where bombs are guided by GPS systems and war vehicles are equipped with massive amounts of data, neglecting cyberspace represents a major security vulnerability for the international order in terms of the risks it poses, and a significant loss in terms of opportunities (Roscini, 2014: 2). *Cyber humanitarian intervention (CHI)* is also a concept that is quite important in this regard and should not be neglected. CHI can be defined as interventions using preventive cyberspace to prevent repressive regimes from committing crimes against humanity, such as genocide, ethnic cleansing, and discriminatory violence, against their own societies or against the people of another state (Güler, 2015: 139). Considering the dependence of the perpetrators of such crimes on digital platforms and online networks in directing their actions, planning, or seeking support today, the necessity of CHI becomes apparent.

Although R2P is an important principle in the UN, one of the main reasons it remains ineffective in the face of systematic human rights violations today is that the principle is often associated with a military response. However, if R2P can be implemented without the use of military force, preventing the crisis from escalating would result in a less complex process for both the target country and the intervening states. The UN has emphasised the importance of the prevention phase by publishing a report entitled *Framework of Analysis for Atrocity Crimes: A Tool for Prevention*. The report states that mass human rights violations generally occur in countries experiencing a certain level of instability or crisis, and that preventing the crisis from escalating to the point of requiring military intervention could avert not only loss of life but also physical, psychological and social trauma. On the other hand, the report states that the cost of prevention is lower than the cost of continuing crises and evaluates the limited options for preventive action (UN, 2012: 2). Therefore, one of the most important issues neglected in the literature on R2P is the question of what preventive interventions might be. Considering the negative aspects of technology that facilitates, deepens and covers up the aforementioned human rights violations, it may be appropriate to evaluate CHI as an antidote for preventive purposes.

One of the most fundamental operations of the CHI is undoubtedly to provide uninterrupted digital access to information in crisis areas. The ability to securely transmit and receive data, coordinate actions in real time, and maintain situational awareness in large and complex crisis areas is the cornerstone of the modern digital world. Without secure and resilient communications, even the most advanced autonomous systems and AI-powered platforms become isolated, vulnerable entities. Considering the possibility that perpetrators may deliberately damage communication infrastructure to avoid repercussions for their actions, it is essential that affected communities have access to internet-based communication channels to make their voices heard, demonstrate the depth of the crisis to the global public, and provide evidence of the elements of the actions.

The internal conflicts that took place in Libya in 2011 and ended Muammar Gaddafi's nearly half-century rule with his death are an example of the regime's blocking of communication channels. The regime had always sought to maintain its monopoly over the internet, blocking websites that produced content inconsistent with its policies or that were critical of it, and imposing harsh penalties on individuals who made critical comments. As a result of the

crackdown, internal unrest began in February 2011, leading to an internet blackout that lasted until August 2011. In Libya, only 17 per cent of the population had access to the internet due to high internet costs, while mobile phone ownership was widespread among almost the entire population. Consequently, the regime not only cut off internet access but also restricted access to CHI cards. After the regime's collapse, archives were found containing files on the online activities of Libyan dissidents communicating with foreigners (Freedom House, 2012).

Syria is another country where the cyber domain is controlled by regime leader Bashar al-Assad through public institutions such as the Syrian Telecommunications Establishment (STE). Although the number of Syrian citizens with internet access reached 4 million in 2010, the regime has always monitored user activity and required businesses such as internet cafes to record customer information and online activities. STE has utilised advanced technologies to block the public's phone calls, text messages, emails and internet access. In 2007, the regime introduced a nationwide surveillance system capable of actively monitoring the internet without individuals' knowledge, resulting in the procurement of devices capable of network filtering, blocking, and surveillance (Helwani, 2024: 249).

It is possible to multiply the policies implemented by repressive regimes to prevent their citizens from communicating with the outside world. What is important here is what steps countries that desire peace and wish to prevent crises will take in the face of repressive regimes. To prevent the blocking of the internet and other communication infrastructures in crisis areas by regime interventions and to ensure the healthy exchange of information, *satellites* can be considered within the scope of CHI. Currently, many states use this satellite technology within the scope of national cyber security. For example, the United States uses satellites for observation, communication and mapping. The Russia-Ukraine war has also highlighted the importance of satellites. In the war that began in February 2022 with Russia's invasion of Ukraine, Russian cyber attacks dealt a heavy blow to Ukraine's communications infrastructure, causing serious communication disruptions between army units and rendering military equipment that required network connectivity unusable. At this point, help for the Ukrainian army came from Starlink, the world's largest satellite constellation owned by SpaceX. With more than 20,000 Starlink terminals provided to Ukraine, the satellite became an indispensable communication infrastructure for the army (Abels, 2024: 843). Furthermore, Starlink satellites were utilised during the Los Angeles wildfire that began on 7 January 2025, replacing the damaged internet and communication infrastructure to ensure both firefighting

teams remained in contact and national and international media could broadcast from the disaster area (Conklin, 2025).

The second type of operation related to CHI could be the implementation of applications for *online surveillance* in crisis areas. Surveillance applications can be used to collect data from crisis areas in a digital environment and to identify threats and risks. Such applications include technologies such as advanced cyber intelligence systems, artificial intelligence-powered data analytics, and satellite imaging. Surveillance operations are critical in identifying atrocities faced by civilians, documenting human rights violations to increase the accountability of perpetrators in international courts, and enabling rapid intervention by the international community. In this context, for example, Zhengyang Hou and colleagues (2024: 1) use image processing techniques to detect destruction in civil war zones using satellite imagery, converting image pixels into information with an application they call PtNet and presenting it through a detection scheme called TKDS. The authors emphasise that real-time detection of damage that may occur in current and future countries due to civil unrest, earthquakes or extreme weather events is of vital importance.

During the Cold War, the primary purpose of surveillance satellites, whose importance grew, was to detect and classify rival states' nuclear-tipped missiles or submarines, warplanes, military equipment, and other communication infrastructure. However, with the technology of the previous century, images were exposed and captured on film, and it took days for the film rolls to reach experts and for the films to be developed. With the digital era, film rolls have been replaced by surveillance technologies with digital sensors that continuously capture images. In the following period, imaging radars with higher resolution capabilities, able to focus on a target, detect different radiation levels in the monitored area, and scan a wider area, were developed, such as Germany's SAR-Lupe, Italy's COSMO-SkyMed, Israel's TecSAR, China's YaoGan, and India's Cartosat-2. (Norris, 2011: 44-46).

The third CHI method could involve preventing social media and other internet-based posts containing hate speech and violent content in order to break the perpetrator's digital assault, and ensuring that supportive content for the victim is included. In the digital age, hate speech content increases social polarisation and can be a powerful factor capable of triggering crises of violence in societies. Violent rhetoric spreading through social media applications radicalises individuals, can lead to increased othering of minorities, and can disrupt social harmony. Therefore, blocking such content through cyber intervention is necessary to prevent

crises from escalating. Content blocking within the framework of CHI must be carried out with great care, as it treads a fine line between freedom of expression and the preservation of peace. CHI, which is not a censorship mechanism that restricts freedom of expression, should aim to combat disinformation, identify digital environments that encourage hate speech, and raise awareness both in victimised communities about the crisis and in other communities around the world about victimised communities. Digital literacy enhancement education programmes can also be beneficial in this regard within the scope of CHI.

The effect of violent content spreading rapidly on social media, thereby deepening crises, is clearly evident in the crimes against humanity committed by the Myanmar army against the Rohingya minority in 2017. The Myanmar army launched an ethnic cleansing operation against Rohingya Muslims, while Facebook, a social media application owned by Meta Technologies, encouraged and reinforced this ethnic cleansing with its algorithms. Radical Buddhist nationalist groups and Myanmar army personnel spread a great deal of misinformation on the app, claiming that Muslims would take over Myanmar as invaders in the near future. They shared photos of human rights activists defending the rights of the Rohingya people within the Myanmar population and threatened them with death. In a report published in 2022, Amnesty International acknowledged that Meta contributed to the atrocities in Myanmar with its dangerous algorithms for profit. Rohingya activist Mohammed Showwife accused Facebook of destroying the dreams of the Rohingya people, who aspire to live like everyone else (Amnesty International, 2022).

Violent content is not limited to social media. Looking further back in history, the 1994 Rwandan genocide confronts humanity. In attacks carried out by Hutus against Tutsis, approximately one million Rwandans were killed and two million people were forced to flee their country. In Rwanda, where two ethnic groups had coexisted peacefully in the past, the fact that ordinary civilians attempted to kill each other with any object they could find highlights the role of communication tools in triggering the genocide. After the death of President Habyarimana in a plane crash, the Hutus were gripped by the fear that the Tutsis would seize power and begin discriminatory activities. During this period, the Hutus used the radio to incite and direct the genocide. Radio broadcasts via Radio Télévision Libre des Milles Collines, calling on other Hutus to take action against the Tutsis, were constructed around memories such as Rwanda's colonial history, suggesting that the only way out of this cycle of the past was through genocide, triggering absolute violence between the two ethnic groups (Kellow and Steeves, 1998: 107).

The final alternative type of operation related to CHI may involve targeted cyber interventions aimed at terminating the perpetrators' actions. By incorporating both traditional warfare techniques and modern cyber space elements, it can directly damage the perpetrators' communication channels. As this method resembles a military intervention rather than a preventive measure, it also carries the risk of harming civilians. The detonation of radios used by Hezbollah by Israel on 18 September 2024 (Aljazeera, 2024), the 2010 Stuxnet Operation by the US targeting Iran's nuclear facilities, which damaged one-fifth of the gas centrifuges (Willett, 2024: 69), or Iranian hackers attempting to infiltrate rural water flow and wastewater treatment systems in Israel (Heller, 2020) are examples of direct, targeted cyber actions. Due to the potential for direct or indirect harm to civilians, their implementation is highly challenging.

The most successful example of a targeted cyber operation is Operation Glowing Symphony, conducted in 2016 by the US Cyber Command and the US National Security Agency. The primary objective of the operation was to target ISIS's global media operations and propaganda, destroying materials and disrupting its digital recruitment and financial activities. As part of the operation, ten accounts used by the organisation to spread its propaganda were listed and phishing emails were used. This allowed the operation teams to gain control, enabling them to freely navigate ISIS networks and plant malicious software on servers. First, ISIS networks were mapped, propaganda content was removed, and the organisation's propaganda methods, such as the Amaq Agency app, were blocked. The teams then moved on to creating technical errors and problems that would often appear to be IT issues, creating a psychological effect within the organisation, such as confusion, anger and deception. This forced the organisation's digital managers to use vulnerable and unreliable tools that would reveal their physical locations, making them targets for kinetic attacks (Raston, 2019: Cohen and Bar'el, 2017: 36).

There are fundamental similarities and differences between CHI and Traditional Humanitarian Intervention (THI). Firstly, both CHI and THI are based on international norms and aim to protect humanitarian values, relying on the obligation of states or international actors to intervene in the face of systematic human rights violations or crimes of mass atrocity. Due to the lack of sufficient interest in CHI in the literature on international law and international relations, there is no study that comprehensively outlines the differences and similarities between CHI and THI. However, considering the similarities and differences noted by Kallberg (2016: 84), it can be seen that the cyber warfare-conventional warfare

distinctions frequently studied in the disciplines are in line with the characteristics of CHI and THI.

CHIs, like cyber conflicts or wars, differ from THIs in terms of the environment in which they are created, the techniques and tools used, and their strategies. Firstly, CHIs promise to have a significant impact in weakening the pressure mechanisms of regimes or other criminal actors with their sudden and unexpected operational capacity. Since any operation undertaken using THI methods will require a specific preparation process, perpetrators can take a defensive position in the event of diplomatic signals or military movements. CHI breaks the control mechanisms of perpetrators, particularly those based on communication and information gathering, ultimately undermining their capacity to maintain pressure. Furthermore, physical boundaries are of no significance for CHI (Healey, 2016: 44-45).

Another fundamental characteristic that distinguishes CHI from THI is that human interventions in the digital environment involve much lower costs and risk rates compared to conventional methods. While military or direct kinetic interventions typically require significant economic investment, logistical support, and extensive operations, digital interventions can yield effective results even with limited resources. For example, according to a report by Richard Norton Taylor and Peter Capella (1999) in The Guardian, the cost of NATO's Operation Allied Force intervention in Kosovo exceeded £30 billion. While military operations involve numerous complex processes, such as the logistics of military units, the maintenance of air operations, and ensuring the safety of the civilian population in the region during the operation, CHI generally requires software-based strategies and thus requires fewer material resources. For example, providing a global VPN or encryption tools against the censorship practices of an oppressive regime is much less costly than air operations. Furthermore, since CHI does not require a physical presence in crisis areas, it does not carry the risk of conflict. In THI, intervention forces must be present on the ground and may therefore face situations such as becoming direct targets or being exposed to retaliatory attacks (Li and Liu, 2021: 8183-8184).

One of the fundamental reasons why CHI is less costly than THI is that CHI can be implemented by a much wider range of actors. THI is carried out by large, centralised and fixed actors, based on states sending their troops to military coalitions formed under the umbrella of international organisations. In CHI, however, in addition to the states that will implement digital interventions, there is the possibility that individuals, civil society

organisations, private companies and hacker groups may take part under the supervision of international and regional organisations. This diversity gives CHI the characteristics of speed, flexibility and low cost, while at the same time raising questions about who the implementing actor for CHI will be.

Any CHI action undertaken to weaken and eliminate the perpetrators' means of coercion in crisis regions should be the responsibility of states and regional/global organisations. The concentration of power in cyberspace by actors such as private companies or hackers, without state or organisational oversight, carries the risk of drawing states into a conflict zone, raising concerns that such cyber interventions will contribute to international instability rather than peace (Pattison, 2020: 251). However, states or organisations may employ private technology companies or individuals with cyber capabilities to carry out humanitarian interventions on their behalf. For example, the United States obtained surveillance opportunities during the Kosovo War through a contract with DynCorp, a private military contractor, to monitor the withdrawal of Serbian military forces from Kosovo. As can be seen, states or organisations can form a cyber humanitarian intervention team under their own identity, but they can also utilise the private sector and, in some cases, even opt for a hybrid structure (Rhiannon, 2021: 187).

The ethical and moral assessment of CHI's implementing actors within the context of international law is also important. In international law, the principle of jus ad bellum specifies when and under what conditions a state or the international community may legitimately resort to the use of force. Conventionally, the use of force between states is prohibited under Article 2(4) of the UN Charter, and this provision has become a jus cogens norm. The only exceptions to the prohibition on the use of force are situations approved by the UN Security Council and situations involving elements of legitimate defence. However, when considering CHI, it is possible that it could be evaluated within the framework of the jus ad bellum principle, as it does not involve the use of physical force, unlike traditional military interventions, and is carried out with the aim of deterring human rights violations by repressive regimes.

Another fundamental principle of international law is *jus in bello*, meaning that during wartime, the rules governing the conduct of war require that combatants respect human rights and civilians. Even if CHI does not involve kinetic operations but rather activities such as data collection or countering disinformation, the level and form of intervention must be

proportionate. For example, CHI actions that affect the health system or basic infrastructure of the country being intervened in may be considered an unacceptable violation under International Humanitarian Law. In this context, the principle of jus in bello requires careful consideration to ensure that CHI operations only target the perpetrators of atrocities and do not harm civilians.

From an ethical perspective, CHI can serve as a deterrent against human rights violations by oppressive regimes and can provide critical evidence for international justice mechanisms. However, the risk of such interventions being misused should not be overlooked. Cyber operations conducted unilaterally, particularly by certain states or international organisations, can be manipulated for political gain, even if they are claimed to be carried out for humanitarian reasons. Therefore, international oversight mechanisms and transparency principles must be implemented to ensure that CHI maintains its legitimacy within a legal and ethical framework.

Conclusion

This study has examined whether a cybersecurity-based humanitarian intervention can be situated within the framework of the Responsibility to Protect (R2P), addressing the question not merely at the level of technical feasibility or operational effectiveness, but through its normative, legal, and ethical constraints. The central finding of the analysis suggests that while the use of cyber technologies in humanitarian intervention may appear theoretically possible, such an approach remains in significant tension with the normative foundations upon which R2P is built. Consequently, the issue is less about whether a cyber intervention can be conducted, and more about whether such an intervention can be defined as legitimate, constrained, and genuinely protective within the scope of R2P.

The R2P doctrine conceptualizes the prevention of atrocity crimes as a collective responsibility, yet it deliberately leaves unresolved the question of which instruments may be legitimately employed to fulfil this responsibility. Proposals for cyber humanitarian intervention draw upon this ambiguity, presenting cyber operations as a seemingly less intrusive alternative to traditional military intervention and as a means of avoiding the political and humanitarian costs associated with kinetic force. However, this study demonstrates that the inherent characteristics of cyber operations—namely their opacity, difficulties of attribution, and indeterminate scope—risk undermining rather than reinforcing the principles of legitimacy, transparency, and accountability that R2P seeks to uphold.

A critical dimension of the research question concerns whether cyber humanitarian intervention genuinely serves the protection of civilians, or whether it merely transforms intervention into a more invisible and less regulated practice. Cyber operations may indeed contribute to civilian protection through early warning systems, information gathering, or the documentation of violations. Yet the same tools can easily be repurposed for operations that infringe upon state sovereignty, disrupt critical infrastructure, or generate wide-ranging indirect effects on civilian populations. This raises unresolved questions regarding how core R2P principles such as last resort and proportionality can be meaningfully applied in the cyber domain.

The study further reveals that the normative vacuum surrounding the legal status of cyber operations in international law effectively shifts cyber humanitarian intervention from a rule-based framework into a realm of political discretion. Given that R2P practices remain contested even in the context of conventional interventions, their extension into cyberspace - an arena characterized by blurred boundaries and limited accountability- risks facilitating the normalization of intervention under increasingly permissive conditions. In this sense, cyber humanitarian intervention may be interpreted not as an evolution of R2P, but as an indicator of its normative erosion.

Accordingly, the answer to the research question must be cautious and conditional. While cybersecurity-based intervention under R2P may be technically conceivable, it is difficult to argue that such interventions can presently be considered genuinely “humanitarian” within the existing international legal order and prevailing power structures. On the contrary, the discourse of cyber humanitarian intervention may function to lower the threshold for intervention and weaken mechanisms of accountability, particularly in the context of the digital reconfiguration of sovereignty.

In conclusion, rather than framing cyber humanitarian intervention as a normative advancement, this study positions it as a contested domain that exposes the inherent limitations and contradictions of the R2P doctrine. The incorporation of cyber technologies into humanitarian intervention can only be justified under conditions of clearly articulated norms, robust oversight mechanisms, and genuinely collective decision-making processes. Absent these safeguards, cyber humanitarian intervention risks becoming a legitimizing discourse for new, less visible forms of intervention, rather than a meaningful instrument for the protection of civilians.

References

Abels, J. (2024). Private Infrastructure in Geopolitical Conflicts: The Case of Starlink and The War in Ukraine. *European Journal of International Relations*, *30*(4), 842–866.

Akyeşilmen, N. (2018a). *Disiplinlerarası Bir Yaklaşımla Siber Politika & Güvenlik*. Ankara: Orion Kitabevi.

Akyeşilmen, N. (2018b). Cyber Good Governance: A New Challenge in International Power Politics?. *Cyberpolitik Journal*, *3*(5-6), 2-21.

Al Jazzeera. (2024). Hezbollah Walkie-talkies Exploded Too, What to Know about Israel's Attacks. https://www.aljazeera.com/news/2024/9/18/more-devices-exploding-across-lebanon-whats-happening

Alexander, K. (2024). The limits of international law: the Responsibility to Protect (R2P), Israel and the International Court of Justice. https://www.realinstitutoelcano.org/en/commentaries/the-limits-of-international-law-the-responsibility-to-protect-r2p-israel-and-the-international-court-of-justice/

Al-Rawi, A. (2016). Video Games, Terrorism, and ISIS's Jihad 3.0. *Terrorism and Political Violence*, *30*(4), 740–760.

Amnesty International. (2022). *Myanmar: Facebook's systems promoted violence against Rohingya; Meta owes reparations – new report*. https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/ (Erişim Tarihi: 17.01.2025)

Badescu, C. G. (2011). *Humanitarian Intervention and the Responsibility to Protect: Security and Human Rights*. Abingdon: Routledge.

Balkin, J. M. (2012). The First Amendment is an Information Policy. *Hofstra Law Review*, *41*(1), 1-41.

BM. (2012). Framework of Analysis for Atrocity Crimes A Tool for Prevention. https://www.globalr2p.org/resources/framework-of-analysis-for-atrocity-crimes-a-tool-for-prevention/

Coady, C. A. J., Dobos, N. and Sanyal, S. (2018). *Challenges for Humanitarian Intervention: Ethical Demand and Political Reality*. Oxford: Oxford University Press.

Cohen, D. and Bar'el, O. (2017). The Use of Cyberwarfare in Influence Operations. *Yuval Ne'eman Workshop for Science, Technology and Security*. https://en-cyber.tau.ac.il/sites/cyberstudies-english.tau.ac.il/files/media_server/cyber%20center/cyber-center/Cyber_Cohen_Barel_ENG.pdf

Conklin, A. (2025). Los Angeles Wildfires: Elon Musk Personally Delivers Starlinks to California First Responders. https://www.foxnews.com/us/los-angeles-wildfires-elon-musk-personally-delivers-starlink-satellites-california-first-responders

Deng, F. (2010). From Sovereignty as Responsibility' to the Responsibility to Protect. *Global Responsibility to Protect*, *2*(4), 353-370.

Evans, G. (2008). *The Responsibility to Protect: Ending Mass Atrocity Crimes Once and For All*. Washington: Brookings Institution Press.

Freedom House. (2012). Freedom on the Net 2012 – Libya. https://www.refworld.org/reference/annualreport/freehou/2012/en/88761

Gilligan, E. (2013). Redefining Humanitarian Intervention: The Historical Challenge of R2P. *Journal of Human Rights*, *12*(1), 21–39.

Gulati, J. and Khosa, I. (2013). Humanitarian Inter Humanitarian Intervention: Tention: To Protect State So otect State Sovereignty. *Denver Journal of International Law & Policy*, *41*(3), 397-416.

Güler, A. (2015). İnsani Müdahale Aracı Olarak Siber Uzay. *Medeniyet Araştırmaları Dergisi*, *2*(4), 139-149.

Healey, J. (2016). Winning and Losing in Cyberspace. N.Pissanidis, H.Rõigas, M.Veenendaal (Eds.). *8th International Conference on Cyber Conflict: Defending the Core*. Talinn: CCD COE.

Heller, A. (2020). Israeli cyber chief: Major attack on water systems thwarted. https://www.aronheller.com/articles/israeli-cyber-chief-major-attack-on-water-systems-thwarted/

Helwani, I. (2024). Cyberactivism in Syria: Emergence, Transformation, Potentials, and Limitations. *Güvenlik Stratejileri Dergisi*, *20*(48), 239-263.

Henkin, L. (1999). Kosovo and The Law of "Humanitarian Intervention". *American Journal of International Law*, *93*(4), 824-828.

Hou, Z., Qu, Y., Zhang, L., Liu, J., Wang, F., Yu, Q., ... and Zhou, C. (2024). War City Profiles Drawn from Satellite Images. *Nature Cities*, *1*(5), 1-11.

ICISS. (2001*). The Responsibility to Protect*. https://www.globalr2p.org/resources/the-responsibility-to-protect-report-of-the-international-commission-on-intervention-and-state-sovereignty-2001/

Illingworth, R. (2024). Not the 'Fairest Norm of Them All' but Still Needed: On Hobson and Criticism of the Responsibility to Protect. *Journal of Intervention and Statebuilding*, *18*(2), 181–190.

Kallber, J. (2016). Humanitarian Cyber Operations. IEEE Tchnology and Society Magazine, https://cyber.army.mil/Portals/3/Documents/publications/external/Humanitarian%20Cyber%20Operations.pdf?ver=2017-09-26-135216-070

Kardaş, Ş. (2003). Humanitarian Intervention: A Conceptual Analysis. *Alternatives: Turkish Journal of International Relations*, *2*(3&4), 21-49.

Kellow, C. L and Steeves, H. L. (1998). The Role of Radio in the Rwandan Genocide. *Journal of Communication*, *48*(3), 107-128.

Li, Y. and Liu, Q. (2021). A Comprehensive Review Study of Cyber-Attacks and Cyber Security; Emerging Trends and Recent Developments. *Energy Reports*, *7*, 8176-8186.

Liff, A. P. (2012). Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War. *Journal of Strategic Studies*, *35*(3), 401–428.

Lin, H. (2019). The Existential Threat from Cyber-Enabled Information Warfare. *Bulletin of the Atomic Scientists*, *75*(4), 187–196.

Massingham, E. (2009). Military Intervention for Humanitarian Purposes: Does The Responsibility to Protect Doctrine Advance the Legality of the Use of Force for Humanitarian Ends? *International Review of the Red Cross*, *91*(876), 803-831.

Neilsen, R. (2021). Cyber Humanitarian Interventions: the Viability and Ethics of Using Cyber-operations to Disrupt Perpetrators' Means and Motivations for Atrocities in the Digital Age. PhD Thesis, UNSW Australia.

Norris, P. (2011). Developments in High Resolution Imaging Satellites for the Military. *Space Policy*, *27*, 44-47.

Norton, R. and Capella, P. (1999). *Bill for Kosovo war goes over £30bn*. https://www.theguardian.com/world/1999/oct/15/balkans

Paterson, T. and Hanley, L. (2020). Political Warfare in The Digital Age: Cyber Subversion, Information Operations and 'Deep Fakes.' *Australian Journal of International Affairs*, *74*(4), 439–454.

Pattison J. (2020). From Defence to Offence: the Ethics of Private Cybersecurity. *European Journal of International Security*, *5*(2): 233-254.

Pellerin, C. (2011). *Dod Releases First Strategy for Operating in Cyberspace*. https://www.army.mil/article/61720/dod_releases_first_strategy_for_operating_in_cyberspace

Raston, D. T. (2019). How the U.S. Hacked ISIS. https://www.npr.org/2019/09/26/763545811/how-the-u-s-hacked-isis?t=1582468821601

Rid, T. (2011). Cyber War Will Not Take Place. *Journal of Strategic Studies*, *35*(1), 5–32.

Roberts, G. W. (2000). Humanitarian Intervention: Definitions And Criteria. *CSS Strategic Briefing Papers*, *3*(1), 1-2.

Roscini, M. (2014). *Cyber Operastions and The Use of Force in International Law*. Oxford: Oxford University Press.

Stone, J. (2012). Cyber War Will Take Place! *Journal of Strategic Studies*, *36(*1), 101–108.

Thomas, T. (2014). Russia's Information Warfare Strategy: Can the Nation Cope in Future Conflicts? *The Journal of Slavic Military Studies*, *27*(1), 101–130.

Uitermark, J. (2016). Complex Contention: Analyzing Power Dynamics within Anonymous. *Social Movement Studies*, *16*(4), 403–417.

United Nations General Assembly. (2009). Implementing the Responsibility to Protect. *Report of the Secretary-General*, https://documents.un.org/doc/undoc/gen/n09/206/10/pdf/n0920610.pdf

Wheeler, N. J. (2000). *Saving Strangers: Humanitarian Intervention in International Society*. Oxford: Oxford University Press.

Willett, M. (2024). A Short History of Cyber Operations. *Adelphi Series*, *64*(511–513), 63–104.