KONRAD ADENAUER STIFTUNG | UN LIAISON OFFICE NEW YORK

REGULATING THE ROLE AND INVOLVEMENT OF OFFENSIVE PROXY ACTORS IN CYBERCONFLICT

By Eleonore Pauwels

ISBN: 978-1-7369528-4-9

TABLE OF CONTENTS

PREFACE

With Russia's extensive use of hybrid warfare techniques and cyberattacks in the war in Ukraine, the increasing prevalence of cyberoperations in armed conflicts and geostrategic rivalries has garnered enhanced attention. However, the phenomenon as such is not as new as it seems. Digital capacities enhanced by artificial intelligence (AI) and the instrumentalization of cyberspace are utilized by state and non-state actors in their struggle to advance geopolitical goals or business interests. Amid the complexities of hybrid warfare and cyberconflict, the present analysis by Eleonore Pauwels draws particular attention to the technical, legal, and normative dimensions of offensive proxy actors in cyberconflict.

Cyberattacks are not only a challenge in relation to attribution and accountability. They also reveal the gaps and weaknesses in established international norms and legal frameworks, and most prominently, the powerlessness of international humanitarian law.

However, it is an encouraging sign to see a growing awareness among United Nations (UN) Member States. A sense of urgency to address the unregulated actions of offensive proxy actors is developing among the international community.

For two decades (starting in 2004), engagement on cybersecurity at the UN level mainly took place in the Group of Governmental Experts (GGE) on Advancing responsible state behavior in cyberspace in the context of international security. Since 2019, the GGE has been complemented by an Open-ended Working Group (OEWG) on security of and in the use of information and communications technologies (see also the comparative analysis of both groups in *Multilateralism and the Rising Challenges of Global Insecurity*). Both processes are mandated by the General Assembly, and thus, not legally binding.

Although the Security Council took up cyberthreats to international peace and security rather late, discussions have intensified over the last three years. Inter alia, an Arria-formula meeting has been co-organized by Albania and the United States (U.S.) on "The Responsibility and Responsiveness of States to Cyberattacks on Critical Infrastructure." Most recently, the United Kingdom (UK) hosted a high-level briefing on "Artificial Intelligence: Opportunities and Risks for International Peace and Security."

While all of these are encouraging developments, they come rather late in the game and are mainly attempts to catch up with the rapidly evolving phenomena of offensive proxy actors in cyberconflicts. Therefore, we hope that the present study not only illustrates the seriousness of these threats to international security and the resilience of societies, but also provides a set of useful recommendations for furthering the political norm-setting process.

KAS New York wishes you an interesting read!

Andrea Ostheimer
Executive Director
Konrad Adenauer Foundation (KAS)
New York Office

EXECUTIVE SUMMARY

Geopolitical tensions among great powers have intensified over the past decade, ushering in an era of multipolar competition. Traditional geopolitical risks are colliding with new, complex challenges of the 21st century, such as the rise of powerful non-state actors in cyberconflict. In a 2021 report, the International Committee of the Red Cross (ICRC) highlighted that "some non-State actors have the potential to deliver effects through cyber tools comparable to or exceeding those available to many States."[1]

Cyber offense has become a powerful business, a pervasive threat, and global in scope. Research for the present report started with an alarming diagnosis: the rapid proliferation, commoditization, and privatization of offensive cyber capabilities by proxy actors with potentially devastating consequences for international peace and security.[2] What happens when private security firms, rogue non-state actors, and transnational cybercriminal groups exploit and trade cyberweapons as powerful as those used by tech-leading nations? There is a potential for escalation, hybrid warfare, and a sharp decline in world order and in security afforded to civilian populations and industries. The human cost of offensive cyberoperations may become disproportionate.[3]

The stakes are high. Very few strategic tools and legal responses have been used successfully to deter hostile non-state actors in cyberspace and hold them accountable for the harm they impose on societies across the globe. Cyber proxies tend to be under-conceptualized in international legal frameworks and, to some extent, occupy and thrive in a "normative safe zone,"[4] an ungoverned space. **Thus, the present report analyzes the technical, legal, and normative dimensions of the cyber proxy phenomenon as it manifests today in its offensive forms, both in situations of armed conflict and in advanced geostrategic competition.** The overall purpose of this research project is to help delineate technical and legal challenges and prospects for regulating the role and involvement of offensive non-state actors acting as proxy in cyberconflict.

The scope of the report focuses on the array of cyberthreat actors that conduct or contribute to offensive cyberoperations which align with a state or a group within a state, support a state's national interests, or are tacitly permitted by a state. Cyberconflict can be framed as strategic conflict in cyberspace pursued by actors motivated by geostrategic and competitive economic interests, posing a "threat to national and regional peace and security architectures" and impacting

1 International Committee of the Red Cross (ICRC), *Avoiding Civilian Harm from Military Cyberoperations during Armed Conflict: ICRC Expert Meeting 21–22 January 2020 – Geneva*, 2021, p. 34, https://shop.icrc.org/avoiding-civilian-harm-from-military-cyber-operations-during-armed-conflicts-icrc-expert-meeting-21-22-january-2020-geneva-pdf-en.html.

2 Report of the Working Group on the use of mercenaries as a means of violating human rights and impeding the exercise of the right of peoples to self-determination, "The human rights impacts of mercenaries, mercenary-related actors and private military and security companies engaging in cyberactivities," A/76/151, 15 July 2021. See McGuire, M., *Nation States, Cyberconflict, and the Web of Profit*, 2021, https://threatresearch.ext.hp.com/wp-content/uploads/2021/04/hp-bps-web-of-profit-report_APR_2021.pdf. See also Microsoft, "Response to United Nations (UN) Working Group on the Use of Mercenaries," October 2021, https://www.ohchr.org/sites/default/files/Documents/Issues/Mercenaries/WG/CyberMercenaries/MSFT-Response.pdf. See also Microsoft, *Microsoft Digital Defense Report 2022*, 2022, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE5bUvv?culture=en-us&country=us.

3 See Gisel, L. and Olejnik, L., *The Potential Human Cost of Cyber Operations: ICRC Expert Meeting 14-16 November 2018 – Geneva*, ICRC, 2019, https://www.icrc.org/en/document/potential-human-cost-cyber-operations.

4 See Maurer, T., *Cyber Mercenaries: The State, Hackers, and Power* (Cambridge: Cambridge University Press, 2018), p. 129.

> What happens when private security firms, rogue non-state actors, and transnational cybercriminal groups exploit and trade cyberweapons as powerful as those used by tech-leading nations?

"strategic targets, democratic processes and wider civilian populations."[5] For the purpose of this report, we define "hostile" or "offensive cyber operations" as operations "that alter, disrupt, deceive, degrade or destroy computer systems or networks, or otherwise undermine the confidentiality, integrity and availability of computer systems or networks for individuals and communities."[6] This category of hostile activities in cyberspace can be part of cyber warfare but also occur in adversarial situations that do not clearly meet the threshold of an armed conflict. Cybersecurity, armed conflict, and legal experts use the term "grey zone" to "describe how actors may be using hostile cyber (or other) operations in a harmful manner yet designed to avoid eliciting a conventional military response by the target, including through blurring the identification by the target of the applicable legal framework."[7]

The strategic goals of the report are twofold. First, by providing in-depth evidence on how states increasingly rely on offensive cyber proxy activity, the report aims to inform policymakers, legal experts, civil society, and multilateral institutions about the emerging strategies at play to obfuscate technical and legal attribution and escape accountability. The Technical Section also addresses geostrategic and governance implications, including those relevant to disarmament and non-proliferation regimes. **Second**, by analyzing how international law applies to cyber proxy offensive operations, the report aims to identify legal ambiguities and potential opportunities to support ongoing normative and policy processes at the multilateral level. Two Case Studies focus respectively on situations of armed conflict and grey zone operations. In particular, there is a pressing need to build collective capacity to hold cyber proxies accountable for civilian harm by strengthening mechanisms for prevention, investigation, prosecution, and remedy. The report provides states with governance recommendations for increased normative cooperation grounded in international law, as well as collaboration in non-proliferation and cybercrime prevention.

The Framing Section reviews how cyber proxy activity has been conceptualized in the past and whether these frameworks resist the evolving nature of cyberconflict. Experts in cybersecurity and international relations have long postulated several hypotheses to explain nation states' increasing reliance on cyber proxies. In terms of **political and economic cost management**, proxies in the cyber-arms industry allow states not only to avoid direct military action but also to efficiently acquire and employ capabilities that might be prohibitively expensive to develop in-house within military units. Using cyber proxies may also constitute a winning strategy to benefit from **plausible deniability**. Cyberweapons are stealthy and effective and can be launched below the threshold of war with a high degree of automation, anonymity, and opacity. Their use can be outsourced in multi-stage operations to proxy actors located at arms' length of a nation state and in different jurisdictions. As such, nation states can achieve an "opportunistic dissociation" from the means of cyber aggression—which is much more difficult to do in kinetic situations—and therefore potentially escape attribution of responsibility and liability. The UN Working Group on the use of mercenaries offers another substantial argument that refers to **the state-centric nature of multilateral governance and international law**: "unlike States that are subject to international human rights and humanitarian law protocols, they [cyber proxies] tend to operate outside the purview of such protocols, making attribution, arrests and prosecution difficult."[8]

[5] van der Waag-Cowling, N., "Stepping into the breach: military responses to global cyber insecurity," ICRC Humanitarian Law & Policy Blog, 2021, https://blogs.icrc.org/law-and-policy/2021/06/17/military-cyber-insecurity/.

[6] Resolution adopted by the United Nations General Assembly on 16 December 2021, A/76/151, "Use of mercenaries as a means of violating human rights impeding the exercise of the right of peoples to self-determination," p. 5, https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/401/86/PDF/N2140186.pdf?OpenElement.

[7] ICRC, *Avoiding Civilian Harm from Military Cyberoperations during Armed Conflict*, p.15.

[8] See A/76/151, p. 11-12, para. 37.

The Technical Section sheds light on the harmful convergence between the cyberarms and cybercrime industries and the offensive proxy capacities these industries bring to an increasing number of nation states and potentially violent actors. This section covers, among others, the following questions: How is great power competition shaping the private market of offensive cyber capabilities? How are nation states engaging with the underground cyberarms industry and the interconnected cybercrime economy? How are these recent technical trends impacting the normative framework for responsible behavior in cyberspace? Technical findings unveil the polymorphous and adaptive nature of cyber proxy relationships. In particular, it demonstrates **how an array of nation states is increasingly engaging in various types of knowledge and technological transfer with the underground cyberarms and cybercrime industries to further their geostrategic interests and enable disruptive forms of economic, industrial, and political warfare**. This section uncovers the tacit strategies used by some nation states and their cyber proxies to benefit from both cyberarms and cybercrime, but also to fuel those nefarious industries with the goal to cause and amplify economic harm, chaos, and insecurity in cyberspace. **The increased permeability between the clandestine cyberarms and cybercrime industries and the blurring of their actors' modi operandi severely complicates the production of evidence for legal attribution and regulation across borders and jurisdictions.**

> "Cyber proxies tend to be underconceptualized in international legal frameworks and, to some extent, occupy and thrive in a "normative safe zone," an ungoverned space.

Regulating the role and involvement of proxy actors in the provision of offensive cyberservices requires clarifying the interpretation of and existing legal ambiguities on how international law applies to cyberspace. The **Legal Section** does so by reviewing Article 8 of the International Law Commission's (ILC) Articles on State Responsibility, as well as the different bodies of international law (the UN Charter, international human rights law (IHRL), international humanitarian law (IHL), The Rome Statute, customary international law, and the normative acquis). As explained by the UN Working Group on the use of mercenaries, multilateral discussions need to (a) define, in more precise terms, what constitutes offensive cyber activities including cyberwarfare and cyberattacks; (b) develop agreed methods to identify the source of cyberattacks and other cyber activities and to attribute such attacks or activities to particular persons or entities; (c) qualify, in legal terms, the relationship between the non-state actor and the state on behalf of which such activities are undertaken, if at all; and (d) determine whether particular cyber activities constitute involvement or direct or indirect participation in ongoing hostilities.[9]

Three Case Studies are presented to address: (1) the proliferation of cyber proxies in the ongoing invasion of Ukraine by the Russian Federation and (2) the use of cyber proxies in grey zone ransomware operations on civilian critical infrastructures. A short synopsis will also elaborate the role of influence operations and cybersurveillance in conflict.

The Governance Section provides a succinct assessment of the potential legal and operational responses by states to counter cyber proxies' offensive activity, including the normative tools and practices that could provide a basis for strengthening security and accountability in cyberspace. The ongoing war of aggression in Ukraine and its cyber component has prompted momentous legal and policy discussions. The ongoing conflict is the first instance of integrating cyberwarfare into an armed conflict and marks a turning point in collective defense collaborations. This final section closes with recommendations for increased normative cooperation grounded in international law, as well as collaboration in non-proliferation and cybercrime prevention.

[9] See Idem, p. 12, para. 38.

SUMMARY OF RECOMMENDATIONS

INTERNATIONAL LAW AND SECURITY IN CYBERSPACE

- To provide accountability in cyberspace, states and non-state actors should aim to attribute offensive operations by cyber proxies. To the extent possible, governments should coordinate and collaborate on the evidentiary process necessary for attribution, investigation, and prosecution and engage in proportionate collective response.
- States should investigate, prosecute, and impose sanctions for alleged violations of international humanitarian law and human rights abuses by cyber proxies, including private sector offensive actors, and provide effective remedies to victims.
- To clarify how international legal frameworks apply to cyberspace, states should define how they understand their obligations under international law. In the case of offensive operations by cyber proxies, it may be relevant for states to clarify their position on rules and principles of customary international law (including sovereignty and due diligence), as well as Article 8 of the International Law Commission's *Draft Articles on Responsibility of States for Internationally Wrongful Acts*. Article 8 establishes the principle of attribution of conduct to a State.
- To better anticipate evolving threats by cyber proxies and more proactively achieve accountability, states should engage in multilateral dialogues that can help them identify and address the strategies, behaviors, and modus operandi of cyber proxies. Such dialogues could be instrumental to better understand the different ways that offensive operations by cyber proxies may violate IHL and IHRL and elaborate new norms, remedy, and reparation mechanisms.

SYNERGIES WITH CYBERCRIME PREVENTION

- States should build collective resilience to cybercrime operations enabled or sanctioned by other states, counter illicit finance that underpins the cybercrime ecosystem, work with the private sector to disrupt and defend against cybercrime and ransomware as a service, and pursue the actors responsible to the full extent permitted under each partner country's applicable laws and relevant authorities.
- States should address the industrialization of cybercrime across appropriate multilateral formats to establish broader-based practices, actions, and norms and cooperate internationally across all elements of the cybercrime threat ecosystem.

SYNERGIES WITH DISARMAMENT AND NON-PROLIFERATION

- States should implement a transparent and operational framework to determine what elements of offensive cyberactivity constitute inherently governmental functions and what elements constitute closely associated functions that can be performed by private sector actors. For instance, when a State aims to prevent a potential cyberattack and neutralize an adversary's system, can this function strictly be operated by military personnel or can it be executed by a cybersecurity contractor?
- States should refrain from recruiting, using, financing, and trading with private sector cyber offensive actors and effectively regulate these offensive actors as well as private military and security companies.
- Regulating the underground cyberarms industry and its exchanges with hostile states and

non-state actors is a massive and complex challenge for which traditional non-proliferation approaches are inadequate. Any regulatory regime for offensive cyber capabilities needs to move beyond export control, severely target and reign in private sector offensive actors, and oversee innovation in civilian sectors.

- The sensitive nature of cyber defense and cyber offense services—increasingly converging with AI technologies—requires a rigorous set of actions across the military and civilian sectors for companies to meet their responsibilities and be in full compliance with international human rights law, international humanitarian law, and international criminal law.

A UN PERMANENT ACCOUNTABILITY MECHANISM FOR CYBERSPACE

A growing number of experts have argued in favor of the development of a standing accountability body to support responsible state behavior in cyberspace, a permanent UN mechanism to deal with cyberspace as a domain of conflict.[10] In the *New Agenda for Peace* published in July 2022, the UN Secretary-General calls for establishing an independent multilateral accountability mechanism for the malicious use of cyberspace by States to reduce incentives for such conduct.[11] In the past, UN working groups have provided strategic forums for policy and normative discussions but have been limited in their capacity to ensure accountability in cyberspace.[12]

As James Lewis explains, legal and political attribution of offensive cyberoperations should remain a sovereign responsibility and a state should ultimately remain in charge of the evidentiary process and the political analysis (trade-off) that comes with attribution.[13] Yet, the multilateral dimension can offer strategic support in "developing common evidentiary standards and information-sharing mechanisms for coordination of collective attribution."[14] As Lewis elaborates, "Coordinated attribution of malicious activity will require better information sharing between partners, and perhaps new mechanisms for sharing and harmonization, but will greatly strengthen the political effect of any accusation."[15] To increase accountability, states will also need to collaborate on "a broadly accepted menu of possible consequences and an ability to ensure that any consequences imposed are both proportional to the initial incident and consistent with international law and practice."[16]

[10] See Lewis, J., *Creating Accountability for Global Cyber Norms*, Center for Strategic and International Studies (CSIS), 23 February 2022, https://www.csis.org/analysis/creating-accountability-global-cyber-norms. See also CyberPeace Institute and Moriani, L., "Untangling Accountability in Cyberspace," CyberPeace Institute, 21 July 2022, https://cyberpeaceinstitute.org/news/untangling-accountability-in-cyberspace/. See also *Microsoft Digital Defense Report 2022*, p. 53.

[11] See *Our Common Agenda: Policy Brief 9—A New Agenda for Peace*, United Nations, July 2023, p. 27, https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-new-agenda-for-peace-en.pdf.

[12] The UN Open-ended Working Group on developments in the field of information and telecommunications in the context of international security (OEWG) is an initiative established by the United Nations to address issues related to security and information technologies. The group aimed to promote discussions among Member States, industry stakeholders, and civil society on the responsible use of cyberspace, norms, rules, and principles governing state behavior in cyberspace, and the protection of critical infrastructure. The OEWG provides a platform for multilateral dialogue and cooperation to enhance cybersecurity but remains limited in preventing potential threats in the digital domain and helping ensure accountability. For an analysis, see Pauwels, E., *Multilateralism and the Rising Challenges of Global Insecurity*, Konrad-Adenauer-Stiftung, 2022, https://www.kas.de/en/web/newyork/single-title/-/content/multilateralism-and-the-rising-challenges-of-global-cyber-insecurity. The UN has also been engaged in discussions and initiatives related to the use of mercenaries and private military and security companies, particularly under the Working Group on the use of mercenaries as a part of its mandate to address issues concerning private military and security companies, including those operating in cyberspace. The focus of such discussions typically revolves around human rights implications, accountability, and regulation of these private actors in various contexts, including cyber operations.

[13] Lewis, *Creating Accountability for Global Cyber Norms*, p. 4-7.

[14] Idem, p. 9.

[15] Ibid.

[16] Ibid.

This is a crucial collaborative endeavor, which could help address some of the most worrisome trends illustrated in this report, in particular the merger and exploitation of the cyberarms and cybercrime industries by nation states. Such a multilateral accountability mechanism would be instrumental:

- To build capacity and strengthen methods and practices to monitor offensive activity by cyber proxies, and support the evidentiary process required for coordinated attribution, investigation, and prosecution efforts, both in situations of peace and conflict; this could lead to a coordinated capacity for technical, legal, and political attribution that could benefit states with less expertise and capabilities.
- To collaborate on a range of internationally lawful responses with important implications to hold states and non-state actors accountable for hostile behaviors in cyberspace.
- To better analyze and anticipate current and evolving forms of offensive cyberoperations and the modus operandi, strategies, and behaviors of cyber proxies; such anticipatory analysis and foresight capacity could support prevention and mitigation of civilian harm and would progressively constitute an "institutional memory" of evolving threats in cyberconflict.
- To develop understanding of the evolving forms of dual-use technologies in cyberspace and related technological and knowledge transfer between actors; such interest in adaptive governance and responsible innovation would help modernize disarmament and non-proliferation efforts.
- To support ongoing normative efforts that aim to clarify how international law applies to cyberspace, in particular discussions to reaffirm states' obligations and responsibilities (including in their relationships with proxies), and to clarify the under-conceptualized, under-regulated zone that non-state actors occupy in cyberspace (for example, the legal definition of cyber proxies or cyber mercenaries).
- To support capacity-building efforts that involve countries most impacted by the digital and cybersecurity divides.

EXISTING UN FORUMS FOR POLICY DEBATE

The UN Open-ended Working Group on developments in the field of information and telecommunications in the context of international security (OEWG) is an initiative established by the United Nations to address issues related to security and information technologies. The group aimed to promote discussions among Member States, industry stakeholders, and civil society on the responsible use of cyberspace, norms, rules, and principles governing state behavior in cyberspace, and the protection of critical infrastructure. The OEWG provides a platform for multilateral dialogue and cooperation to enhance cybersecurity but remains limited in preventing potential threats in the digital domain and helping ensure accountability.

The UN has been engaged in discussions and initiatives related to the use of mercenaries and private military and security companies—particularly under the **UN Working Group on the use of mercenaries**—as a part of its mandate to address issues concerning private military and security companies, including those operating in cyberspace. The focus of such discussions typically revolves around human rights implications, accountability, and regulation of these private actors in various contexts, including cyberoperations.

FRAMING THE CYBER PROXY PHENOMENON

Over the last several years, we have witnessed **an accelerated blurring of the traditional boundaries that have kept the world in relative order**. Distinctions between offensive and defensive technologies, between public and private sectors, and between criminal endeavors, state power, and military agency have become less clear-cut. Cyber proxy and cyber mercenary activities have long challenged expert analysis and rules of governance, further blurring those already fading lines.

How has nation state cyber proxy activity been conceptualized in the past decade? And does this framework resist the evolving nature of cyber-conflict?

The concept of "**cyber proxies**" is central to understanding the complex set of alliances and arrangements that exist and keep evolving between states and non-state actors in cyberspace. Yet, salient academic efforts to capture cyber proxy activity have revealed how this phenomenon tends to escape fixed definition and categories. In 2021, the UN Working Group on the use of mercenaries clearly synthesized the difficulty with framing cyber proxy activity, stating "It is also difficult to ascertain the exact extent and nature of the provision of those services, given the highly sensitive nature of such operations and the secrecy and opaqueness that characterize the cyberindustry. More research is needed to identify which actors are delivering what kind of services. Current research on how States and non-State actors contract for cyber capabilities and what kind of services they are purchasing is both imperfect and incomplete."[17]

Starting a decade ago, several experts have eloquently deconstructed existing misconceptions about the nature of cyber proxy activity, refusing to frame it primarily as a top-down, state-centric activity. In 2011, with deep practical knowledge of the cyber domain, Jason Healey approached states' involvement with proxies not as rigid, codified forms of deputization but as a spectrum of engagement, from coordinated to tacit.[18] In 2015, Nicolò Bussolati brought to light early analysis of the challenges to international law posed by non-state actors' engagement in cyberwarfare.[19] And in 2016, Kubo Mačák provided an in-depth legal analysis of the ambiguities in attribution of cyber proxy activity by decoding the standards of attribution built into the body of the law of state responsibility.[20] In his 2018 monograph *Cyber Mercenaries: The State, Hackers, and Power*, Tim

> "We have witnessed an accelerated blurring of the traditional boundaries between offensive and defensive technologies, between public and private sectors, and between criminal endeavors, state power, and military agency.

[17] See A/76/151, p. 11, para. 34.

[18] Healey, J., "The Spectrum of National Responsibility for Cyberattacks," *The Brown Journal of World Affairs*, Vol. 18, No. 1 (Fall/Winter 2011): p. 57–70, https://www.jstor.org/stable/24590776.

[19] Bussolati, N., "The Rise of Non-State Actors in Cyberwarfare," in Ohlin, J.D., Govern, K., and Finkelstein, C. (eds.), *Cyber War: Law and Ethics for Virtual Conflicts* (Oxford: Oxford University Press, 2015), p. 102-126, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2764185.

[20] Mačák, K., "Decoding Article 8 of the International Law Commission's Articles on State Responsibility: Attribution of Cyber Operations by Non-State Actors," *Journal of Conflict and Security Law*, Vol. 21, Issue 3 (Winter 2016): p. 405-428, https://academic.oup.com/jcsl/article/21/3/405/2525375.

Maurer evokes early debates about cyberconflict in 2013 and explains how "the debate was state-centric, while the media was full of reports about the significant role non-state actors play in this field, including private companies such as Gamma International and Vupen, hacktivist groups from Anonymous to the Syrian Electronic Army, and cybercriminals operating with impunity from different hotspots around the world."[21]

Since then, a new strain of research and analysis has confirmed that proxy relationships between states and an array of non-state actors have grown increasingly complex. In 2018 and 2019 respectively, David Sanger and Ben Buchanan wrote gripping accounts that illustrate the importance of plausible deniability and the strategic logic for states to "function," almost as a symbiosis with non-state actors in modern geopolitical conflicts. In 2019, Andreas Krieg and Jean-Marc Rickli focused on the game-changing element that emerging technologies, AI, and automation represented when outsourcing cyberwarfare to proxies.[23] In 2021, Nicole Perlroth published the synthesis of 10 years of research outlining the role of hackers and private sector offensive actors in the clandestine cyberarms industry and the global cyberweapons arms race it has sparked.[24] In 2021, Michael McGuire provided a criminological perspective in a striking report based on field interviews showing the integration of cybercrime—the "Web of Profit"—in today's statecraft.[25] In 2022, Justin Sherman published an enlightening policy brief with a focus on the Russian web of non-state actors—from front companies to state-tapped individuals, hacktivists to cybercriminals—and their complex relationships with the state. Sherman emphasizes that "proxy as a universal term fails to capture the gradations of the State's involvement with hackers, assuming a top-down hierarchical relationship that is not always present in Russia."[26]

The present report converges with these previous and recent research efforts by unveiling the polymorphous and adaptive nature of cyber proxy relationships. Still, in an attempt to characterize cyber proxy activity, the work of Maurer remains particularly useful and relevant for two reasons.[27] **First**, Maurer's conceptualization is open and comprehensive enough to reflect the evolving nature of the phenomenon. Cyber proxies can be considered as "intermediaries that conduct or directly contribute to an offensive cyber action that is enabled knowingly, whether actively or passively, by a beneficiary."[28] The notion therefore encompasses the different contemporary ways that states can either mandate and assist in the conduct of cyberattacks, support and fund hostile cyber activities, or merely allow such actions to be waged. To frame nation states' relationships with cyber proxies, Maurer establishes a **typology** that includes contractual delegation of authority, orchestration, and sanctioning. ***Delegation*** represents a type of legal deputization where "clear responsibilities are assigned to proxies through the channels of municipal law and policy, for example to undertake pre-emptive strikes against perceived cyberthreats on critical infrastructure."[29] Municipal law often provides the legal framework within a state's jurisdiction that allows its government to authorize and conduct cyber operations, including pre-emptive strikes, in accordance with its domestic

21 See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. X.

22 See Sanger, D., *The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age* (New York, NY: Random House/Crown, 2018). See Buchanan, B., *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics* (Cambridge, Massachusetts: Harvard University Press, 2020).

23 Krieg, A. and Rickli, J., *Surrogate Warfare: The Transformation of War in the Twenty-First Century* (Washington, D.C.: Georgetown University Press, 2019), https://www.jstor.org/stable/j.ctvf34hnd.

24 Perlroth, N., *This Is How They Tell Me the World Ends – The Cyberweapons Arms Race* (New York, NY: Bloomsbury Publishing, 2021).

25 See McGuire, *Nation States, Cyberconflict, and the Web of Profit.*

26 Sherman, J., *Untangling the Russian Web: Spies, Proxies, and Spectrums of Russian Cyber Behavior*, Digital Forensics Research (DFR) Lab, Cyber Statecraft Initiative, The Atlantic Council, Issue Brief, September 2022, p. 2, https://www.atlanticcouncil.org/wp-content/uploads/2022/09/Untangling-the-Russian-Web-Spies-Proxies-and-Spectrums-of-Russian-Cyber-Behavior-1.pdf.

27 See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*.

28 Idem, p. XI.

29 See A/76/151, p. 10, para. 31.

laws and regulations. ***Orchestration*** does not involve contractual or legal delegation, nor oversight, but consists in providing cyber proxies with enabling support that can vary in focus and intensity, such as political, financial, or logistical means. ***Sanctioning*** implies that a nation state does not multilaterally acknowledge, but rather overlooks and tolerates, cyber proxy activity that emanates from its territory and may serve its strategic interest.

Second, Maurer avoids the conceptual trap of categorizing cyber proxies by intent, as motivations and goals are rarely static and clear-cut for those actors. The next section of this report will illustrate the increasing permeability between cyber proxies' intents, business models, strategies, and behaviors. Some hackers' groups employ their offensive services in an informal, opaque relationship with a nation state while also pursuing financial revenue strategies through cybercrime. Maurer therefore focuses on the **strategic logic for nation states** to employ proxies in cyberspace: to conduct tactical and persistent threat operations beyond kinetic reach and conventional military capabilities, to dissuade and weaken adversaries, and ultimately to "project power" in cyberspace.

Experts in cybersecurity and international relations have long postulated several hypotheses to explain nation states' increasing reliance on cyber proxies. In terms of **political and economic cost management**, proxies in the cyberarms industry allow states not only to avoid direct military action but also to efficiently acquire and employ capabilities that might be prohibitively expensive to develop in-house within military units. Using cyber proxies may also allow nation states to benefit from **plausible deniability**. Cyberweapons are stealthy and effective and can be launched below the threshold of war with a high degree of automation, anonymity and opacity. Their use can be outsourced in multi-stage operations to proxy actors located at arms' length of a nation state and in different jurisdictions. Consequently, nation states achieve an "opportunistic dissociation" from the means of cyber aggression—which is much more difficult to do in kinetic situations—and therefore potentially escape attribution of responsibility and liability. The UN Working Group on the use of mercenaries offers another compelling argument that speaks to the state-centric nature of multilateral governance and international law: "unlike States that are subject to international human rights and humanitarian law protocols, they [cyber proxies] tend to operate outside the purview of such protocols, making attribution, arrests, and prosecution difficult."[30]

> "Using cyberproxies allows nations states to benefit from plausible deniability. Cyberweapons can be launched below the threshold of war with a high degree of automation, anonymity, and opacity.

WHAT ARE THE TYPES OF OFFENSIVE AND DEFENSIVE CAPABILITIES OUTSOURCED IN CYBERSPACE?

Cyberservices include both the provision of expertise and related support and the provision of cyberproducts (e.g., spyware and software) that can be harnessed by states. Defensive services—from antivirus software, patches, and firewalls to more sophisticated algorithmic programs for cyberthreat detection—are provided by cybersecurity firms to public and private sector actors. Active cyber defense has increasingly become part of what states call "persistent engagement," which refers to techniques at the convergence of AI, cyber intelligence, cyber protection, and cyber analytics to proactively and predictively combat cyberattacks and protect data assets. These specific services, whether active or passive, fall within existing legal boundaries for cybersecurity operations.

Yet, experts like Perlroth also describe the growth of "a wildly lucrative, entirely unregulated gray market for insanely dangerous digital weapons that private hackers develop and then sell to the highest bidder."[31] Private sector firms, groups of hackers, and other rogue operators all compete to provide offen-

[30] See Idem, p. 11, para. 37.

[31] See Tepperman, J., "The Most Serious Risk Facing the United States," *The New York Times*, 9 February 2021, https://www.nytimes.com/2021/02/09/books/review/this-is-how-they-tell-me-the-world-ends-nicole-perlroth.html.

sive cyberservices. Malicious and offensive services carried out by state-sponsored actors or proxies working for states include the targeting of digital assets and digital assets providers, surveillance, and industrial espionage, as well as cyberattacks on critical infrastructures, elections, and information operations. The UN Working Group on the use of mercenaries emphasizes that "Both democratic and non-democratic States acquire offensive technologies from external providers as do States with in-house cyber capabilities as well as those without such resources."[32] The multiple offensive cyberoperations that have been observed in recent years include *inter alia* the below categories and have increasingly targeted civilian populations and critical entry-points in digital and physical infrastructures.

> Private sector firms, groups of hackers, and other rogue operators all compete to provide offensive cyberservices.

- **Ransomware** involves malicious actors breaking into computer networks and locking digital information until the victim pays for its release. Recent high-profile attacks have cast a spotlight on this rapidly expanding criminal industry. In 2021, FIN7 (also known or related to DarkSide), a sophisticated advanced persistent threat (APT) group connected with the Russian government, conducted a double extortion attack on the Colonial Pipeline Company. The oil company was threatened with the leakage of stolen data in addition to the data on its systems remaining encrypted unless a ransom was paid. Colonial Pipeline paid a ransom of 75 bitcoins (then equivalent to 4.4 million USD).[33]
- **Supply chain attacks** use malicious implants or other vulnerabilities inserted prior to a system's installation with the goal to infiltrate and corrupt datasets, hardware, software, operating systems, or services at any point during the life cycle. In 2021, the supply chain of the software producer, Kaseya Ltd, was infiltrated by ransomware, which then corrupted the computer systems of its clients through software update. The attack was perpetrated by the REvil (i.e., Ransomware Evil) group, a Russian- speaking and Russia-based ransomware gang.[34] It disrupted the operations of around 1,500 companies and affected thousands of victims, including nurseries, schools, pharmacies, and supermarkets in 17 countries.
- **Zero-day vulnerability attacks** exploit a previously unknown hardware, firmware, or software vulnerability. In 2021, four zero-day exploits of vulnerabilities in Microsoft Exchange were utilized to gain initial access to servers and install backdoor for data-exfiltration and ransom-seeking on more than 5,000 unique servers in more than 115 countries. More than 10 APT groups were involved in the attacks, including the prominent HAFNIUM and BARIUM (also known as APT41 or the Winnti Group), both allegedly sponsored by the Chinese government.[35]
- **A distributed denial of service (DDoS)** technique consists in taking down a digital service by flooding it with so much traffic data that the operating system is unable to maintain functionality. In February 2022, as Russian forces gathered along the Ukrainian border, DDoS attacks allegedly by Russian threat actors targeted Ukraine's armed forces, defense ministry, public radio, and the two largest national banks.[36]

[32] See A/76/151, p. 7, para. 19.

[33] Campbell, I. C., "Colonial Pipeline CEO confirms company paid $4.4 million ransom it wasn't supposed to pay," *The Verge*, 19 May 2021, https://www.theverge.com/2021/5/19/22443933/colonial-pipeline-ransom-4-million-hack-gas-shortage.

[34] Osborn, C., "Updated Kaseya ransomware attack FAQ: What we know now," *ZDNet*, 23 July 2021, https://www.zdnet.com/article/updated-kaseya-ransomware-attack-faq-what-we-know-now/.

[35] "Number of APT groups exploiting the latest Exchange vulnerabilities grows, with thousands of email servers under siege, ESET discovers," *ESET*, 10 March 2021, https://www.eset.com/uk/about/newsroom/press-releases/number-of-apt-groups-exploiting-the-latest-exchange-vulnerabilities-grows/.

[36] Antoniuk, D., "DDoS attacks hit Ukrainian government websites," *The Record*, 14 February 2022, https://therecord.media/ddos-attacks-hit-websites-of-ukraines-state-banks-defense-ministry-and-armed-forces/.

- **Malware attacks** consist in injecting malicious code into computer systems for destructive purposes, such as deleting datasets, running intrusive programs, or corrupting operating systems. In 2017, the NotPetya malware was spread by a general update to a tax accounting software used by many Ukrainian businesses and impacted companies around the world. While masquerading as ransomware, NotPetya irreversibly encrypted every infected machine's operating system, thus effectively destroying those computers. The attack was attributed by the U.S. and UK governments to the APT group Sandworm (also called IRIDIUM and Unit 74455), which is allegedly a Russian cybermilitary unit of the GRU (the Main Directorate of the General Staff of the Armed Forces of the Russian Federation and the organization in charge of Russian military intelligence).[37] The estimated global economic losses caused by the NotPetya attack exceeded 10 billion USD.

WHO ARE THE THREAT ACTORS USED IN CYBER PROXY ACTIVITY?

The following types of non-state actors may be harnessed as proxies:

- **APT groups** are stealthy threat actors that conduct long-term, resource-intensive operations to collect strategic intelligence and penetrate the cyber defenses of potential targets in the public and private sectors. APT groups are often closely associated with a state. They have the in-house capacity to deploy sophisticated AI and cyber offense strategies. In April 2022, Microsoft's digital security unit reported that several APT groups representing Russian government security services had been active through 2021 to compromise "organisations that could provide valuable intelligence on a Ukrainian military, diplomatic, or humanitarian response to Russian military action."[38] State-sponsored APT groups have been involved in the advanced attacks that targeted the U.S. company SolarWinds in 2020 and Microsoft Exchange servers in 2021.[39] In both cases, the threat actors conducted indiscriminate cyberoperations that resulted in harm to thousands of civilian institutions, including schools, medical facilities, and critical infrastructure platforms.

- **Cybermilitias** function often on a voluntary basis and coalesce around varied forms of organization, from a simple forum to a more cohesive non-state entity. Members of cybermilitias may not exhibit the same level of sophisticated skills as APT groups. They may not necessarily benefit from a state's sustained financial support and may not be engaged in furthering long-term geostrategic interests. Still, cybermilitias may be used in cyber espionage and surveillance, information operations, and, in certain circumstances, offensive cyberoperations against critical digital assets and infrastructures. It is important to note that university students and youth movements are sometimes involved in cybermilitias. Starting with "information warfare" units in the late 1990s, China has built a strong structure of cybermilitias that recruit hackers, both in the private sector and within highly-skilled university programs.[40] In 2011, reports revealed that members of the pro-Kremlin youth movement, Nashi, had been involved in cyber and information operations.[41] As early as 2011, the Syrian Electronic Army gathered hackers, including youth, to conduct surveillance and cyber-

[37] Marsh, S., "US joins UK in blaming Russia for NotPetya cyber-attack," *The Guardian*, 15 February 2018, https://www.theguardian.com/technology/2018/feb/15/uk-blames-russia-notpetya-cyber-attack-ukraine. See also "Foreign Office Minister condemns Russia for NotPetya attacks," Government of the United Kingdom, 15 February 2018, https://www.gov.uk/government/news/foreign-office-minister-condemns-russia-for-notpetya-attacks.

[38] Microsoft Digital Security Unit, *Special Report: Ukraine: An overview of Russia's cyberattack activity in Ukraine*, April 2022, p. 6, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4Vwwd.

[39] See McGuire, *Nation States, Cyberconflict, and the Web of Profit*, p. 13. See "Russia: UK exposes Russian involvement in SolarWinds cyber compromise," Government of the United Kingdom, 15 April 2021, https://www.gov.uk/government/news/russia-uk-exposes-russian-involvement-in-solarwinds-cyber-compromise. See also "Analyzing attacks taking advantage of the Exchange Server vulnerabilities," Microsoft Defender Threat Intelligence, 25 March 2021, https://www.microsoft.com/en-us/security/blog/2021/03/25/analyzing-attacks-taking-advantage-of-the-exchange-server-vulnerabilities/.

[40] See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 114-116.

[41] See Idem, p. 97.

attacks against media organizations and opponents seen as critical of the Assad regime.[42]

Another example of formal cybermilitias is the Estonian Defense League's cyber unit, born in the wake of the 2007 cyberattacks that targeted the country. The Estonian government helped recruit civilian cybersecurity experts to form this cyber unit, which is supposed to function as a regular army in situations of cyberconflict.[43] There is a parallel with Kyiv's "IT Army," a large group of approximately 400,000 civilians who, in the ongoing conflict with the Russian Federation, volunteered their services to support Ukraine's government in offensive cyberoperations. Ukraine's "IT Army" is reported to operate at the direction of the government, with specific tasks and a target list assigned through a Telegram channel, including requests to conduct cyberattacks against Russian and Belarusian targets.[44]

- Groups of **hacktivists** may include civilians who voluntarily engage in hacking for ideological, political, religious, or patriotic reasons. Such groups usually act independently from the strategic agenda of a state. For instance, the "Cyber Partisans" is an independent Belarusian hacktivist group that "has claimed responsibility for several major cyberattacks, including a high-profile operation against the Belarusian railway system that reportedly halted Russian ground artillery and troop movement into Ukraine."[45] Another example is the global hacktivist group Anonymous that came to life in response to the Islamic State's (IS) violent propaganda on social media.[46] Progressively, Anonymous organized a form of digital resistance, combining cyber espionage, hacking techniques, and automated bots to compromise IS websites, networks, and social media presence.
- **Cybercriminals**, increasingly organized in **cybercrime syndicates**, are criminal groups that use the holding and theft of corporate, institutional, or population data, and digital assets as an extortion mechanism. They are profit-driven individuals or groups that either operate for their own benefits or increasingly lend their extortion mechanisms as a service to benefit other state-sponsored rogue actors. *The Microsoft Digital Defense Report 2022* shows how cybercriminals continue to function as extremely sophisticated profit enterprises. For instance, the report insists on the evolution from what was portrayed as "ransomware gangs" to a full-spectrum ransomware economy where "separate entities build malware, gain access to victims, deploy ransomware, and handle extortion negotiations."[47] The Microsoft report explains how key affiliates from Conti, one of the most active ransomware groups in the past years, dissolved operations in mid-2022 to "re-emerge months later and redistribute their technical capabilities and resources to new groups."[48]
- **Private military and security companies** constitute a growing category of actors that can deploy offensive cyber capabilities. They range from small companies with cutting-edge expertise in cyber and AI technologies to larger corporate groups in the military and defense sector. Some of the largest defense contractors active across the full spectrum of cyberoperations are Raytheon, Lockheed Martin, Booz Allen Hamilton, Science Applications International Corporation (SAIC), and CACI International. Other private sector offensive actors include the NSO Group (based in Israel), Hacking Team (based in Italy), and the DarkMatter Group (based in the United Arab Emirates). The 2021 UN Working Group on the use of mercenaries emphasizes that "The evolving threat of the

[42] See Idem, p. 88.

[43] See Idem, p. 20, 103.

[44] See Lonergan, E., "Cyber Proxies in the Ukraine Conflict: Implications for International Norms," Council on Foreign Relations, 21 March 2022, https://www.cfr.org/blog/cyber-proxies-ukraine-conflict-implications-international-norms.

[45] See Smeets, M. and Achberger, B., "Cyber hacktivists are busy undermining Putin's invasion," *The Washington Post*, 13 May 2022, https://www.washingtonpost.com/politics/2022/05/13/cyber-attack-hack-russia-putin-ukraine-belarus/.

[46] See Pauwels, E., *The New Geopolitics of Converging Risks: The UN and Prevention in the Era of AI*, United Nations University Centre for Policy Research, 2019, p. 16, https://collections.unu.edu/eserv/UNU:7308/PauwelsAIGeopolitics.pdf.

[47] See *Microsoft Digital Defense Report 2022*, p. 11.

[48] Ibid.

© iStock/gorodenkoff

privatization of cybersecurity attacks through a new generation of private companies referred to as so-called 'cybermercenaries' is proliferating, and there is an increasingly blurred line separating the private and national spheres."[49] Cybersecurity experts have voiced concerns underlining that "cyber mercenaries bring world-class capabilities to countries with low human rights protection, rule of law, and good governance."[50] According to the CyberPeace Institute, "private operations that are run on behalf of state actors in a self-regulated market provide a stress test to states' will and capacity to monitor and enforce its obligations to respect, protect and fulfil human rights; some deliberately choose to use mercenaries in an attempt to escape accountability."[51]

The term **"cyber mercenaries"** poses important definitional problems. The elements that frame the concept of mercenary are enumerated in Article 47 of Additional Protocol I to the Geneva Convention of 1949. Six specific criteria have to be met for an actor to be considered as a mercenary: special recruitment; direct participation in hostility; desire for private gain as primary motivation; neither a national of a party to conflict nor a resident of territory controlled by a party; not a member of the armed forces of a party to the conflict; and not sent by another state on official duty as a member of its armed forces. Yet, the application of these criteria to cyber offense does not reflect what is happening in cyberspace. The collusion and permeability between different types of cyber proxies has implications that might not fit the above criteria related to special recruitment, hierarchy, nationality, and residence. The criteria of private gain as primary motivation also poses problems. Some APT groups might demonstrate mixed behaviors of being involved in state-sponsored cyber offense, but also committing cybercrime. Patriotic hackers may conduct offensive cyberoperations for ideological reasons and not for private gain. There are also several levels

[49] See Note 9, A/76/151, p. 8, para. 23.

[50] See Microsoft, "Response to the United Nations (UN) Working Group on the Use of Mercenaries," p. 1.

[51] See CyberPeace Institute, "Mercenary-Related Activities in Cyberspace."

of dissociation and a certain amount of opacity in the supply chain for offensive and intrusive cyber-operations, which means that not every "supplier" will have a good sense of what the code is intended for and whom it is aiming to target. In its 2021 report, the UN Working Group on the use of mercenaries concurred with the above arguments, recognizing that the scope of the definition of mercenary and mercenary-related activities did not reflect the activities conducted by non-state actors in cyberspace.[52] The Working Group decided to focus on "the range of military and security services provided in cyberspace which can generate mercenary-related activities in order to stimulate a discussion on how better to frame and address them."[53]

Microsoft has proposed framing the concept of mercenarism in cyberspace as the adversarial activities of private military and security companies (PMSCs) that manufacture and sell cyberweapons or what Microsoft researchers call "private sector offensive actors" (PSOAs).[54] Such a conceptualization also reflects the growing importance of the cyberarms industry.

GLOSSARY

- **Cybercrime** is described by the United Nations Office of Drugs and Crime (UNODC) as an "act that violates the law, which is perpetrated using information and communication technology (ICT) to either target networks, systems, data, websites and/or technology or facilitate a crime."
- **Cyber espionage** is framed by UNODC and security experts as unauthorized methods of intelligence collection. It can be perpetrated by "government actors, state-sponsored or state-directed groups, or others acting on behalf of a government" for economic gain, competitive or military advantage, or political reasons.
- **Cyberwarfare** is framed by the ICRC as "cyber operations during armed conflicts" or "methods of warfare that are used against a computer, a computer system or network, or another connected device, through a data stream."[55] The 2021 UN Working Group on the use of mercenaries defines the concept as "a method of warfare that can not only infiltrate, disrupt, damage or even destroy military or civilian objects, but also cause serious human harm."[56]
- **Offensive cyberoperations** are defined by the U.S. National Institute of Standards and Technology as "cyberspace operations intended to project power by the application of force in or through cyberspace." For the purpose of this report, we define "hostile" or "offensive cyber operations" as operations "that alter, disrupt, deceive, degrade or destroy computer systems or networks, or otherwise undermine the confidentiality, integrity and availability of computer systems or networks for individuals and communities."[57]
- **Offensive cyber proxy actors** can be defined as cyberthreat actors that conduct or contribute to offensive cyberoperations which align with a group within a state, support a state's national interests, or are tacitly permitted by a state or a non-state actor (such as violent and armed non-state actors).

[52] A/76/151, p. 5, para. 7.

[53] Idem, p. 6, para. 14.

[54] See Microsoft, "Response to the United Nations (UN) Working Group on the Use of Mercenaries," p. 2.

[55] See ICRC, International Humanitarian Law and Cyber Operations during Armed Conflicts (ICRC position paper submitted to the Open-Ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security and the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security, November 2019), *International Review of the Red Cross* (2020), 102 (913), https://international-review.icrc.org/sites/default/files/reviews-pdf/2021-03/ihl-and-cyber-operations-during-armed-conflicts-913.pdf.

[56] A/76/151, p. 5.

[57] Ibid.

- **Hybrid warfare** involves the combination of conventional and unconventional strategies, methods, and tactics in contemporary warfare, as well as the psychological or information-related aspects of modern conflicts.[58]
- In its 2021 report "Harmful Information," the ICRC defines **information operations** as "the strategic and calculated use of information and information-sharing systems to influence, disrupt or divide society."[59] In armed conflict and other situations of violence, information operations can involve surveillance and profiling of specific populations' subgroups, disinformation campaigns, and the spread of hate speech targeted at ethnic and religious communities. **Misinformation** consists in false information that is disseminated, without the intent to cause harm, often by individuals who have not done fact-checking. **Disinformation** refers to false information that is conceived and spread with a deliberate intent to cause harm or deceive. **Hate speech** includes "all forms of expression (text, image, audio) that spread, incite, promote or justify hatred and violence based on intolerance, usually against identity traits (gender, religion, ethnicity, sexual orientation, etc.)."[60]

[58] See Bilal, A., "Hybrid Warfare – New Threats, Complexity, and 'Trust' as the Antidote," *NATO Review*, 30 November 2021, https://www.nato.int/docu/review/articles/2021/11/30/hybrid-warfare-new-threats-complexity-and-trust-as-the-antidote/index.html.

[59] ICRC, *Harmful Information – Misinformation, disinformation and hate speech in armed conflict and other situations of violence: ICRC initial findings and perspectives on adapting protection approaches*, July 2021, p. 18, https://www.icrc.org/en/publication/4556-harmful-information-misinformation-disinformation-and-hate-speech-armed-conflict.

[60] Ibid.

TECHNICAL SECTION: THE UNDERGROUND CYBERARMS AND CYBERCRIME INDUSTRIES

Regulating cyber proxies and cyber mercenaries requires first an understanding of the ramifications and inner workings of the underground cyberarms and cybercrime industries. Within these complex, opaque, and often fluid networks, it becomes increasingly challenging to map cyberthreat actors, their interests, practices, and evolving relationships with nation states. The strategies and behaviors of states and non-state actors, including intense knowledge and technical transfer, contribute to enhancing stealth, opacity, and adaptivity within these clandestine industries. Increased permeability between cyberthreat actors and the blurring of their modus operandi severely complicates legal attribution and regulation across borders and jurisdictions.

BLURRED LINES

While cybersecurity technologies are inherently dual-use, the lines that previously separated cyber defense and cyber offense have become increasingly blurred. Core AI and cybersecurity capabilities developed by the private sector are drastically augmenting functionality across a wide spectrum of civilian and commercial applications that could be abused by cyber proxies (e.g., the growing Internet of Things). Cyber proxy actors have also been able to acquire available defense systems and repurpose them for offensive use. APT groups affiliated with China, for instance, have reverse-engineered antivirus software sold by Western companies and have learned directly from "red-teaming" techniques that white hat hackers use to help companies defend their digital assets.[61]

- The rapid convergence of AI, automation, and cybersecurity innovations has not only blurred the boundary between offense and defense, but has also increasingly **merged civilian and military technologies**, creating new dependencies between the digital architectures that power private, public, and national security systems and infrastructures. For instance, critical civilian infrastructures and their cyber resilience systems may be privately owned and operated. Similarly, parts of strategic military cyber defense may depend on private security and military companies. In this configuration, the role and legal status of private sector contractors is increasingly ambiguous. Their services may oscillate between defense and offense, civilian and military spheres, which has implications for how they can be targeted through international law.
- As the trade for cyber defense and cyber offense has become global, the **notion of sovereign control over dual-use technologies is largely being challenged** with implications for national security. A large number of private sector offensive actors will sell expertise, malware, and services to the highest bidder without considering national security concerns. Others will simply not properly execute a due diligence assessment that would prevent offensive cyber expertise from being used in attacks against a country's infrastructure or population.

[61] Hiroaki, H. and Lee, T., " Hack the Real Box: APT41's New Subgroup Earth Longzhi," *TrendMicro*, 9 November 2022, https://www.trendmicro.com/en_us/research/22/k/hack-the-real-box-apt41-new-subgroup-earth-longzhi.html.

As this section will unveil, another trend which complicates attempts to govern and control the proliferation of offensive cyberservices is the growing technological transfer and skill-combination that exist between criminal capabilities, state power, and military agency.

> Private sector offensive actors sell expertise, malware, and services to the highest bidder and without consideration of their potential use against infrastructure or populations.

The implications of rapidly expanding, unregulated markets for cyber offense remain corrosive to international peace and security. Offering the means of cyberaggression to every actor who can afford it is already transforming how contemporary conflicts are fought and may rewrite the rules of the global order. In the last five years, several **geostrategic and technological trends** have accelerated the proliferation of cyberthreat actors, increased the sophistication of cyberweapons, and augmented the set of vulnerabilities that can be targeted across nations' digital and physical infrastructures.

TRENDS

How is great power competition shaping the private markets of offensive cyber capabilities?

- **What has made cyberwarfare both more pervasive and accessible to nation states is the commodification and privatization of offensive cyber capabilities through a growing cyberarms industry that operates mostly in shadow and secrecy.** Since the beginning of the 21st century and in their race to establish dominance in cyberspace, tech-leading nations have created and fueled a lucrative and vastly expanding grey market for cyberweapons. To further geostrategic interests, national defense and intelligence programs have harnessed the expertise and recruited the services of hackers ("brains for hire"), start-ups, and large private security and military companies. Progressively, as they stockpiled exploits and intrusion techniques, tech-leading nations lost monopoly over their use. A growing number of states have recognized the opportunity and begun engaging with this grey market—buying codes and data, contracting services and talent—to advance domestic, global, and regional interest. Their primary strategic logic remains to avoid attribution of wrongful conduct by using non-state actors for cyberattacks. In the space of the last decade, cybersecurity experts have confirmed a rapid, uncontrolled proliferation of non-state actors able to develop or acquire the expertise to craft, repurpose, use, and trade cyberweapons sometimes purchased with after-sales services.[62]
- Today, an expanding number of countries rely on the cyberservices of private contractors—from large PMSCs to smaller start-ups—which operate in a global market and, often, with close ties ("revolving doors") to intelligence and security professions. Researchers have helped shed light on the nature of the services and the types of companies that make up the underground cyberarms ecosystem.[63] For instance, Zerodium (formerly Vupen) based in France and Maryland in the U.S., provides zero-day exploits that can be used for offensive cyber missions and network operations.[64] ReVuln, operated from Malta by Italian hackers, "specializes in finding remote vulnerabilities in industrial control systems that can be used to access—or disrupt—water treatment facilities, oil and gas pipelines and power plants."[65] The Miami-based company, Immunity Inc., harnesses,

[62] See Perlroth, *This Is How They Tell Me the World Ends*. See also McGuire, *Nation States, Cyberconflict, and the Web of Profit*. See also Greenberg, A., *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers* (New York, NY: Doubleday/Penguin Random House, 2019). See also Maurer, *Cyber Mercenaries: The State, Hackers, and Power*.

[63] Idem.

[64] See Perlroth, *This Is How They Tell Me the World Ends*, p. 219.

[65] Perlroth and Sanger, "Nations Buying as Hackers Sell Computer Flaws," *The New York Times*, 13 July 2013, https://www.nytimes.com/2013/07/14/world/europe/nations-buying-as-hackers-sell-computer-flaws.html.

exploits, and develops techniques, training, and after-sale services that provide governments across the world with cyberattack capabilities ("penetration testing"). It has provided large security and defense contractors in the U.S. and abroad with training in zero-day exploitation techniques.[66] Acquired in 2010 by the large Computer Science Corps, Vulnerability Research Labs has operated in Maryland with a fine team of U.S. National Security Agency (NSA) recruits and a global network of subcontractors and specialized in hunting, weaponizing, and testing powerful zero-day vulnerabilities.[67] Gamma Group is a commercial spyware group based in the UK that sells surveillance services to law enforcement, intelligence, and military agencies.[68]

The company CyberPoint, later renamed DarkMatter, operates from the United Arab Emirates and recruited former NSA engineers to work on sophisticated spyware and cyber intrusion tools.[69] The firm became well-known for the scandal "Project Raven": cutting-edge cybertheft and spying tools developed by former American cyber espionage agents were used to keep thousands of civil society activists, journalists, and political figures under tight digital surveillance.[70] The Milan-based Hacking Team is another start-up that used to sell sophisticated spyware and malware to security agencies in the U.S. and European countries as well as to many other states, including dictatorships.[71] When Hacking Team was hacked in 2015, the intrusion not only showed that the company's spyware was harnessed for mass surveillance, but also led to Hacking Team's library

> "Cybersecurity experts have confirmed a rapid, uncontrolled proliferation of non-state actors able to develop or acquire the expertise to craft, repurpose, use, and trade cyberweapons sometimes purchased with after-sales services.

of zero-day vulnerabilities being leaked and exploited by threat actors across the world. The NSO Group, a competitor based in Israel, has also made a highly lucrative and global business of targeted surveillance—for instance, charging more than one million USD to install its remote zero-click[72] hacking spyware ("Pegasus") on just a dozen iPhones and Android phones.[73] In 2021, the Biden administration placed NSO and another Israeli firm, Candiru, on a U.S. Department of Commerce blacklist that banned American companies from doing business with the hacking firms.[74] In October 2022, the White House announced it would "stand against digital authoritarianism" and fight the "illegitimate use of technology, including commercial spyware and surveillance technology."[75] Yet, recent investigation shows that the cyber surveillance business is booming with clients in the U.S. government and many other governments, with new companies emerging and sophisticated technologies being acquired.[76] Such forms of stealthy, precision surveillance can have corrosive implications for civilian populations across the globe.

66 See Perlroth, *This Is How They Tell Me the World Ends*, p. 151-152 and 155-157.

67 See Idem, p. 139-145.

68 See Idem, p. 183. See also Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 151.

69 See Perlroth, *This Is How They Tell Me the World Ends*, p. 163-164.

70 See Pauwels, *The New Geopolitics of Converging Risks*, p. 16.

71 See Perlroth, *This Is How They Tell Me the World Ends*, p. 172-178.

72 "Zero-click" technology can stealthily and remotely extract everything from a target's mobile phone, without the user having to click on a malicious link to give Pegasus remote access.

73 See Perlroth, *This Is How They Tell Me the World Ends*, p. 179-189.

74 Sanger, Perlroth, Swanson, A., and Bergman, R., "U.S. Blacklists Israeli Firm NSO Group Over Spyware," *The New York Times*, 03 November 2021, https://www.nytimes.com/2021/11/03/business/nso-group-spyware-blacklist.html.

75 Sanger, "Biden's National Security Plan Focuses on China, Russia and Democracy at Home," *The New York Times*, 12 October 2022, https://www.nytimes.com/2022/10/12/us/politics/biden-china-russia-national-security.html.

76 Mazzetti, M., Bergman, and Stevis-Gridneff, M., "How the Global Spyware Industry Spiraled Out of Control," *The New York Times*, 08 December 2022, https://www.nytimes.com/2022/12/08/us/politics/spyware-nso-pegasus-paragon.html.

- **The clandestine cyberarms industry is increasingly merging with the illicit, interconnected cybercrime economies that are thriving across the world.** Nation states tend to cultivate a larger and more complex ecosystem of cyber proxies where APT groups, private cybersecurity firms, private sector offensive actors, and cybercriminal groups collide, compete, and even collaborate. This sinister and opportunistic combination is happening through new forms of collusion and dependency.

  First, prominent states increasingly leverage the offensive services of different types of private sector actors in the underground cyberarms industry. They have used IT firms as front companies to conduct covert adversarial cyberoperations and espionage.[77] They have also nurtured ties, knowledge, and technology transfers between private cybersecurity firms and private groups engaged in hostile mercenary activity.[78] Such practices are facilitated by a lack of clarity over the legal status of military and security services provided in cyberspace. As the UN Working Group on the use of mercenaries notes, "private actors can be engaged by States and non-state actors not only to protect their own networks and infrastructures but also to carry out cyberoperations designed to weaken the military capacities and capabilities of enemy armed forces or to undermine the integrity of another State's territory."[79]

  Second, more nation states are not only acquiring techniques and vulnerabilities usually harnessed in cybercrime, but are also recruiting the services of cybercriminals. It has become increasingly common for cybercriminal syndicates to act as proxies for states. As early as 2017, the U.S. Department of Justice indicted two Russian Federal Security Service (FSB) officers and the criminal hackers they hired for conducting computer hacking, economic espionage, and other criminal offenses.[80] In 2021, the U.S. Department of Treasury announced new sanctions against the Russian Intelligence Services for perpetrating malign cyber activities compromising U.S. interests and co-opting the cybercriminal services of the ransomware group Evil Corp., noting "To bolster its malicious cyber operations, the FSB cultivates and co-opts criminal hackers, including the previously designated Evil Corp, enabling them to engage in disruptive ransomware attacks and phishing campaigns."[81] Investigation by the U.S. cybersecurity provider FireEye shows that operatives of the Chinese group APT41 have been allowed to moonlight for personal gain.[82] APT41 has penetrated and spied on global tech, communications, and healthcare providers for the Chinese government, while using ransomware against game companies and attacking cryptocurrency providers for personal profit.[83] APT41's links to both cybercrime marketplaces and state-sponsored activity may be a signal that the group enjoys a tacit agreement, which allows it to conduct its own for-profit activities or Chinese authorities are willing to overlook them. Similar patterns exist in Russia where "authorities allow the Russian cybercriminal apparatus to thrive for a variety of reasons, including the fact that cybercrime brings money to Russia, and the talent base

[77] U.S. Department of the Treasury, "Treasury Sanctions Russia with Sweeping New Sanctions Authority," 15 April 2021, https://home.treasury.gov/news/press-releases/jy0127.

[78] Schroeder, E., Wilde, G., Sherman, and Herr, T., *Hackers, Hoodies, and Helmets: Technology and the Changing Face of Russian Private Military Contractors*, DFRLab, Cyber Statecraft Initiative, The Atlantic Council, Issue Brief, July 2022, p. 7, https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/technology-change-and-the-changing-face-of-russian-private-military-contractors/.

[79] A/76/151, p. 7, para. 21.

[80] U.S. Department of Justice, "U.S. Charges Russian FSB Officers and Their Criminal Conspirators for Hacking Yahoo and Millions of Email Accounts," 15 March 2017, https://www.justice.gov/opa/pr/us-charges-russian-fsb-officers-and-their-criminal-conspirators-hacking-yahoo-and-millions.

[81] U.S. Department of the Treasury, "Treasury Sanctions Russia with Sweeping New Sanctions Authority," 15 April 2021, https://home.treasury.gov/news/press-releases/jy0127.

[82] Menn, J., Stubbs, J., and Bing, C., "Chinese government hackers suspected of moonlighting for profit," *Reuters*, 07 August 2019, https://www.reuters.com/article/us-china-cyber-moonlighters-idUSKCN1UX1JE.

[83] Fraser, N., Plan, F., O'Leary, J., et al., "APT41, A DUAL ESPIONAGE AND CYBER CRIME OPERATION," *Mandiant*, 07 August 2019, https://www.mandiant.com/resources/blog/apt41-dual-espionage-and-cyber-crime-operation.

© iStock/standret

it cultivates gives the Kremlin proxies to tap as needed."[84] Such findings point to the fact that some of the world's most sophisticated cyber proxies increasingly pose a threat to civilian populations and to a large number of companies beyond those traditionally targeted by state-sponsored offensive cyberoperations.

In June 2022, at a Wall Street Journal Pro Cybersecurity Forum, U.S. senior officials confirmed that "the lines between criminal hacking groups and intelligence operations in countries like Russia, Iran and China have increasingly blurred, making Washington's job in curbing cyberattacks all the harder."[85] Attributing cyberattacks becomes even more complex with the transfer of expertise and techniques, the honing on similar targets, and the permeability between nation states' intelligence services, APT groups, and cybercriminals "for hire." In practice and on the frontline of cyberspace, there is growing competition and collusion between APT groups, private sector offensive actors, and cybercriminal syndicates that organize to monetize their services. This merging of sophisticated cyberthreat actors represents both a powerful challenge to governance and an acute, pervasive threat to civilian life and security.

- **Equally sobering, cybersecurity experts have warned of a problematic "industrialization" of the cybercrime economy (or its sub-economic ecosystems).** The *Microsoft Digital Defense Report 2022* notes that "cybercrime continues to rise as the industrialization of the cybercrime economy lowers the skill barrier to entry by providing greater access to tools and infrastructure."[86] The increasing number of online services (cybercrime as a service) that allows for

[84] See Sherman, *Untangling the Russian Web: Spies, Proxies, and Spectrums of Russian Cyber Behavior*, p. 5.

[85] Uberti, D., "Line Between Criminal Hackers and Nation-State Threats Blurs, U.S. Officials Say," *The Wall Street Journal*, 01 June 2022, https://www.wsj.com/articles/line-between-criminal-hackers-and-nation-state-threats-blurs-u-s-officials-say-11654109885.

[86] See *Microsoft Digital Defense Report 2022*, p. 7.

> Cybersecurity experts have warned of a problematic "industrialization" of the cybercrime economy (or its sub-economic ecosystems).

the outsourcing and automation of cybercrime activities is evidence of this industrialization. Diverse types of cyberattacks—from ransomware to phishing to DDoS—are now commoditized under affordable subscription models that come with encrypted safeguards and one year of 24/7 support. For instance, Microsoft's Digital Crimes Unit observes that "DDoS subscription service offers different architectures and attack methods, so a purchaser simply selects a resource to attack and the seller provides access to an array of compromised devices on their botnet to conduct the attack."[87] On its own, the rising industrialization of cybercrime has destructive potential, posing a nation-level threat to governments, industries, and populations. But its merger with nation-state proxy activity and the cyberarms industry constitutes a major threat amplification and creates even more difficulties in applying legal and accountability frameworks at international and national levels.

- **Both the underground cyberarms and cybercrime industries function primarily as competitive markets and are mostly and effectively ungoverned spaces.** Compared to other arms trades, these markets are extremely difficult to control. Crafting cyber weapons does not depend on trading physical equipment (e.g., guns and missiles), but on transferring data, malicious code, and dual-use expertise (for instance, how to design algorithms behind customized, adaptive malware or how to exploit vulnerabilities in computer codes). In a nutshell, hackers mainly sell services, skills, knowledge, and information that can be weaponized. When they provide access to technologies (such as tools for penetration testing, malware, or spyware), some may be "rented" but remain private intellectual property, others may be protected and sealed by encryption.[88] Some technical services can be "disguised" as legitimate defense programs, others transferred and acquired through third-party resellers in clandestine, opaque networks. An important number of offensive cyber technologies is also directly designed and customized through open-source platforms and tools (such as "proof-of-concept" malware on GitHub) and within dark web communities and underground market places.[89] Obviously, all these platforms escape regulation. Another proliferation challenge relates to the nature of cyber exploits: if they are disclosed, the algorithms, exploits, or techniques can be reverse-engineered and repurposed for adversarial or criminal operations. For instance, in 2020, the cybersecurity firm FireEye was hacked and its penetration testing tools stolen and repurposed.[90]

> An important number of offensive cyber technologies is also directly designed and customized through open-source platforms and tools and within dark web communities and underground market places.

Central to the dual-use paradigm in export control is the fact that knowledge and techniques in cybersecurity (e.g., how to protect from critical vulnerabilities), AI, and automation (e.g., how to automate cyberthreat detection) are crucially needed for cyber defense and innovation, yet they are also critical to design the most efficient cyberthreats we face today. Such intangible transfer of dual-use knowledge and expertise challenges non-proliferation architectures and traditional governance agreements such as export controls. As pointed by Perlroth, security researchers have long argued that "restrictions

87 See Idem, p. 19.

88 Interview with Sultan Meghji, cybersecurity expert at Carnegie Endowment for International Peace, 09 December 2022.

89 Ibid.

90 Sanger and Perlroth, "FireEye, A Top Cybersecurity Firm, Says It Was Hacked by Russians," *The New York Times*, 09 December 2020, https://www.nytimes.com/2020/12/08/technology/fireeye-hacked-russians.html.

on zero-days would actually handicap cybersecurity, in that it would keep researchers from sharing vulnerability research and malware across borders."[91] The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies was not originally conceived for cyberspace and does not capture most offensive cybertechnologies, including new techniques aimed at surveillance, data-exfiltration, obfuscation, or destruction.[92] Moreover, implementing the Wassenaar Arrangement would require participating states to develop national policies on domestic and international use, procurement, or production of advanced cyberweapons.

> "Such intangible transfer of dual-use knowledge and expertise challenges non-proliferation architectures and traditional governance agreements such as export controls.

In 2021, the European Union (EU) aimed to curb "surveillance for hire" and adopted a dual-use regulation, which introduces a new "end-use control" on cyber surveillance equipment.[93] Pursuant to the regulation, the exporter must undertake a due diligence assessment to ensure that the exported item is not to destined for internal repression or the commission of serious violations of human rights and international humanitarian law. However, such regulatory control faces serious challenges, in particular from the secrecy and opaqueness that characterize the cyber surveillance industry and its complex supply chains. Beyond the European Union, other cyber offense markets remain effectively unregulated and have welcomed providers that left the EU.[94] Investigations by private and public actors have demonstrated, for instance, how the NSO Group's powerful spyware, Pegasus, based on zero-click infection method, has enabled pervasive surveillance and human rights violations.[95] Self-regulation efforts by the NSO Group and required defense approval by Israel's Ministry of Defense have not been sufficient to protect populations and uphold human rights.[96]

How is the industrialization of cyber offense impacting cyber proxies' strategies and behaviors?

- **Pressure to monetize expertise and trade offensive cyberservices has led to increased knowledge transfer and interconnection between different types of cyberthreat actors and cyber proxies.** Competition, emulation, and more importantly, transfer of expertise is taking place between dark web communities, transnational cybercriminal gangs, APT groups, and other state-sponsored hostile actors. While some dark web communities remain closed-off and accessible only to privileged clients, the cybercrime underground is a complex and evolving ecosystem with no sharp borders, but rather, connections and transfer of expertise across hackers' communities.[97] Expertise in offensive cyber capabilities tends to move from dark web communities to the surface web. The underlining logic is that brokers of illicit exploits, such as "proof-of-concept" malware and zero-day vulnerabilities, need to market their products, and therefore need to advertise them on the surface web (including through online markets, social media channels, chat rooms, forums, paste sites, and open-source platforms like GitHub).

91 See Perlroth, *This Is How They Tell Me the World Ends*, p. 149.

92 See Idem, p. 150-151.

93 Council of the European Union, "Trade of dual-use items: new EU rules adopted," 10 May 2021, https://www.consilium.europa.eu/en/press/press-releases/2021/05/10/trade-of-dual-use-items-new-eu-rules-adopted/.

94 See Perlroth, *This Is How They Tell Me the World Ends*, p. 151.

96 Human Rights Council (41st Session), *A/HRC/41/35: Surveillance and human rights - Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 2019, p. 10, para. 31-32, https://www.ohchr.org/en/documents/thematic-reports/ahrc4135-surveillance-and-human-rights-report-special-rapporteur.

96 Ibid.

97 See Recorded Future, *EXECUTIVE BRIEF: Inside the Dark Web*, 2022, https://go.recordedfuture.com/hubfs/white-papers/inside-dark-web.pdf?hsLang=en.

The present report highlights the "industrialization" of cyber offense based on what experts detect on the ground: increased forms of trading, collaboration, and sharing between cyberthreat actors, including APT and cybercrime groups.[98] For instance, in 2022, the cybersecurity firm TrendMicro tracked multiple malicious campaigns by several APT41 subgroups, studying their tactics, techniques, and procedures (TTPs).[99] Such cyber forensics analysis reveals that there is substantial sharing of techniques and tools, such as the commodity malware licensed as "Cobalt Strike" in clandestine markets between China's APT groups.

A similar fusion of tools and techniques takes place in Russia's cyberthreat actors' ecosystem. The cybersecurity firm, Mandiant, reports instances of multiple Russian threat groups engaging in intrusion operations within close temporal proximity, sometimes even using the same access method within hours or minutes of each other.[100] The Russian criminal APT group, FIN7, is extremely polymorphous and a good example: it has not only incorporated and merged numerous cybercrime units, but also used a front information technology (IT) company as cover, developed ties with the cybercrime syndicate REvil, adopted offensive techniques from ransomware-as-a-service operations, and extended its attack surface to different civilian sectors, including the oil industry.[101] Equally sobering is the fact that FIN7 has shared its advanced offensive techniques with other ransomware groups, enabling them to be more effective and de facto augmenting the potential for chaos, civilian harm, and insecurity in cyberspace. Expert Adam Flatley, Vice President of Threat Intelligence at [redacted], explains how "they ([FIN7] are making the ransomware problem worse by continuing to bring their nearly nation-state level talent to the table as a force multiplier."[102]

Rigorous documentation of technical evidence is critical for modern attribution of cyber offense. As explained by TrendMicro, with the increasing transfer of expertise and techniques, "tool-based attribution and analysis will likely become more complicated and will be a challenge to threat researchers in figuring links among different groups."[103]

- **Most sequences in the anatomy of a cyberattack—design, intrusion, deployment, extortion—can be outsourced to different cyber-skilled actors, which augments the opacity of the supply chain for offensive and intrusive cyberoperations**. It becomes increasingly complex to trace which actors are designing, deploying, sponsoring, and ordering offensive operations and which "suppliers" are entangled in operations that could make them liable for the harm.

  Moreover, cyber-skilled actors in the clandestine cyberarms industry and the criminal underground are profit-driven. No protocols (such as "Know Your Client"), due diligence, or ethical code will prevent them from selling their offensive malware and techniques even to opaque buyers.[104] In these fluid networks, providers communicate with beneficiaries by relying on multiple fake identities and several levels of intermediaries and dissociation. They connect, operate, and deploy from dislocated geographies and may use different adversarial vectors. And since an actor can buy the components of a cyber-exploit kit from different groups, no "supplier" will have a sense of what the code is

98 Interview with Sultan Meghji on 09 December 2022. See also McGuire.

99 See Hiroaki and Lee, "Hack the Real Box: APT41's New Subgroup Earth Longzhi."

100 Arghire, I., "FIN7 Cybercrime Operation Continues to Evolve Despite Arrests," *SecurityWeek*, 06 April 2022, https://www.securityweek.com/fin7-cybercrime-operation-continues-evolve-despite-arrests/.

101 See Greig, J., "FIN7 cybercrime cartel tied to Black Basta ransomware operation: report," *The Record*, 03 November 2022, https://therecord.media/fin7-cybercrime-cartel-tied-to-black-basta-ransomware-operation-report/. See also Cimpanu, C., "FIN7 hacker trialed in Russia gets no prison time," *The Record*, 30 November 2021, https://therecord.media/fin7-hacker-trialed-in-russia-gets-no-prison-time/.

102 See Arghire, "FIN7 Cybercrime Operation Continues to Evolve Despite Arrests."

103 See Hiroaki and Lee, "Hack the Real Box: APT41's New Subgroup Earth Longzhi."

104 See Perlroth, *This Is How They Tell Me the World Ends*, p. 14-17, 144-145, and 155-157.

> Most sequences in the anatomy of a cyberattack—design, intrusion, deployment, extortion—can be outsourced to different cyber-skilled actors, which augments the opacity of the supply chain for offensive and intrusive cyberoperations.

intended for and whom it aims to target. Such opacity within the supply chains for offensive and intrusive cyberoperations drastically complicates potential due diligence efforts by providers of cyber defense and cyber offense services. A good example of "dissociation" is a previously unknown Lebanon-based group, which Microsoft has assessed as being a cyber proxy for Iran's Ministry of Intelligence and Security. According to Microsoft, "Such collaboration or direction from Tehran would align with revelations since late 2020 that the Government of Iran is using third parties to carry out cyber operations, likely to enhance Iran's plausible deniability."[105]

Many APT groups have also shifted from customized malware (crafted in-house) to commodity malware (sold in underground markets) and open-source penetration testing tools to obfuscate their cyber intrusions. The other added benefit to using open-source exploit kits is that development and incremental innovation is done by someone else at no cost.

The integration of AI to improve the stealth of cyberattacks may provide another way to divert and obfuscate attribution of responsibility, especially when autonomous malware or automated methods of operation (cybercrime as a service) are used. In an AI-led cyberattack, autonomous malware could infect a large network of computers and devices and command the network to compromise other targets. To trace, document, and characterize the automated attack, investigators would have to obtain the cooperation of several countries and jurisdictions. The *Microsoft Digital Defense Report 2022* discusses the multi-jurisdictional nature of offensive cyberoperations and cybercrime as a service with threat actors collaborating across languages and time zones, noting "[f]or example, one cybercrime-as-a-service website administered by an individual in Asia maintains operations in Europe, and creates malicious accounts in Africa."[106]

In general, trends in decentralization, automation, and outsourcing make it difficult to detect and trace offensive cyber capabilities, neutralize large groups of threat actors, and attribute wrongful and illicit conduct across jurisdictions. In its 2021 report, the UN Group on the use of mercenaries notes that "the possibility that cyber proxies may move across borders and thus escape regulatory control and accountability mechanism is serious cause for concern."[107]

How are nations states engaging with the underground cyberarms industry and the interconnected cybercrime economy?

- **Nation states are increasingly trading or sharing exploits, as well as contracting and outsourcing services to a growing number of clandestine proxy actors in the underground cyberarms and cybercrime industries.**[108] States are not only adopting techniques, tools, expertise, and services provided by cybercriminals, but also leaking datasets, penetration tools, and cyber exploit kits to malicious groups, enhancing their offensive skills through covert technical and financial support.[109] The 2021 cyberattack on Microsoft Exchange Servers orchestrated by Hafnium, a Chinese APT group, is one prominent example of strategic sharing of worldwide threats. According to cybersecurity experts, it is likely that HAFNIUM

[105] See *Microsoft Digital Defense Report 2022*, p. 48.

[106] See Idem, p. 18.

[107] A/76/151, p. 11, para. 36.

[108] Interview with Sultan Meghji on 09 December 2022. See also McGuire, *Nation States, Cyberconflict, and the Web of Profit*, p. 18-19. See also Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 27-28.

[109] See McGuire, *Nation States, Cyberconflict, and the Web of Profit*, p. 18-19.

leaked the unpatched Microsoft vulnerabilities to be aggressively exploited by a half-dozen APT groups in the same region and some with ties to China's government.[110] The 2021 attack aimed at economic and industrial espionage resulted in hacks of hundreds of thousands of victims across 115 countries. The intrusion targeted both public and private sectors and a range of industries, including agriculture, biotechnology, aerospace, defense, legal services, power utilities, and pharmaceutical.

Beyond serving the strategic interests of nation states, such massive exploitation tactics seem designed to increase—or at least, contribute to—economic loss, disruption, and levels of cyberinsecurity for governments, populations, and industries worldwide.[111] As Maurer observes, "the relevance of proxies lies not only in their ability to cause harm but also in their ability to wield power more broadly."[112]

Adversarial methods such as ransomware or DDoS that are industrialized by cybercriminal groups are increasingly used in nation state attacks against geostrategic targets, including governments (such as *inter alia* the U.S., Ukraine, Albania, Israel, Canada, and Australia) and international institutions (including the UN, the International Monetary Fund, and the World Health Organization).[113] Iran- and North Korea-based APT groups have harnessed commodity ransomware tools to damage targeted systems, including critical infrastructures such as transportation, energy, and water treatments, within regional and international reach.[114] Zero-day vulnerabilities and malicious tools (malware strains, bots, keyloggers, and spyware) usually sold in underground cybercrime communities are now weaponized by nation states' proxy actors. Some dark web marketplaces also list specialized services according to the needs, targets, and languages of domestic state actors. For instance, a penetration testing tool like PowerShell Empire that was harnessed by cybercriminals in 2018 ended up in the hands of nation state proxies, such as the China-linked APT group Gadolinium involved in a 2020 set of cyberattacks to acquire intellectual property data related to COVID-19 vaccines' production.[115]

> "Adversarial methods such as ransomware or DDoS that are industrialized by cybercriminal groups are increasingly used in nation-state attacks against geostrategic targets.

As McGuire explains, **offensive tools and techniques developed by nation state proxies are also traded, leaked to cybercrime syndicates, or incidentally benefit those groups.**[116] In 2017, the EternalBlue exploit stockpiled by the U.S. National Security Agency was first leaked by a Russian-sponsored APT group, the Shadow Brokers, and then turned into an extremely lucrative cybercrime tool. Global losses from the leak of EternalBlue have amounted to several billion dollars for public and private sector actors, while it exceeded $500 million revenues for cybercriminals. In 2020, the SolarWinds supply-chain-attack that targeted the U.S. government and that was attributed by the U.S. and UK to the Russian Foreign Intelligence Service has led to sensitive data being "leaked" and highly priced on the dark web.

[110] Goodin, D., "There's a vexing mystery surrounding the 0-day attacks on Exchange servers," *Ars Technica*, 11 March 2021, https://arstechnica.com/gadgets/2021/03/security-unicorn-exchange-server-0-days-were-exploited-by-6-apts/.

[111] Ravich, S. and Fixler, A. (eds.), *The Attack on America's Future – Cyber-Enabled Economic Warfare*, Foundation for Defense of Democracies, October 2022, https://www.fdd.org/analysis/2022/10/28/the-attack-on-americas-future-cyber-enabled-economic-warfare/.

[112] See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. XII.

[113] See McGuire, *Nation States, Cyberconflict, and the Web of Profit*, p. 18.

[114] See *Microsoft Digital Defense Report 2022*, p. 47-51.

[115] See McGuire, *Nation States, Cyberconflict, and the Web of Profit*, p. 19.

[116] Ibid.

> Offensive tools and techniques developed by nation state proxies are also traded, leaked to cybercrime syndicates, or incidentally benefit those groups.

- **There are signals that some nation states may abuse confidence-building measures and normative requirements to their advantage. In a recent report, Microsoft accused APT groups in China of abusing the country's vulnerability disclosure requirements in an effort to stockpile and harness critical zero-day vulnerabilities**.[117] In 2021, the Cyberspace Administration of China (CAC) announced more stringent procedures related to disclosing vulnerabilities for companies based on its territory.[118] For instance, Alibaba was sanctioned for reporting directly to Apache—and not first to the Chinese government—the vulnerability in the open-source software tool Log4j, which is widely used by companies worldwide.[119] Microsoft observed that the aggressive use of zero-day exploits by China-linked APT groups over the 2021-2022 period corresponds to the first full year of application of the new CAC rule for reporting software vulnerabilities.[120]

- **In recent years, nation states have found new ways to monetize the revenues of the cybercrime economy, using such cash flow to support in-house development of sophisticated cyberweapons and to invest in other military and tech-driven domains**. Cyber industrial espionage, data theft and trading, illicit mining of cryptocurrencies, and provision of offensive cyberservices are all methods employed in some nation states' arsenal to increase profit. In 2017, the North Korean APT Group Lazarus generated approximately 571 million USD from targeting cryptocurrency operations.[121] In 2016, North Korean APT groups siphoned 81 million USD from Bangladesh's Central Bank.[122] In 2021-2022, North Korea's APT group COPERNICUM specialized in large state-sponsored theft operations against cryptocurrency companies worldwide to generate domestic revenue.[123] A confidential 2021 UN report mentions that North Korea's cyber proxy actors allegedly generated hundreds of millions of dollars of revenue throughout much of 2020 to fund the country's nuclear and ballistic missile programs in violation of international law.[124] Similar profit-making strategies—while not on the same scale—have been used by the Iranian APT group PHOSPHORUS to execute ransomware attacks on critical infrastructures in the United States and other Western nations.[125] Russia's very active cybercrime ecosystem is also a growing source of revenue. Decades ago, Russia built revenue-generating capacities, and already in 2014, its cybercrime ecosystem generated billions in USD.[126] As Sherman mentions, "In 2021 alone, it was reported that 74% of global ransomware revenue went to Russian hackers, to the tune of 400 million USD in cryptocurrencies."[127]

Nation states are also turning to increased asymmetric industrial competition and cyber-enabled

[117] See *Microsoft Digital Defense Report 2022*, p. 39-40.

[118] Cimpanu, "Chinese government lays out new vulnerability disclosure rules," *The Record*, 13 July 2021, https://therecord.media/chinese-government-lays-out-new-vulnerability-disclosure-rules/.

[119] Greig, "Chinese regulators suspend Alibaba Cloud over failure to report Log4j vulnerability," *ZDNet*, 22 December 2021, https://www.zdnet.com/article/log4j-chinese-regulators-suspend-alibaba-partnership-over-failure-to-report-vulnerability/.

[120] Greig, "Microsoft accuses China of abusing vulnerability disclosure requirements," *The Record*, 03 November 2022, https://therecord.media/microsoft-accuses-china-of-abusing-vulnerability-disclosure-requirements/.

[121] See McGuire, *Nation States, Cyberconflict, and the Web of Profit*, p. 20.

[122] Ibid.

[123] See *Microsoft Digital Defense Report 2022*, p. 49.

[124] Roth, R. and Berlinger, J., "North Korean hackers stole more than $300 million to pay for nuclear weapons, says confidential UN report," *CNN*, 09 February 2021, https://edition.cnn.com/2021/02/08/asia/north-korea-united-nations-report-intl-hnk/index.html.

[125] See *Microsoft Digital Defense Report 2022*, p. 46-47.

[126] Maurer, "Why the Russian Government Turns a Blind Eye to Cybercriminals," Carnegie Endowment for International Peace, 02 February 2018, https://carnegieendowment.org/2018/02/02/why-russian-government-turns-blind-eye-to-cybercriminals-pub-75499.

[127] See Sherman, *Untangling the Russian Web: Spies, Proxies, and Spectrums of Russian Cyber Behavior*, p. 9.

economic warfare.[128] For instance, industrial cyberespionage by China's APT groups tends to align with the country's Five-Year Plan, focusing particularly on the economic and military domains where competitive advantage could be gained, such as the pharmaceutical/health, biotechnology, AI, 5G and advanced ICT, defense, and aerospace sectors.[129]

IMPLICATIONS

How are recent geostrategic and technological trends influencing cyberconflict and impacting the regulation of cyber proxy activity?

Cyber proxy activity is becoming increasingly difficult to decrypt, trace, and attribute. The frameworks used to categorize forms of deputization in cyberspace do not capture well the increased permeability and the intense knowledge and technical transfer that exist among non-state actors, as well as between states and non-state actors. The polymorphous, multi-jurisdictional nature of cyber proxy activity therefore drastically complicates technical and, even more so, legal attribution of wrongful conduct in cyberspace.

The consequence is that deniability remains more than ever a winning strategy for states using cyber proxies to advance their geostrategic interests, particularly in the absence of an independent and multilaterally-recognized attribution authority. The normative gap that will persist for the coming years in this regard—coupled with the potential involvement of decentralized private actors in the design, management, and procurement of dual-use technologies—will give room to new types of abuses being left unaddressed.

Increasingly, a powerful and strategic logic behind aggressive cyber proxy activity is to conduct economic warfare. A 2022 report from the Foundation for Defense of Democracies warns of the implications of cyber-enabled economic warfare. Highlighting the example of China, the report notes that "China seeks to dismantle the U.S. and allied stake in these markets [5G, AI, cloud-computing, semiconductors, etc.] through cyber espionage and sabotage as well as non-market coercion so that Beijing can control key nodes in the global economy."[130] In general, China has expanded its global offensive cyberoperations, targeting an increasing number of countries, particularly in South Asia, but also in Africa, the Middle East, and Oceania.

> "Deniability remains more than ever a winning strategy for states using cyberproxies to advance their geostrategic interests.

Another implication is the potential erosion of states' will to cooperate around norms of responsible behavior and confidence-building measures. In countries where the private sector is not independent from state power, confidence-building and reporting measures could potentially be abused to stockpile and weaponize critical cyber exploits. Harnessing zero-day vulnerabilities has been part of China's increasingly aggressive behavior in cyberspace. According to U.S. National Security Agency cyber chief Rob Joyce, in 2021-2022, China's APT groups were "really brazen, doubling down on their activities to steal intellectual property and compromise sensitive networks; they establish persistence and move laterally across the interconnected networks so malicious state sponsored activity is a major threat to U.S. critical infrastructure, election systems, national security systems and the Department of Defense along with the defense industrial base that we help protect."[131]

Despite alleged Russian arrests of prominent cybercriminals, the rapid resurgence and re-combination of skills, expertise, and cybercrime as a service shows little prospect of an effective governance effort to

[128] "Russian, Chinese hackers targeted Europe drug regulator: newspaper," *Reuters*, 06 March 2021, https://www.reuters.com/article/us-eu-cyber-idUSKBN2AY0F1.

[129] Culafi, A., "The wide web of nation-state hackers attacking the U.S.," *TechTarget*, 20 April 2021, https://www.techtarget.com/searchsecurity/news/252499613/The-wide-web-of-nation-state-hackers-attacking-the-US.

[130] Ravich and Fixler, *The Attack on America's Future – Cyber-Enabled Economic Warfare*, p. 8.

[131] Smalley, S., "Chinese state-sponsored hackers have become more brazen, prompting an NSA advisory," *Cyberscoop*, 6 October 2022, https://cyberscoop.com/chinese-state-sponsored-hackers-more-brazen/.

curb cartels and their profit-making activity.[132] Some experts even allude to the fact that official arrests of cybercriminals by Russia have been part of a strategic move to "clean up," or better, to reshape the cyber proxy ecosystem.[133] The current geopolitical context with Russia's invasion of Ukraine offers even less hope.

The intangible transfer of knowledge, expertise, and technology taking place in the underground cyberarms and cybercrime industries also directly challenges the application of export control and market regulation mechanisms within non-proliferation and disarmament architectures. The 2022 Microsoft report notes that "because these offensive cyber capabilities are no longer highly classified capabilities created by defense and intelligence agencies, but rather commercial products now offered to companies and individuals, any regulatory regime for cyberweapons needs to move beyond export control."[134] There is a need for regulatory innovation on this national and international security front.

The increased connection between nation state-sponsored cyber offense with the underground cybercrime economy indicates an opportunity to look at how international and regional cooperation mechanisms to prevent cybercrime could be leveraged to regulate cyber proxy activity.[135] While they have sometimes lacked coordination, capacity, and financial support, law enforcement and security authorities have still collaborated for decades at domestic, regional, and international levels (for instance, with the support of Europol and INTERPOL and in collaboration with the private sector). Their strategies have included collaborations across jurisdictions to seize cybercrime infrastructure (for instance, servers and websites of underground marketplaces or networks of compromised computers also called "botnets"), as well as to share intelligence, coordinate operations, arrest, and extradite cybercriminals. Cybercrime prevention might therefore provide opportunities to respond to the proliferation of offensive cyber proxies more rapidly than through UN cyber diplomacy, including inter-state normative processes and inter-state positioning on international law. The next section will assess the efficacy of renewed efforts in the cybercrime prevention domain with the International Counter Ransomware Initiative and Task Force.

The complex, interrelated trends that drive cyber offense and cyber proxy activity should raise questions about the role of the civilian private sector and its importance in strengthening cyber-resilience, protecting governments, populations, and industries. There is a need to assess the leverage that those actors have—in partnership with defense and security authorities—when it comes to technical and legal attribution, naming and shaming, indictments and prosecution, and collective and coordinated responses in active cyber defense.

The implications of rapidly expanding, unregulated markets for cyber offense remain corrosive to international peace and security with a potential rise in cyber-mercenary operations and cyberterrorism. Offering the means of cyberaggression to every actor who can afford it is already transforming how contemporary conflicts are fought. States and non-state actors are drastically empowered through cyber offense; the relationship between both is less asymmetrical, with increased diffusion of power. As a result, the potential beneficiaries of the underground cyberarms and cybercrime industries may include private mercenary groups, terrorist groups, transnational illicit networks, and proxy forces involved in conflict. **First**, such "diffusion of cyber power" will rapidly reach increasing numbers of **private sector offensive actors and private groups associated with mercenary activity** (for instance, the Wagner Group which is already actively involved in spreading global disinformation campaigns and leveraging influence operations).[136] Harmful implications are already seen through infor-

[132] Cimpanu, "FIN7 hacker trialed in Russia gets no prison time."

[133] See Sherman, *Untangling the Russian Web: Spies, Proxies, and Spectrums of Russian Cyber Behavior*, p. 5-6.

[134] See *Microsoft Digital Defense Report 2022*, p. 52.

[135] See Hakmeh, J. and Vignard, K., *ICTs, International Security, and Cybercrime*, United Nations Institute for Disarmament Research, October 2021, https://unidir.org/publication/icts-international-security-and-cybercrime.

> The implications of rapidly expanding, unregulated markets for cyberoffense remain corrosive to international peace and security with a potential rise in cyber-mercenary operations and cyberterrorism.

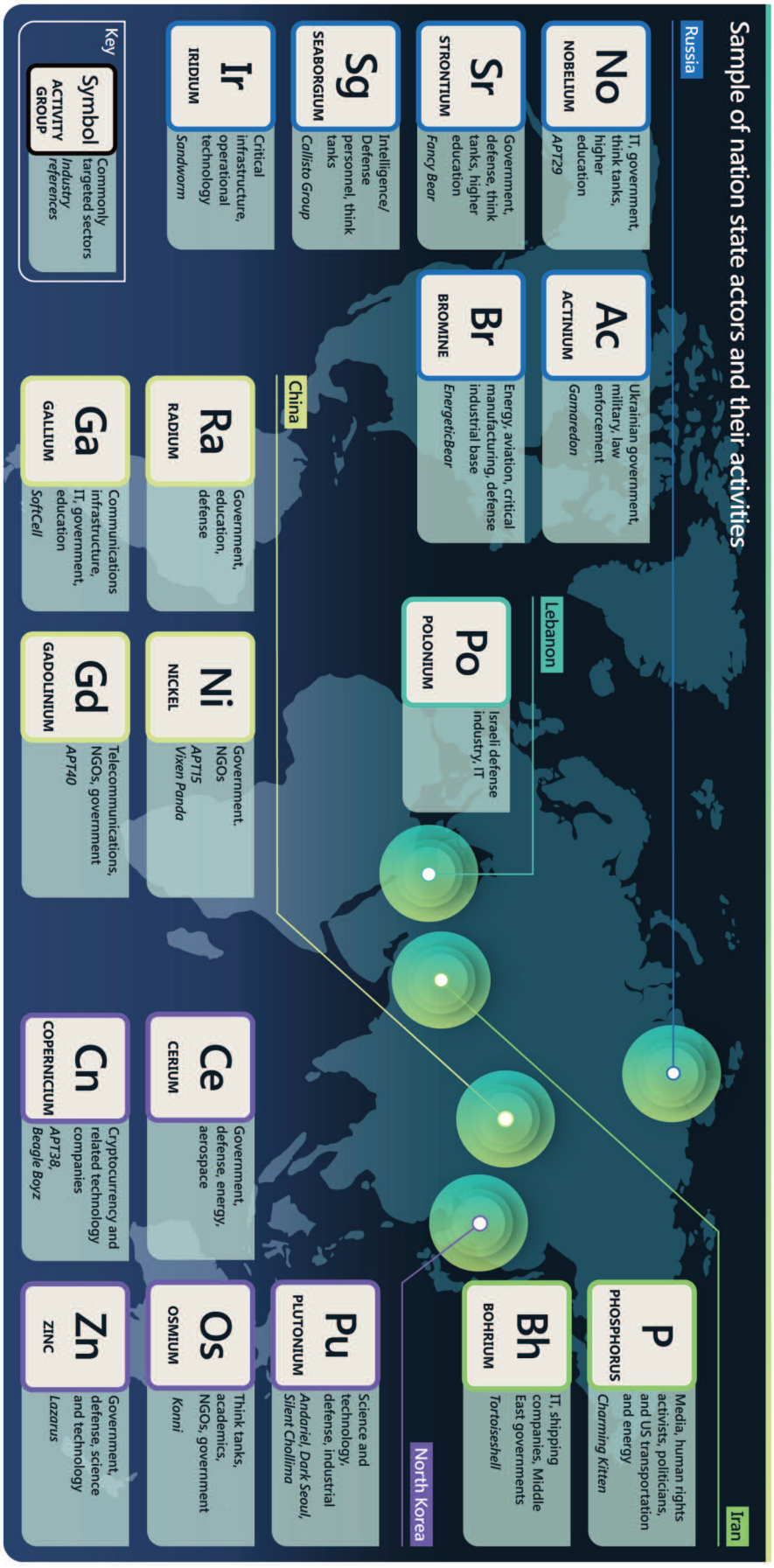
mation operations waged in different countries across the globe. **Second, terrorist organizations** may acquire, exploit, or outsource cyber services to support their offensive agenda (for example, disinformation, surveillance, ransomware, and zero-day attacks as a service). And for their defensive and operational tactics, terrorist organizations may learn to use cybercrime techniques (such as monetizing data leaks and crypto-jacking) to fund their physical, on-the-ground operations. Therefore, there might be an increasing correlation between criminal accessibility and mercenary and terrorist capability.

In general, several factors augment the risk of escalation and the prospect to face more advanced cyberconflict: (1) proliferation of harmful cyber proxy activity, (2) lack of control over unregulated non-state actors that compete and are often in tension, and (3) high level of strategic engagement and rising stakes for nation states that sponsor such activity. The consequences could be devastating. The ICRC has confirmed that cyber operations without—or with unclear—links to armed conflicts have had damaging impacts on civilian infrastructure, from hospitals, water, and electrical facilities to nuclear and petrochemical plants. It has further noted that known cyber incidents of all types "offer a chilling warning about the potential humanitarian impact of military cyber operations in contemporary and future armed conflicts."[137] The UN Working Group on the use of mercenaries concurred with this analysis in its 2021 report, warning that "emerging forms of warfare can have a significant impact on both military objectives and civilian populations and can result in violations of international humanitarian law as well as the rights and freedoms of individuals in the context of armed conflicts and otherwise."[138]

[136] U.S. Department of State, "Wagner Group, Yevgeniy Prigozhin, and Russia's Disinformation in Africa," 24 May 2022, https://www.state.gov/disarming-disinformation/wagner-group-yevgeniy-prigozhin-and-russias-disinformation-in-africa/.

[137] ICRC, *Avoiding Civilian Harm*, p. 6.

[138] A/76/151, p. 17, para. 61.



Source: *Microsoft Digital Defense Report 2022*, p. 34. The periodic table of chemical elements is the taxonomy used for decades by Microsoft to identify and name cyberthreat actors.

LEGAL SECTION: APPLICATION OF INTERNATIONAL LAW TO CYBER PROXY OFFENSIVE OPERATIONS

The previous section has shed light on the harmful convergence between the cyberarms and cybercrime industries and the offensive proxy capacities they bring to a growing number of nation states and potentially violent actors. We face an unprecedented confluence of powerful technologies, strategic ambiguity, political and economic warfare, espionage, and increased funding of military programs, including in some countries by illicit means such as revenue-generating cybercrime capacities. Such confluence is starting to pose unique challenges in how to regulate cyber proxy activity, especially finding common normative approaches which can lessen tensions between nations states and help strengthen accountability.

Regulating the role and involvement of cyber proxies in offensive cyberoperations on the basis of international law poses complex and unresolved challenges. **Three types of legal ambiguities** enter into play: ambiguities related to the interpretation of international law; ambiguities as to how international law applies to cyberspace and cyberwarfare; and ambiguities related to wrongful conduct in cyberspace by *non-state* actors. Public international law aims to regulate the interaction between states, not natural persons or private entities. Under international law, the existing standards and tests for attributing the acts of non-state actors to states remain extremely high, require substantial technical evidence, and are therefore difficult to apply in practice. As Maurer summarizes, "A state can only be held accountable for the offensive actions of a cyber proxy if that proxy is under tight control of a government and if the effect of the action causes significant harm."[139] Yet, determining whether wrongful conduct in cyberspace can be attributed to a state or a non-state actor has important legal consequences. Such an attribution process is critical in deciding which body of international law and related thresholds are relevant, and to ensure that actors who conduct unlawful cyberoperations can be held accountable.

> "Determining whether wrongful conduct in cyberspace can be attributed to a state or a non-state actor has important legal consequences.

THE ATTRIBUTION PROBLEM

Can the source of cyberattacks or other offensive cyberoperations be identified? Can cyberattacks or other offensive cyberoperations be attributed to particular persons or entities? Is it possible to identify and qualify with evidence the relationship between the cyber proxy actor(s) and the state on behalf of which such activities are undertaken, if at all?

- Cybersecurity companies have developed sophisticated methods in forensics analysis to track malware families and threat actors' "tactics, techniques and procedures" and are lending expertise to other actors in the public and private sectors. Yet, **evolving behaviors**

[139] See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 123.

and obfuscation strategies by cyber proxy actors pose a challenge to cyber forensics analysis.[140] For instance, APT groups rely on open-source and commodity malware usually sold in dark web markets and previously mostly exploited by cybercriminals. They use automated AI-led attacks, cybercrime as a service, impersonation, and false flags to lead to misattribution; they try to avoid the presence of watermarks and signatures on their offensive tools to reduce their exposure. It is not uncommon for members of APT groups to spend three days of the week "working" as cyber proxies for a nation state and two moonlighting for personal gain by conducting ransomware or crypto-theft operations. Another source of ambiguity stems from the fact that companies in the cyberarms industry often implement substantial parts of their offensive services, taking on practical responsibility in operations. Moreover, zero-day vulnerabilities are not exclusively traded by private sector offensive actors—which regulators could attempt to track—but also shared between APT groups and surreptitiously leaked to cybercriminal syndicates. **Consequently, there are less and less clear patterns of deputization, delegation, subordination, and control between nation states and their proxies**. Rather, what we see emerging are clandestine and polymorphous ecosystems for cyber offense.

- **Another problem in identifying the source and actors behind cyberattacks lies in extraterritoriality**.[141] Proxy actors conducting offensive cyberoperations for a nation state tend to operate at distance from that state and can design their adversarial vectors to transit through the digital systems of different states before reaching the targeted victims. They may also use different covert operative cells in hotspots across the globe. Some countries are considered safe havens, because they have not criminalized malicious cyber activities in their domestic laws or do not have the capability to effectively enforce cybercrime laws. As a result, pervasive challenges remain. Only a few tech-leading nations and their companies have the expertise to perform robust attribution—lending such capacity can lead to geopolitical tensions—and there is **no overarching attribution authority that is recognized multilaterally**.

- The reliance of States on **cyber proxies with unclear legal status** also augments the potential to obfuscate political and legal responsibility. For instance, a substantial number of companies in the cyberarms industry present themselves as providers mainly of cybersecurity and defensive services, concealing their offensive role.[142] Certain private sector offensive actors might not be officially registered as private military and security companies or as procuring mercenary-related services, and their status might change depending on the country that hires their services.[143] They may act without clear or traceable supervision and with something close to impunity.[144] The CyberPeace Institute mentions the need for "more clarity over the legal status of military and security services provided in cyberspace in order to determine the potential and actual impact upon human security, dignity, and equity."[145] The Institute emphasizes how "The obscurity of the cyber capabilities market makes it impossible to track their deployment in practice and eludes attempts to introduce oversight mechanisms, which are needed in order to regulate the use of cyber capabilities themselves."[146]

[140] Interview with Sultan Meghji on 09 December 2022. See McGuire, *Nation States, Cyberconflict, and the Web of Profit*. See also Perlroth, *This Is How They Tell Me The World Ends: The Cyberweapons Arms Race*. See *Microsoft Digital Defense Report 2022*.

[141] See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 132-137.

[142] See Perlroth, *This Is How They Tell Me The World Ends: The Cyberweapons Arms Race*, p. 149-177.

[143] Consider the example of the renowned Wagner Group that has been tied to the Russian military intelligence agency (GRU). As explained by Professor Kimberly Marten in her 2022 testimony to the U.S. House of Representatives, "The Wagner Group is neither a legally established private military and security company (PMSCs remain illegal in Russia), nor a true mercenary group," but an "opaque mechanism to serve at the behest of the Russian state, often with funding and contracts signed by foreign authoritarian regimes (such as in Sudan, CAR and Mali)." Testimony available at https://barnard.edu/news/professor-marten-publishes-policy-memo-yevgeny-prigozhin-and-wagner-groups-involvement-africa.

[144] A/HRC/41/35, p. 3, para. 1; p. 6, para. 15-17; and p. 10, para. 29-32.

[145] CyberPeace Institute, "Mercenary-related Activities in Cyberspace."

[146] Idem.

In a nutshell, the increased permeability between different types of cyberthreat actors and the blurring of their modus operandi severely complicates the production of evidence for legal attribution and regulation of cyber proxy activity across borders and jurisdictions. Production of evidence, multilateral processes, and legal frameworks are therefore seriously challenged when it comes to ascribing legal responsibility and accountability for wrongful conduct, determining liability for harmful impact, and obtaining remedial action. As stated by the UN Working Group on the use of mercenaries, "the issue of the attribution of cyberoperations and the matter of the intentional dissociation of such operations from State armed forces, such that there can be 'plausible deniability,' is patently a serious problem in advancing regulation."[147]

> "The increased permeability between different types of cyberthreat actors and the blurring of their modus operandi severely complicated the production of evidence for legal attribution and regulation of cyberproxy activity across borders and jurisdictions.

The below sections will highlight the specific instances when international law could be leveraged to regulate cyber proxies' offensive activity. The goal is to begin clarifying answers to the following questions: When is a state responsible for the actions of a non-state actor under existing bodies of international law? And in situations where state responsibility cannot be established, can cyber proxy actors be prosecuted for conducting offensive cyberoperations and under which conditions?

The below sections will also unveil the remaining ambiguities and limits in the international legal framework through Case Study 1 (cyber proxy activity with a nexus to international armed conflict), Case Study 2 (cyber surveillance and information operations in situations of armed conflict), and Case Study 3 (offensive cyberoperations in the grey zone).

JUS AD BELLUM AND STATE RESPONSIBILITY

Relevant to the regulation of offensive cyberoperations is Article 2(4) of the Charter of the United Nations, which prohibits the threat or use of force against the territorial integrity or political independence of any State.[148] Article 51 of the UN Charter also states that the use of force is only permitted in self-defense, in response to an "armed attack," or with the authorization of the UN Security Council.[149] The argument that there exists a "gravity threshold," below which the prohibition of the use of force is inapplicable, has gained ground in legal doctrine. Under certain circumstances, offensive cyberoperations may qualify as "use of force" or meet the threshold of an "armed attack." As Michael Schmidt explains, "a victim State may respond in self-defense if a non-state actor conducts a cyber armed attack on behalf *or with the substantial involvement* of a State."[150] The 2021 report by the UN Working Group on the use of mercenaries confirms that "Whether such cyber activities meet the relevant thresholds, in particular regarding the principles of necessity and proportionality, is a question of fact and degree but there can be little doubt that, given the nature and effects of modern cyber activities, they could satisfy those thresholds in particular circumstances."[151]

UN Member States are progressively positioning themselves and indicating how they would interpret this notion of "fact and degree" or "scale and effects." For instance, France and Norway have expressed the opinion that cyberoperations could amount to "use of force" if they targeted on a large scale and

[147] Idem, p. 12, para. 39.

[148] Charter of the United Nations, Article 2(4), https://www.un.org/en/about-us/un-charter/full-text.

[149] Idem, Article 51.

[150] Schmidt, M., "Expert Backgrounder: NATO Response Options to Potential Russia Cyber Attacks – Understanding the Legal Framework" (see Section B on "Response Options"), *JustSecurity*, 24 February 2022, https://www.justsecurity.org/80347/expert-backgrounder-nato-response-options-to-potential-russia-cyber-attacks/.

[151] A/76/151, p. 13, para. 41.

severely harmed their national economy.[152] Germany has suggested precise criteria that could be taken into account in assessing scale and effects: "The determination of a cyber operation as having crossed the threshold of a prohibited use of force is a decision to be taken on a case-by-case basis. Based on the assessment of the scale and effects of the operation, the broader context of the situation and the significance of the malicious cyber operation will have to be taken into account. Qualitative criteria which may play a role in the assessment are, inter alia, the severity of the interference, the immediacy of its effects, the degree of intrusion into a foreign cyber infrastructure and the degree of organization and coordination of the malicious cyber operation."[153]

In his analysis, Schmidt mentions that "the United States and numerous other States, including key NATO [North Atlantic Treaty Organization] members (the United Kingdom and Germany), take the position that so long as cyberoperations reach the armed attacked level of severity, the victim has the right of individual self-defense and may look to other States for assistance in collective self-defense."[154] In June 2022, NATO's Strategic Concept confirmed that a single or cumulative set of hostile cyberoperations could reach the level of armed attack and lead the North Atlantic Council to invoke Article 5 of the North Atlantic Treaty, which serves as the fundamental basis for the collective defense of NATO member states.

Qualifying offensive cyberoperations as an internationally wrongful act is one important step. **When it comes to hostile cyber proxy activity, another question is whether the wrongful conduct of non-state actors would engage the relevant provisions of the law on the use of force**. The UN Charter regulates the conduct of sovereign states, not the conduct of private entities. As stipulated in Article 8 of the ILC Articles on State Responsibility: "The conduct of a person or group of persons shall be considered an act of a State under international law if the person or group of persons is in fact acting on the instructions of, or under the direction or control of, that State in carrying out the conduct."[155] In attributing cyber proxy activity to a state, Mačák points to the fact that "from the perspective of international law, the type or even the lack of any domestic legal status of such group is of no consequence for attribution purposes."[156] Thus, **a subordinate relationship between the state and its cyber proxies is central for allowing the application of the three "attribution criteria"—instructions, direction, and control—which all share the same underpinning**.

- Regarding the interpretation of this first criteria, these **"instructions"** must "manifest the will of the State to authorise the unlawful conduct, however broadly they may be phrased."[157] There is also a need to obtain proof: material evidence to trace the offensive cyberoperation back to the instructing state. For instance, outside of a declared armed conflict, public incitements to perpetrate cyberattacks on another country's infrastructure would not be sufficient to be considered as instructions, per Article 8.[158]
- The second criteria of **"direction"** implies that an ongoing, sustained relationship of subordination

[152] See A/76/136, *Official compendium of voluntary national contributions on the subject of how international law applies to the use of information and communications technologies by States*, submitted by participating governmental experts in the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security established pursuant to General Assembly resolution 73/266, 13 July 2021, https://ccdcoe.org/uploads/2018/10/UN_-Official-compendium-of-national-contributions-on-how-international-law-applies-to-use-of-ICT-by-States_A-76-136-EN.pdf.

[153] The Federal Government of Germany, *On the Application of International Law in Cyberspace*, Position Paper to be annexed to the 2021 UN GGE Report on *Advancing Responsible State Behaviour in Cyberspace in the Context of International Security*, March 2021, p. 6, https://www.auswaertiges-amt.de/blob/2446304/32e7b2498e10b74fb17204c54665bdf0/on-the-application-of-international-law-in-cyberspace-data.pdf.

[154] Schmidt, "Expert Backgrounder: NATO Response Options to Potential Russia Cyber Attacks – Understanding the Legal Framework."

[155] See International Law Commission (ILC)'s Articles on State Responsibility, Article 8, p. 47, https://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf.

[156] See Mačák, "Decoding Article 8 of the International Law Commission's Articles on State Responsibility: Attribution of Cyber Operations by Non-State Actors," p. 414.

[157] Idem, p. 416.

[158] See Idem, p. 415-416.

between the state and its cyber proxies can be documented through evidence (for instance, financial transactions for offensive cyberservices that would recur over time).[159]

- The third criteria of **"control"** constitutes a very high standard, even in the analogue world, but is even more complex to meet in modern cyber proxy activities. The test of "effective control" requires that a state plans and controls the modalities, targets, and operational aspects of the overall offensive cyberoperation. Mačák explains how "the State in question must go beyond merely supporting the relevant non-state actor, whether this takes the form of financing, organizing, training, supplying [or] equipping the latter."[160] Based on legal interpretation of Article 8, the longterm patronage and support provided both by Russia and China to their APT groups and cybercriminal proxies would probably not be sufficient to pass the test of effective control.[161]

ATTRIBUTION IN PRACTICE: As described in the Technical Section, current trends in decentralization, automation (autonomous malware and bots), and outsourcing (cybercrime as a service) make it increasingly complex to trace the source of offensive cyber capabilities, obtain material evidence to prove instructions, direction, or control, and therefore attribute state responsibility for wrongful and illicit conduct across jurisdictions.

Two recent examples of alleged nation state cyberattacks illustrate that the direct transfer of expertise and technology between threat actors constitutes another problem in proving a government's involvement and meeting attribution criteria. For instance, according to several cybersecurity experts, the Chinese APT group HAFNIUM may have leaked unpatched Microsoft vulnerabilities to be exploited in a short time period by a half-dozen APT groups in the same region.[162] But while the 2021 attack on Microsoft Exchange Servers resulted in hacks of hundreds of thousands of victims across critical sectors and in 115 countries, it is unlikely that the attack would pass the thresholds of "use of force" or "armed attacks."

In 2021, the U.S. Federal Bureau of Investigation (FBI) confirmed that the Russian criminal APT group, FIN7 (also known as DarkSide), conducted a harmful ransomware attack on the U.S. Colonial Pipeline Company.[163] The threat actor is known for its obfuscation strategies, for example, sharing advanced offensive techniques with other ransomware groups and enabling them to thrive and be more effective. According to cybersecurity experts, FIN7 or DarkSide is tolerated by the Russian government, "a private, for-profit criminal organization that operates under the benign neglect of Russian authorities."[164] The malicious group also falls under what Maurer defines as sanctioning or passive support; that is, when a state cultivates and enables an ecosystem in which cybercriminal groups can generate revenue by specifically hacking the state's geopolitical adversaries. Without attributing the cyberattack to the Russian state, U.S. President Joe Biden asserted that Russia has "some responsibility to deal with this [incident]."[165]

What we increasingly see is the emergence of a loose, nebulous category of cyberactivity that defies the attribution standards of international law: threat actors who operate remotely, or from a given state's territory, conduct offensive cyberoperations which align with this state's intelligence services or another state-aligned group but are not

[159] See Idem, p. 417-418.

[160] Idem, p. 421.

[161] Ibid.

[162] Goodin, "There's a vexing mystery surrounding the 0-day attacks on Exchange servers," *ArsTechnica*, 11 March 2021, https://arstechnica.com/gadgets/2021/03/security-unicorn-exchange-server-0-days-were-exploited-by-6-apts/.%20.

[163] Federal Bureau of Investigation, "FBI Statement on Compromise of Colonial Pipeline Networks," 10 May 2021, https://www.fbi.gov/news/press-releases/fbi-statement-on-compromise-of-colonial-pipeline-networks.

[164] Rivero, N., "Hacking collective DarkSide are state-sanctioned pirate," *Quartz*, 10 May 2021, https://qz.com/2007399/the-darkside-hackers-are-state-sanctioned-pirates.

[165] "Biden Says Russia Has 'Some Responsibility' In Pipeline Ransomware Attack," *Radio Free Europe/Radio Liberty*, 10 May 2021, https://www.rferl.org/a/fbi-confirms-darkside-hacker-group-pipeline-cyberattack-russia/31248174.html.

© iStock/the-lightwriter

coordinated or directly supported by the state in question.

What we can also conclude from the application of Article 8 of the ILC's Articles on State Responsibility is that there is an **increasing grey zone where cyber proxies' strategies and behaviors do not meet the high standards required to ascribe state responsibility. Nation states are certainly exploiting this "unregulated zone" of cyberspace**.

INTERNATIONAL HUMANITARIAN LAW

During times of armed conflict, international humanitarian law (IHL)—including the core principles of distinction, proportionality, and precautions—provides a comprehensive regulatory framework that can be applied to offensive cyberoperations.[166] **IHL binds all parties to an armed conflict and thus establishes an equality of rights and obligations between states and non-state actors**. Yet, for IHL to apply, there must be a clear nexus between harmful cyberoperations and ongoing hostilities. In this regard, it is important to note that **only cyberoperations that are conducted in support of kinetic operations are governed by IHL**.[167]

Under IHL treaties and customary law of armed conflict, certain types of nefarious cyber activities are prohibited including (1) cyber capabilities that qualify as weapons and are by nature indiscriminate, (2) direct attacks against civilians and civilian objects, (3) indiscriminate attacks that do not distinguish

[166] See ICRC, *International humanitarian law and cyber operations during armed conflicts*. See also "International humanitarian law (jus in bello)," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/International_humanitarian_law_(jus_in_bello).

[167] See Gisel, L., Rodenhuser, T. and Dormann, K., "Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts," *International Review of the Red Cross* (2020), 102 (913), p. 287–334, https://international-review.icrc.org/sites/default/files/reviews-pdf/2021-03/twenty-years-ihl-effects-of-cyber-operations-during-armed-conflicts-913.pdf. See in particular p. 297-310.

between military objectives and civilians or civilian objects, (4) disproportionate attacks that may cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, (5) military cyberoperations that would not take all feasible precautions to avoid or at least minimize incidental civilian harm, (6) attacks that would destroy, remove, or render useless objects indispensable to the survival of the population, and (7) attacks that would target humanitarian and medical services, including digital medical records.[168] Only a few other types of cyberoperations (even if they do not meet the definition of "military operation") might be regulated by IHL based on the principle of distinction. For instance, a limited range of psychological cyberoperations—such as severely harmful types of propaganda—could be under the protective reach of IHL if they are directed at civilians and if "they amount to prohibited acts or threats of violence the primary purpose of which is to spread terror among the civilian population or encourage IHL violations."[169]

A few states have shared positions on how IHL applies to cyberoperations during armed conflicts, and a wealth of debates among legal experts exist on the matter. Among the below issues that lack consensus or require further analysis, the first two concern the *types of IHL violations* that could be perpetrated by cyber proxies. The last two are relevant to determining whether offensive operations conducted by cyber proxies constitute *direct participation in hostilities* and if proxy actors can be considered combatants or mercenaries.

The "attack threshold": There is no internationally agreed definition on what constitutes a cyberattack or cyber hostilities within IHL.[170] Important technical questions persist about how to define and qualify—in the context of an armed conflict—technical terms

> "There is no internationally agreed definition on what constitutes a cyberattack or cyberhostilities within international humanitarian law.

such as "attack" when they rely exclusively on cyber means. Yet, it is increasingly recognized that cyberoperations designed to bring physical destruction or death meet the attack threshold. In its 2019 Position Paper, the ICRC underlines the importance of considering "harm due to the foreseeable direct and indirect (or reverberating) effects of an attack, for example, the death of patients in intensive care-units caused by a cyberoperation on an electricity network that results in cutting off a hospital's electricity supply."[171] Consensus among states is still lacking as to whether cyberoperations that would not cause physical damage but would result in disruption and loss of essential services, or in erosion of public trust in critical systems, would qualify as an attack and thus violate IHL. Yet, a 2020 ICRC report explains that "the ICRC has taken the position that an operation designed to disable a computer or a computer network during an armed conflict constitutes an attack as defined in IHL, whether or not the object is disabled through destruction or in any other way."[172] ICRC legal experts insist that "under an overly restrictive understanding, a cyber operation that is directed at making a civilian network (electricity, banking, communications, or other network) dysfunctional, or risks causing this incidentally, might not be covered by essential IHL rules protecting the civilian population and objects."[173] This distinction will be relevant to Case Study 1.

Civilian data as a protected "object": A second salient and unresolved question is whether datasets can be considered as "objects" and whether adversarial cyberoperations targeting essential civilian

[168] See ICRC, *International humanitarian law and cyber operations during armed conflicts*, p. 5-6.

[169] See Idem, p. 5, including reference to Art. 51(2) AP I; Rule 2 ICRC Customary IHL Study. See also ICRC, *Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts*, p. 326.

[170] See ICRC, *International humanitarian law and cyber operations during armed conflicts*, p. 7. See also ICRC, *Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts*, p. 310-315.

[171] ICRC, *Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts*, p. 313.

[172] Ibid.

[173] Idem, p. 314.

datasets for manipulation or destruction would then violate IHL.[174] While IHL affords protection to digital medical records, other large populations' datasets are not explicitly covered by IHL, such as biometric ID, social security, financial, and civil registry data.[175] According to the ICRC, "excluding essential civilian data from the protection afforded by IHL to civilian objects would result in an important protection gap."[176] While the question remains unresolved, a number of states, including Finland, France, Germany, Norway, and Romania, have argued that civilian datasets should be afforded under IHL the same level of protection as civilian objects.[177]

Direct participation in hostilities: While there is no internationally agreed definition of what constitutes "direct participation in hostility," ICRC refers to three criteria to assess if an act meets the threshold.[178] The first concerns a **threshold of harm**. The harmful cyberactivity would either have to "adversely affect the military operations or military capacity of a party to an armed conflict" or "cause death, injury or destruction on persons or objects protected against direct attacks" (cf. the above mentioned "attack threshold").[179] The second relates to a need to prove **direct causation** between the hostile cyberactivity and the resulting harm.[180] For instance, the cyberattack launched by proxies has to be an integral part of a coordinated military operation. The third constitutes a **belligerent nexus**, as indicated by the ICRC. The offensive cyberoperation must be specifically designed to directly cause the required threshold of harm in support of a party to the conflict and to the detriment of another.[181] Under certain circumstances, a cyberoperation that would cause serious harm to a state's infrastructure could be considered as direct participation in hostilities by cyber proxies in the context of an armed conflict. As mentioned by the UN Working Group on the use of mercenaries, "it is a matter of fact and degree whether any particular cyberactivity is likely to cause harm to a party to a conflict, with a sufficient nexus between the act and the armed conflict."[182]

The Cyber Law Toolkit provides enlightening examples of how direct participation in hostilities could be understood in cyberspace.[183] Consider a situation where an APT group acts as proxy for State A, which is opposed to State B in an ongoing armed conflict. The APT group uses offensive cyber methods (such as cyber theft or intrusion) to exfiltrate strategic military information about State B and transmit this information to State A for tactical use on the battlefield. Now, consider another situation where the APT group designs sophisticated influence operations, including spreading wrong information about State A's defense strategies or its concentration of troops at certain locations. Such influence operations directly affect and severely compromise State B's military operations. Both types of offensive cyber conduct may qualify as direct participation in hostilities.

[174] Idem, p. 317. See also "Scenario 12: Cyber operations against computer data," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/Scenario_12:_Cyber_operations_against_computer_dataScenario 12 on Computer DataScenario 12 on Computer Data. See Mačák, K., "Unblurring the lines: military cyberoperations and international law," *Journal of Cyber Policy,* Vol. 6, No. 3 (2021), p. 411-428, in particular p. 421-422, https://ore.exeter.ac.uk/repository/bitstream/handle/10871/128291/Macak%20(2021)%20Unblurring%20the%20lines%20military%20cyber%20operations%20and%20international%20law.pdf?sequence=1.

[175] See ICRC, *International humanitarian law and cyber operations during armed conflicts*, p. 8.

[176] ICRC, *Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts*, p. 319.

[177] Mačák, "Unblurring the lines," p. 421.

[178] Melzer, N., *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law* (ICRC, May 2009), see in particular p. 46-64, https://shop.icrc.org/interpretive-guidance-on-the-notion-of-direct-participation-in-hostilities-under-international-humanitarian-law-pdf-en.html.

[179] Idem, see in particular p. 47.

[180] Idem, p. 51

[181] Idem, p. 58.

[182] A/76/151, p. 14, para. 47.

[183] See "Scenario 1: Legal status of cyber operators during armed conflicts," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/Scenario_18:_Legal_status_of_cyber_operators_during_armed_conflict.

Qualification as mercenaries or mercenary-related actors: Direct participation in hostilities is necessary to meet the criteria for classification as a mercenary under Article 47 of the Additional Protocol I to the Geneva Conventions of 1949.[184] Qualification as mercenaries implies that these types of actors do not benefit from the inherent protections related to the status of combatant. Even more, cyber proxies that would qualify as mercenaries could be prosecuted for directly participating in an armed conflict even if it cannot be proven that their participation in hostility has been contracted by a state or a non-state actor.[185] Common Article 1 of the Geneva Conventions also stipulates that States have to ensure IHL compliance by proxy actors that would be conducting operations on their behalf.[186]

Interpreting what "mercenarism" means in cyberspace and cyberwarfare is a complex and enduring challenge. As previously explained, the application of the traditional definition of mercenaries to cyber offense does not reflect well what is actually happening in cyberspace. Because of the opacity and levels of dissociation characterizing the murky ecosystems in which cyber proxy actors thrive, it might be difficult to prove that they meet the above-mentioned necessary threshold of "direct causation" and "belligerent nexus."

ATTRIBUTION IN PRACTICE: Determining who the "operator" of a cyberattack is, and whether the direction of the operation can be attributed to a state or non-state actor party to the conflict, has severe legal consequences in the context of armed conflict. If the entity responsible for planning and executing a given cyberoperation cannot be attributed, it becomes problematic to prove the nexus between the operation and an armed conflict, and assess whether IHL is even applicable to the operation. The problem resides not only with identifying the source of the cyberattack, but also with qualifying in legal terms (and with material evidence) the nature of the relationship with the state on behalf of which the cyberoperation is conducted. Attribution is often necessary to ascribe responsibility to a state. In that case, the ICRC stipulates that "under international law, a state is responsible for conduct attributable to it, including possible violations of IHL," and this includes: (1) violations committed by its organs, including its armed forces; (2) violations committed by persons or entities it empowered to exercise elements of governmental authority; (3) violations committed by persons or groups acting in fact on its instructions, or under its direction or control; and (4) violations committed by private persons or groups which it acknowledges and adopts at its own conduct."[187]

INTERNATIONAL CRIMINAL LAW

A limited category of harmful cyberoperations can also be regulated under international criminal law, which applies to any natural person who commits an international crime. Under this regime, cyber mercenaries and individuals engaging in cyber proxy groups may be prosecuted for conducting cyber and information operations that would constitute war crimes, crimes against humanity, and genocide. To prove individual responsibility for these international crimes, two elements have to be established: *actus reus* (the physical parts of the crime) and *mens rea* (the intent to commit the crime). The principle of command responsibility (Article 28 of the Rome Statute), established in customary international law, stipulates that military commanders may be held

[184] Article 47, Protocol I Additional to the Geneva Conventions of 1949, https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-47#:~:text=1.,or%20a%20prisoner%20of%20war.&text=(f)%20has%20not%20been%20sent,member%20of%20its%20armed%20forces.

[185] A/76/151, p. 14, para 48.

[186] Hill-Cawthorne, L., "GCIII Commentary: Common Article 1 and State responsibility," ICRC Humanitarian Law & Policy Blog, 28 January 2021, https://blogs.icrc.org/law-and-policy/2021/01/28/gciii-commentary-common-article-1-state-responsibility/.

[187] ICRC, *Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts*, p. 309-310.

criminally responsible for crimes committed by armed forces under their effective command and control.[188] As specified by the UN Working Group on the use of mercenaries, "the superiors of such individuals who are implicated in the commission of devastating cyberattacks, or fail to prevent such malicious cyberattacks, should not evade accountability."[189]

In 2019 and 2020, a Council of Advisers on the Application of the Rome Statute to Cyberwarfare provided critical insights into how the ICC may regulate cyberoperations that have the potential to cause grave suffering of the civilian population, including suffering equal to that caused by the most serious international crimes.[190] For instance, members of the Council confirmed that a cyberoperation altering or deleting **civilian medical data** may be considered a violation of IHL, and therefore possibly a war crime.[191] The Council also specified conditions under which cyberoperations could lead to crimes against humanity: by producing a mass-casualty event related to the **malfunction or failure of critical infrastructures** (such as nuclear power plants, dams, and large networks of hospitals); or by inflicting serious and **systematic harm to the mental health** of a targeted group to the extent that it would amount to torture or persecution.[192] In particular, the Council agreed with the UN's special rapporteur on torture and other cruel, inhuman, or degrading treatment or punishment that "cybertechnology can also be used to inflict, or contribute to, severe mental suffering while avoiding the conduit of the physical body, most notably through intimidation, harassment, surveillance, public shaming and defamation, as well as appropriation, deletion or manipulation of information."[193] Regarding the crime of genocide, members of the Council concluded that cyberoperations may not only contribute to severe psychological and mental harm, but also help initiate and amplify physical acts of violence that could threaten the destruction of a specific minority.[194]

> How to qualify, document, and attribute international crimes in the digital context and how to proceed across jurisdictions will continue to create legal ambiguities and challenges.

ICC proceedings require very high evidentiary standards for attribution, and this is particularly relevant to the involvement of cyber proxies. How to qualify, document, and attribute international crimes in the digital context and how to proceed across jurisdictions will continue to create legal ambiguities and challenges. The characteristics of cyber proxy operations, such as the different ways for non-state actors to hide or falsify their identity and obfuscate information related to their conduct, knowledge, and intent, could complicate evidentiary attribution. In particular, the Council of Advisers "considered that if cyberactivity is outsourced by a party to the conflict, as a factual matter, it may be more difficult to establish that the perpetrators of a particular attacks had knowledge that there was an ongoing armed conflict, particularly if the attack is launched from territory outside of the territory of conflict."[195] Such considerations will be developed in Case Study 1.

[188] Command responsibility is a jurisprudential doctrine in international criminal law permitting the prosecution of military commanders for war crimes perpetrated by their subordinates. The first legal implementations of command responsibility are found in the Hague Conventions IV and X (1907). See Article 28, Rome Statute of the International Criminal Court, p. 15, https://www.icc-cpi.int/sites/default/files/RS-Eng.pdf.

[189] A/76/151, p. 15, para. 52.

[190] The Permanent Mission of Liechtenstein to the United Nations, *The Council of Advisers' Report on the Application of the Rome Statute of the International Criminal Court to Cyberwarfare*, August 2021, https://www.regierung.li/files/medienarchiv/The-Council-of-Advisers-Report-on-the-Application-of-the-Rome-Statute-of-the-International-Criminal-Court-to-Cyberwarfare.pdf.

[191] Idem, p. 9.

[192] Idem, p. 65-67.

[193] Bowcott, O., "UN Warns of Rise of "Cyber torture" to Bypass Physical Ban," *The Guardian*, 21 February 2020, https://www.theguardian.com/law/2020/feb/21/un-rapporteur-warns-of-rise-of-cybertorture-to-bypass-physical-ban.

[194] *The Council of Advisers' Report on the Application of the Rome Statute of the International Criminal Court to Cyberwarfare*, p. 80-88, https://www.regierung.li/files/medienarchiv/The-Council-of-Advisers-Report-on-the-Application-of-the-Rome-Statute-of-the-International-Criminal-Court-to-Cyberwarfare.pdf.

[195] Idem, p. 49.

INTERNATIONAL HUMAN RIGHTS LAW (IHRL)

- **Human rights implications of cyber proxy activity:** Offensive cyberoperations, including cyberattacks, intrusion, and surveillance, can cause human rights violations, both in armed conflicts and in peacetime. The ICRC confirms that cyber-operations "without, or with unclear, links to armed conflicts" have had damaging impact on civilian infrastructures, such as hospitals, water and electrical facilities, and nuclear and petrochemical plants.[196] In 2021, the UN Working Group on the use of mercenaries noted, "Destruction of databases which contain information concerning civilians could quickly bring government services and private businesses to a complete standstill and thus cause more harm to civilians than the destruction of physical objects." In situations where cyberattacks have harmful impact and reverberating effects on critical systems across the medical, energy, water and disaster relief sectors, these attacks may engage and even violate **the right to life and the right not to be subjected to torture and other inhuman or degrading treatment**.

  Aggressive information operations and digital propaganda may affect citizens' **right to information**. Offensive cyberoperations that target and aim to manipulate the functioning of electoral systems may also directly impact citizens' capacity to vote and their fundamental **democratic rights of representation**.

  Some digital vulnerabilities can be exploited for the cybersabotage of IT supply chains and critical infrastructure while others can be harnessed for both **targeted and mass-surveillance of populations**. In its 2019 "Surveillance and human rights" report, the UN Rapporteur on the promotion and protection of the right to freedom of opinion and expression explains that "Governments and private actors are known to purchase security vulnerabilities in commonly available software from security researchers to be utilized as 'zero-day exploits' for the purpose of gaining access to individual communications and devices."[197] The combination of critical vulnerabilities in widely used civilian technologies (e.g., iPhones, Android phones) with very intrusive spyware (such as the ones using zero-click infection method and key-logger) gives offensive actors the capacity to implement precise surveillance but on a large-scale. If the vulnerabilities are not reported to relevant service providers, the **security of increasingly large swaths of populations** can be compromised, including remote, cross-border **access to and exfiltration of their most sensitive financial, health, biometrics, and employment data**. The 2019 "Surveillance and human rights" report describes a private surveillance industry shrouded in secrecy, operating with impunity in an unregulated market, and matching the needs of authoritarian regimes.[198] The report also clearly mentions that "Surveillance of specific individuals – often journalists, activists, opposition figures, critics and others exercising their rights to freedom of expression – has been shown to lead to arbitrary detention, sometimes to torture and possibly to extrajudicial killings."[199]

  Beyond risks to physical safety and security, such new forms of hypercharged surveillance activities can violate individuals' **right to privacy, as well as their freedom of expression and freedom of association**, becoming a powerful medium for repression and censorship. Precision surveillance, increasingly merging with biometrics and behavioral recognition technologies, can also be used to **discriminate against certain groups based on their race, religion, sexual orientation**, or other factors. In certain instances of repression, pervasive surveillance has led to violence and direct harm against vulnerable population sub-

[196] ICRC, *Avoiding Civilian Harm from Military Cyberoperations during Armed Conflict*, p. 6.

[197] Human Rights Council (41st Session), *Surveillance and human rights - Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, p. 6, para. 17.

[198] Idem, p. 3 and p. 6.

[199] Idem, p. 3, para 1.

groups who are targeted. Examples of tech-enabled surveillance used in situations of repression or conflict can be found in Syria, China, and Iran.[200]

- **Attributing responsibility for IHRL violations:** During times of peace, armed conflict, and humanitarian emergencies, international human rights law (IHRL)[201] establishes obligations for states to protect, respect, and fulfil human rights standards. In other words, states bear international responsibility for the violation of human rights obligations that are attributable to them.[202] It follows that states must refrain from conducting, sponsoring, or supporting offensive cyberoperations that would violate the rights of individuals within their jurisdiction. While there is no international consensus on whether human rights obligations apply to a state's extraterritorial cyberactivity, legal experts concur that "human rights obligations do apply to some acts of a State outside its territory."[203] This is relevant for the present report, as there are increasing trends of authoritarian states imposing cyber surveillance and intrusion on victims located in a foreign or third country.

  Recruiting non-state actors to wage cyberattacks on civilian groups does not relieve governments from their obligations under IHRL. States are also bound to take all reasonable measures to protect the rights of individuals within their jurisdiction from violation by cyberoperations when such operations are conducted by other states and non-state actors.[204] This implies that states are required to guarantee human rights compliance by the private sector within their territory through domestic legislation and enforcement. For instance, states must investigate, prosecute, and sanction alleged violations of human rights abuses by cyber proxies, including APT groups and private military and security companies, and provide effective remedies to victims. The obligation to protect also entails implementing preventive measures. Consider, for instance, the following: members of a certain ethnic group living in a defined territory in a State have been the target of pervasive cyber-aggression that interferes with their rights. The State concerned is therefore responsible for taking all reasonable measures to prevent future, similar cyberthreats from occurring.

> "Recruiting non-state actors to wage cyberattacks on civilian groups does not relieve government from their obligations under international human rights law.

In contemporary conflicts, armed non-state violent actors (ANSA) increasingly have access to the services of private sector offensive actors and a globalized market of offensive cyber capabilities. Through resolutions adopted in UN organs such as the Security Council, the General Assembly, or the Human Rights Council, it is increasingly recognized that ANSA do bear human rights obligations, particularly if they exercise either government-like functions or *de facto* control over territory and population. In these situations, ANSA must, at a minimum, respect and protect the human rights of individuals and groups in that territory. States should therefore develop and strengthen clear and formal mechanisms that can help recognize the IHRL obligations of ANSA in cyberspace, including criteria to determine their capacity to hold

[200] For Syria, see Access Now and The UIC John Marshall Law School International Human Rights Clinic (IHRC), supported by Syrian Justice & Accountability Centre and MedGlobal, *DIGITAL DOMINION: How the Syrian regime's mass digital surveillance violates human rights*, March 2021, https://www.accessnow.org/cms/assets/uploads/2021/03/Digital-dominion-Syria-report.pdf. For China, see Feldstein, S., "China's Latest Crackdown in Hong Kong Will Have Global Consequence," Carnegie Endowment for International Peace, July 2020, https://carnegieendowment.org/2020/07/09/china-s-latest-crackdown-in-hong-kong-will-have-global-consequences-pub-82264. For Iran, see Biddle, S. and Hussain, M., "HACKED DOCUMENTS: HOW IRAN CAN TRACK AND CONTROL PROTESTERS' PHONES," *The Intercept*, 28 October 2022, https://theintercept.com/2022/10/28/iran-protests-phone-surveillance/.

[201] See "International Human Rights Law," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/International_human_rights_lawInternational Human Rights Law.

[202] See "International Human Rights Law (Chapter 6)," *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*, ed. Schmitt, M.N. (Cambridge: Cambridge University Press, 2017), https://www.cambridge.org/core/books/tallinn-manual-20-on-the-international-law-applicable-to-cyber-operations/international-human-rights-law/B5CD88C7C704F5D1CE10294578B5BC9A.

[203] See "International Human Rights Law," Cyber Law Toolkit.

[204] Ibid.

human rights obligations.[205] Such normative effort should include collaboration with international and regional judicial authorities to ensure that the potential for human rights violation in cyberconflict is addressed.

While IHRL presents a relatively robust legal framework with monitoring bodies and enforcement mechanisms at the international level, **complex accountability and compliance challenges** remain, such as **(1) clarifying legal attribution (substantiated with evidence) in instances of human rights violations, (2) tracing surveillance and intrusion back to third parties and private sector actors, and (3) dealing with multi-jurisdictional and extraterritorial types of violations**.

In cyberspace, sufficient mechanisms for ensuring corporate responsibility and related compliance are also lacking. The **UN Guiding Principles on Business and Human Rights**, adopted by the Human Rights Council in 2011, confirm that states have the duty to take appropriate measures to prevent, investigate, punish, and redress human rights abuses by third parties. Moreover, the Guiding Principles indicate that states should exercise adequate oversight when carrying out their duty to protect, for instance when they contract with private sector actors for the procurement of services that could impact human rights.[206] The Guiding Principles also encourage business entities to implement a set of policy commitments and processes, such as human rights impact assessment, consultation with affected groups, and grievance mechanisms for affected rights holders. Yet, when it comes to the cybersurveillance industry, the diagnosis is appalling. In a 2019 report, the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression states that "By every measure, the companies would appear to fail to meet even these minimum baselines,"[207] concluding that, in this technological domain, "self-regulation lacks substance."[208]

CUSTOMARY INTERNATIONAL LAW AND THE NORMATIVE ACQUIS

Non-binding in nature, the consensus reports published in 2021 by both UN Cyber Groups (the 6th UN Governmental Group of Experts [GGE] and the 1st Open-Ended Working Group [OEWG] reflect normative convergence, in particular around affirming the "acquis" of 2010-2015. The normative "acquis" is the common term used by UN and government officials to refer to the collective outcomes of the UN's GGEs on responsible state behavior in cyberspace, in particular the three GGEs that agreed on consensus reports in 2010, 2013, and 2015. This fundamental normative framework stipulates, inter alia, (1) the applicability of international law to cyberspace, including the UN Charter and (2) adhesion to 11 non-binding, voluntary norms of responsible state behavior, with the understanding that further norms could be developed and adopted over time. In its 2021 report, the 6th UN GGE reaffirmed that "States must not use proxies to commit internationally wrongful acts," and that "States should seek to ensure that their territories are not used by non-State actors for unlawful use of ICTs."[209] While this recommendation does not constitute a legal obligation for states, it reproves the instances when a state directs (orchestration) or tolerates (passive support) hostile cyber proxy activity. The 2021 UN

[205] A/76/151, p. 20, para. 79.

[206] See Human Rights Council (17th Session), *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, John Ruggie), A/HRC/17/31, 21 March 2011, p. 10, https://www.ohchr.org/sites/default/files/Documents/Issues/Business/A-HRC-17-31_AEV.pdf.

[207] While they do not create legally binding obligations, the UN Guiding Principles on Business and Human Rights encourage businesses to adopt the following practices: develop and implement a human rights policy; conduct human rights due diligence; engage stakeholders to understand and address human rights impacts; report on their due diligence process; and establish grievance mechanisms.

[208] Human Rights Council (41st Session), *Surveillance and human rights - Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, p. 10, para. 32.

[209] See UN General Assembly (76th Session), *Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security*, A/76/135, 14 July 2021, p. 10 and 18, https://front.un-arm.org/wp-content/uploads/2021/08/A_76_135-2104030E-1.pdf.

GGE report goes a step beyond the previous 2015 GGE report, which had framed cyber proxy activity mainly as non-state actors operating under effective control of a state.

Several principles in the acquis are relevant to the protection of critical civilian infrastructure from offensive cyber proxy activity and rely on both a set of negative and positive standards to ensure responsible behavior in cyberspace. In particular, when public critical infrastructure is targeted by a cyberattack, states are expected to provide assistance and, if the cyberthreat is emanating from their sovereign territory, states are expected to help mitigate the malicious cyberactivity in question. One of the norms also encourages responsible reporting of ICT vulnerabilities. Yet, the Technical Section of this report explains how China has been accused of using such normative expectations to stockpile and exploit zero-day vulnerabilities.

Another achievement of both Cyber Groups consists in their convening and consultative function, building processes for states and multi-stakeholder actors to exchange arguments and **clarify legal positions with increased transparency**. Such effort may not only support future political and legal dialogue but also contribute to forming an ***opinio juris*, which could consolidate understanding and implementation of the rules and principles of customary international law**. During the UN Cyber Groups' negotiations, significant tensions emerged in discussions about whether international legal rules and principles should have a binding effect in cyberspace. Existing international law posits that states have jurisdiction over the ICT infrastructure located within their territory, but there is no overarching consensus regarding the principles of sovereignty, due diligence, and possible response options under the form of collective countermeasures.

For instance, the 2021 GGE report recognizes that there is no multinational consensus on whether **sovereignty** is a primary rule of international law or merely a principle with no binding effect [paragraph 71(b)].[210] The UK firmly insists on sovereignty as a non-binding principle while a growing number of states, including China, France, Germany, and several other European countries, argue for a binding status. **Defining the legal scope of cyber sovereignty has critical implications—for example, in distinguishing and qualifying when remote, offensive cyberoperations, including by proxy actors, would constitute a sovereignty breach**. Offensive cyberoperations that would target digital infrastructure and merely result in a loss of functionality or alteration of datasets would not likely constitute a violation of sovereignty. Only France has signaled that a cyberoperation causing harmful effects on its national territory could be considered a sovereignty breach.

The 2021 GGE report did not make further progress on the question of whether it is lawful for states to assume **collective countermeasures**—for instance, by assisting each other in taking countermeasures during conflicts in cyberspace [paragraph 71(e)].[211] Countermeasures constitute "sub-use-of-force actions that alone would be unlawful but for the fact that they are taken in response to an internationally wrongful act of another state and specifically aimed at inducing that state to return to compliance or pay reparations."[212] Again, it may be necessary in the future to delimitate the legal framework that would allow for states to join forces in cyber responses when attacked by state-sponsored cyber proxies. Particularly, it might be useful to clarify the lines between cyber-resilience strategies and more active countermeasures (for instance, cyber network intrusion for "hunt forward" missions that are often framed as "active cyber defense"). Yet, providing such clarity might also have complex implications for strategic ambiguity, compellence, and deterrence in cyberspace.[213]

[210] Idem, p. 17.

[211] Idem, p. 18.

[212] See Corn, G., "International Law's Role in Combating Ransomware," *JustSecurity*, 23 August 2021, https://www.justsecurity.org/77845/international-laws-role-in-combating-ransomware/.

[213] Compellence is a term used in international relations and security studies to describe a military, economic, or diplomatic strategy or action taken by a state or actor to coerce another state or actor to change its behavior or policies by using threats or force. It is important to note that compellence is distinct from deterrence. Deterrence aims to prevent an adversary from taking a specific action in the first place by demonstrating the willingness and capability to respond forcefully. In contrast, compellence seeks to change a target's behavior after it has already taken a particular action or adopted a certain policy.

The most important principle for addressing cyber proxy offensive activities might be **due diligence**. Due diligence is the normative notion that states should be aware of and aim to prevent a situation where hostile cyberoperations would be operated from their territory. Such a normative principle could constitute useful leverage to address problems of extraterritoriality and attribution of responsibility to states. Due diligence would still apply in different complex situations (1) when a state turns a blind eye on offensive cyber proxy activity emanating from its territory, or (2) when cyberthreat actors launch an attack that transits across several countries and jurisdictions.[214] A certain core of UN Member States, including France and Germany, consider due diligence as an important rule of international law that could become customary, but this position is not universal, and for instance, not shared by Israel and the United States. There may be legitimate reasons for nuanced discussions between states and being cautious about "solidifying" the principle of due diligence. On one hand, it could be used by adversarial states as an excuse to impose more and more intrusive monitoring of their own ICT infrastructures, legitimating practices that violate human rights. Yet, due diligence could help strengthen international expectations of responsible behavior, such as when information is shared with a state indicating that its territory is being used for adversarial cyberoperations.

CASE STUDY 1:

PROLIFERATION OF CYBER PROXIES IN THE ONGOING INVASION OF UKRAINE BY THE RUSSIAN FEDERATION

Modern conflicts have increasingly taken on cyber components, and this became more evident when the Russian Federation escalated its war of aggression against Ukraine in February 2022. Several important shifts are currently shaping how the ongoing international armed conflict in Ukraine is fought with potential consequences for the application of IHL and the Rome Statute. Three observations can be made at the onset. **First**, cyberwarfare can become an integral and operational part of an international armed conflict. **Second**, because of the interconnectivity and dual-use nature of cyberspace, offensive cyberoperations in an armed conflict may target and have indiscriminate effects on civilian infrastructures. **Third**, the ongoing conflict between the Russian Federation and Ukraine shows an alarming proliferation of proxy actors involved with different degrees of intensity in cyber and information operations. Together, these three evolving facets of the ongoing conflict have the potential to cause human harm and make difficult the application of IHL and the Rome Statute.

NEXUS TO AN ONGOING INTERNATIONAL ARMED CONFLICT

Cyberspace has clearly emerged as a critical domain of warfare along with land, air, sea, and outer space. Russia's actions in Ukraine have shown how cyberoperations may be conducted in coordination with and in support of wartime kinetic operations. In this specific context, technical investigation may be used to start establishing, with evidentiary support, the existence of a nexus between offensive cyberoperations targeting Ukrainian critical infrastructures and the ongoing armed conflict. For instance, analyses by experts have confirmed that many offensive cyberoperations launched by proxy actors affiliated with Russia's military intelligence have targeted the same categories of Ukrainian systems hit by wartime

[214] See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 130-131.

kinetic operations.[215] These targets include critical systems across public sectors, including communications, banking institutions, electricity grids, water systems, and logistics and transportation infrastructures. As a 2022 Microsoft analysis indicates, "the repeated temporal, sectoral and geographic association of these cyberattacks by Russian military intelligence with corresponding military kinetic attacks indicates a shared set of operational priorities and provides strong circumstantial evidence that the efforts are coordinated."[216] While there are some technical and policy debates about the military effectiveness of Russia's wartime cyberoperations, high-level officials with NATO, the U.S. Cyber Command, and the U.S. National Security Agency have confirmed high intensity in offensive cyberoperations brought to Ukraine and their civilian networks by Russia.[217] An analysis by Nick Beecroft at the Carnegie Endowment for International Peace reports that "Russia launched an intensive campaign of cyberattacks to coincide with the invasion, constituting around 800 attacks against Ukrainian targets up to the end of March [2022]."[218]

HARMFUL IMPACT ON CIVILIAN POPULATIONS AND OBJECTS

Cybersecurity experts across the public and private sectors have started documenting the proliferation of hostile attacks by APT groups, pervasive cyber-espionage activities, and strains of innovative malware that have been deployed in the context of the ongoing conflict in Ukraine.[219] In this context, one cyber proxy actor has been particularly brazen: Sandworm (also called Iridium or Unit 74455) is affiliated with the GRU and was already involved in the 2015/2016 attacks on Ukraine's electric grid and in the widespread 2017 NotPetya attack.

- In the first months of the conflict, waves of offensive cyberoperations targeted communication systems used by Ukraine's military forces, government agencies, and civilian population. On February 24, 2022, a major cyberoperation conducted by the proxy group Sandworm targeted broadband Internet access, disabling modems that communicate with global Internet provider Viasat's KA-SAT satellite network.[220] The cyberattack was planned to coincide with missile strikes throughout Ukraine's territory as Russian troops crossed the border. Widespread reverberating effects extended far beyond military objectives, impacting civilian infrastructure in Ukraine and European countries, including the disruption of emergency health services in France and wind turbines in Germany.
- On April 8, 2022, the same proxy actor, Sandworm, deployed a sophisticated malware designed to manipulate or damage the function of industrial control systems critical to electrical facilities in a region of Ukraine. The cyberoperation was modelled on similar destructive techniques used by Sandworm in the extensive 2015 and 2016 attacks on Ukraine's power grid. The 2022 attack's purpose was to cut off electrical power for 1.5 to 2 million Ukrainians.
- In October 2022, Sandworm used a new attack method—the Prestige ransomware—to paralyze several logistics, aid, and transportation sectors in Ukraine and Poland.

[215] See Bateman, J., *Russia's Wartime Cyber Operations in Ukraine: Military Impacts, Influences, and Implications,* Carnegie Endowment for International Peace, Carnegie Endowment for International Peace Working Paper, December 2022, https://carnegieendowment.org/files/Bateman_Cyber_final.pdf. See Beecroft, N., *Evaluating the International Support to Ukrainian Cyber Defense*, Carnegie Endowment for International Peace, November 2022, https://carnegieendowment.org/2022/11/03/evaluating-international-support-to-ukrainian-cyber-defense-pub-88322. See Watts, C., "Preparing for a Russian cyber offensive against Ukraine this winter," Microsoft On The Issues, 03 December 2022, https://blogs.microsoft.com/on-the-issues/2022/12/03/preparing-russian-cyber-offensive-ukraine/.

[216] See Watts, "Preparing for a Russian cyber offensive against Ukraine this winter."

[217] Freeman, L., "Russian Cyberattacks Need an International Criminal Court Response," Center for European Policy Analysis, 19 July 2022, https://cepa.org/article/russian-cyberattacks-need-an-international-criminal-court-response/.

[218] See Beecroft, *Evaluating the International Support to Ukrainian Cyber Defense*.

[219] See Bateman, Watts, and Freeman. See also Starks, T., "Russian Sandworm Hackers deployed malware in Ukraine and Poland," *The Washington Post*, 11 November 2022, https://www.washingtonpost.com/politics/2022/11/11/russian-sandworm-hackers-deployed-malware-ukraine-poland/.

[220] See "Viasat KA-SAT attack (2022)," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/Viasat_KA-SAT_attack_(2022).


© iStock/EvgeniyShkolenko

- Throughout the first year of the conflict, destructive cyberattacks that involve new malware designed to delete data and destroy computers' key functions targeted close to 50 Ukrainian agencies and enterprises. As indicated by Microsoft, "of the roughly 50 Ukrainian organizations that Russian military operators have hit with destructive wiper malware since February 2022, 55% were critical infrastructure organizations, including in the energy, transportation, water, law enforcement and emergency services, and health care sectors."[221]

EVIDENTIARY PROCESS FOR IHL VIOLATIONS AND WAR CRIMES

The International Criminal Court (ICC) is currently investigating physical war crimes committed in Ukraine during the past annexation of Crimea and in the ongoing conflict. While Ukraine is not a party to the Rome Statute, it has accepted ICC's jurisdiction after Russia's annexation of Crimea. In this context, legal experts have begun framing discussions on potential "cyber war crimes." They are also actively analyzing the type of offensive cyberoperations conducted by Sandworm to assess if these cyber incidents could constitute IHL violations and, even possibly, war crimes.

In March 2022, human rights lawyers and investigators at the Human Rights Center at the University of California Berkeley School of Law submitted a formal request to the Office of the Prosecutor of the ICC in the Hague.[222] The document urges the ICC to consider war crime prosecutions of the cyber proxy actor, Sandworm, for its offensive cyberoperations targeting civilian objects in Ukraine. Strategically, the Berkeley group emphasized the two cyberattacks that caused widespread power blackouts for hundreds of thousands of civilians in the winters of 2015 and 2016. The reason for focusing on those attacks is that they have been documented in detail by technical and law enforcement experts in the grand jury indictment published by the U.S. Department of Justice in October 2020.[223] The indictment clearly attributes the 2015 and 2016 attacks to six Russian

[221] See Watts, "Preparing for a Russian cyber offensive against Ukraine this winter." See also Microsoft Digital Security Unit, *Special Report: Ukraine: An overview of Russia's cyberattack activity in Ukraine*.

[222] Freeman, "Russian Cyberattacks Need an International Criminal Court Response," Center for European Policy Analysis, 19 July 2022, https://cepa.org/article/russian-cyberattacks-need-an-international-criminal-court-response/. See also Greenberg, "The Case for War Crimes Charges Against Russia's Sandworm Hackers," *Wired*, 12 May 2022, https://www.wired.com/story/cyber-war-crimes-sandworm-russia-ukraine/.

[223] U.S. Department of Justice, "Six Russian GRU Officers Charged in Connection with Worldwide Deployment of Destructive Malware and Other Disruptive Actions in Cyberspace," 19 October 2020, https://www.justice.gov/opa/pr/six-russian-gru-officers-charged-connection-worldwide-deployment-destructive-malware-and.

military officers, pointing to the relationship between the proxy group Sandworm and the GRU, and providing thorough evidence about Sandworm's modus operandi.

What the U.S. indictment does not establish is the nexus between the two 2015/2016 attacks and the kinetic fighting in Ukraine that followed the 2014 annexation of Crimea by the Russian Federation. This issue of establishing a nexus with hostilities in Crimea and Eastern Ukraine could be revisited retrospectively by the ICC. Not only did the harmful 2015 and 2016 cyberoperations lead to physical impact (large power outages), they also disproportionately affected civilian populations in Western Ukraine far from the kinetic combat front.[224] Under Article 8 of the Rome Statute, the ICC could pursue war crime charges for Russia's attacks on civilian objects in Ukraine since February 2022. According to Berkeley human rights lawyer Lindsay Freeman, "Pursuing these charges, along with any additional cyberattacks since the start of this year (2022) that meet the requisite threshold, would be an important step towards modernizing international law and accountability mechanisms."[225]

The Berkeley group of human rights lawyers has advanced several arguments to support the importance of considering "cyber war crimes" perpetrated in the ongoing conflict in Ukraine. **First**, an array of offensive cyberoperations in coordination with kinetic strikes in the ongoing conflict has targeted civilian objects (including, according to Ukrainian sources, humanitarian and emergency medical services).[226] **Second**, there is an urgency to collect, classify, and preserve all digital and material evidence that can help support technical investigation and legal attribution. Such fact-finding missions are necessary to substantiate indiscriminate attacks against civilian objects. **Third**, past and ongoing research efforts by ICRC, the CyberPeace Institute, and other groups have already demonstrated the severe human costs of offensive cyberoperations, and these experts have emphasized why stronger accountability mechanisms should be prioritized. In a 2020 report, the ICRC emphasized how "they [cyber incidents] offer a chilling warning about the potential humanitarian impact of military cyberoperations in contemporary and future armed conflicts."[227] Others have warned that cyberoperations designed to corrupt industrial and safety control systems could lead to a humanitarian crisis.[228] **Fourth**, the ongoing conflict is the first instance of integrating cyberwarfare into an armed conflict, but likely not the last. Therefore, it becomes urgent to consider legal and accountability implications. To this end, the 2019-2020 Council of Advisers on the Application of the Rome Statute to Cyberwarfare has already begun delineating the legal challenges to be discussed and clarified to modernize international criminal law.

UNDERSTANDING THE LEGAL THRESHOLDS IN PRACTICE

Several analyses by cybersecurity and legal experts have argued that a range of offensive cyberoperations by proxy actors affiliated with the Russian military have been conducted alongside and in support of kinetic warfare, thereby creating a nexus with the ongoing conflict in Ukraine. However, there is a need to assess whether some of these offensive cyberoperations meet the relevant legal thresholds to qualify as IHL violations and, possibly, war crimes.

IHL violations: First, experts would have to establish whether some of the offensive cyberoperations conducted in the ongoing conflict in Ukraine constitute an "attack" under IHL. As indicated by ICRC, "Concretely, rules such as prohibition on *attacks* against civilians and civilian objects, the prohibition on indiscriminate and disproportionate *attacks*, and the obligation to take all feasible precautions to avoid or at least reduce incidental harm to civilians and damage to civilian objects when carrying out an

[224] Greenberg, "The Case for War Crimes Charges Against Russia's Sandworm Hackers."

[225] Freeman, "Russian Cyberattacks Need an International Criminal Court Response."

[226] Shchyhol, Y., "Vladimir Putin's Ukraine invasion is the world's first full-scale cyberwar," The Atlantic Council, 15 June 2022, https://www.atlanticcouncil.org/blogs/ukrainealert/vladimir-putins-ukraine-invasion-is-the-worlds-first-full-scale-cyberwar/.

[227] ICRC, *Avoiding Civilian Harm from Military Cyberoperations during Armed Conflict*, p. 6.

[228] Caltagirone, S., "Industrial cyber attacks: a humanitarian crisis in the making," ICRC Humanitarian Law & Policy Blog, 03 December 2019, https://blogs.icrc.org/law-and-policy/2019/12/03/industrial-cyber-attacks-crisis/.

attack apply to those operations that qualify as 'attacks' as defined in IHL."[229] It follows that cybersecurity and legal experts would have to provide evidence that some of the cyberoperations targeting Ukrainian civilian infrastructures since the start of 2022 have significantly incapacitated these infrastructures and essential services even if they did not cause physical destruction. Pervasive or significant loss of functionality, as well as damage that would require replacing some of the digital infrastructure involved in essential services, may actually meet the "attack" threshold. Several of the offensive cyberoperations perpetrated by the proxy group, Sandworm, and described above, may have had such destructive impact. For instance, the cyberoperation against the Viasat's KA-SAT satellite network rendered inoperable thousands of broadband modems in Ukraine, including those used by military and other governmental agencies, causing major loss in Internet communications (about 30,000 new modems had to be shipped to customers to bring customers back online).[230] While cyber defense forces were able to thwart Sandworm's cyberoperation targeting industrial control systems of a large Ukrainian energy provider, the purpose of the cyberoperation was to manipulate these safety systems and severely compromise electric power supply for a large part of the population.[231] Among the waves of cyberoperations that targeted data-based systems and services, the Hermetic Wiper malware destroyed the functionality of computer systems of Ukrainian government agencies as well as financial, defense, aviation, IT, and energy service organizations.[232] The data wiper malware also compromised a Ukrainian border control station, critically slowing the processing of refugees fleeing into Romania.

With time, more research could help document the loss of functionality that was caused to digital infrastructures and reverberating effects on critical sectors (documenting impact on emergency, healthcare, and humanitarian services, or water and energy infrastructures). And such crucial cyber forensics evidence could help experts assess whether some of the 2022 offensive cyberoperations qualify as attacks under IHL. Next, cybersecurity and legal experts would then need to assess the extent to which those cyberoperations that constitute "attacks" violate the principles of distinction, proportionality, and precaution. Military professionals have explained how kinetic warfare operations by Russia have amplified and become more indiscriminate as the war has progressed.[233] And, in the case of offensive cyberoperations against critical civilian infrastructures, there might be an argument to make about the need to assess their potential cumulative impact and the subsequent suffering it has imposed on the civilian population.

War crimes: Under Article 8 of the Rome Statute, indiscriminate attacks affecting the civilian population or civilian objects qualify as war crimes. Reports by military and cybersecurity professionals have confirmed that offensive cyberoperations by Russian proxy groups have targeted Ukrainian critical infrastructure. But to be prosecuted as a war crime, experts would have to prove that the principle of distinction has been breached and that such cyberattacks have *indiscriminately* targeted the civilian population or civilian objects. As emphasized by the Council of Advisers on the Application of the Rome Statute to Cyberwarfare, "this [indiscriminate attacks] covers both attacks that are conducted in an indiscriminate manner – in other words, an attack not

[229] ICRC, *Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts*, p. 312.

[230] See "Viasat KA-SAT attack (2022)," Cyber Law Toolkit.

[231] See Bateman, *Russia's Wartime Cyber Operations in Ukraine: Military Impacts, Influences, and Implications*.

[232] See Watts, "Preparing for a Russian cyber offensive against Ukraine this winter." See *Microsoft Digital Defense Report 2022*, p. 41-43. See also "HermeticWiper malware attack (2022)," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/HermeticWiper_malware_attack_(2022).

[233] Shchyhol, "Vladimir Putin's Ukraine invasion is the world's first full-scale cyberwar."

directed specifically at a military objective – and attacks conducted by means and methods of warfare incapable of being directed at specific military objective."[234] As the Council illustrates, "the former can include a discriminate weapon that is used indiscriminately, such as a piece of malware triggered by accessing a website used by both civilian and military actors."[235] The problem in cyberwarfare is that offensive operations tend to target computing infrastructures that are "dual-use," used both for civilian and military functions. The ICC would therefore have to assess whether the 2022 cyberattacks on Ukrainian energy, water, and other critical infrastructures were specifically designed, conducted, and deployed with the goal to target a military objective and achieve a military advantage. One way to address the analysis, for instance, would be for the cyber proxy groups to demonstrate in their defense argument, through a proportionality test, that targeting electric power supplies (even if those were largely serving civilians) could help Russia achieve direct, prominent, or overbearing military advantage over Ukraine's armed forces. Expert analyses have pointed to the fact that offensive cyberoperations by proxy actors affiliated with Russia have disproportionately affected Ukrainian citizens in the ongoing conflict while bringing few military successes.[236] Some have added that targeting the functioning of the power grid in the middle of winter in cold temperatures could be described as a way to harm the civilian population and seriously degrade its conditions and means of survival.[237]

While the facts still need to be documented and backed by evidence, Ukrainian officials have condemned these cyberattacks on their critical infrastructures. Yurii Shchyhol, Head of Ukraine's State Service for Special Communications and Information Protection, wrote for the Atlantic Council:

"Just as the Russian army routinely disregards the rules of war, Russian hackers also appear to have no boundaries regarding legitimate targets for cyber-attacks. Popular targets have included vital non-military infrastructure such as energy and utilities providers. Hospitals and first responders have been subjected to cyber-attacks designed to disrupt the provision of emergency services in the immediate aftermath of airstrikes. As millions of Ukrainian refugees fled the fighting during the first month of the war, hackers attacked humanitarian organizations."[238]

Overall, attributing cyberattacks and ascribing responsibility for IHL violations and war crimes requires a complex evidentiary process. In the ongoing conflict, cybersecurity companies and professionals have lent their cyber forensics capacity to public authorities and have already attributed a range of cyberoperations to proxy actors affiliated with Russia's military and foreign intelligence services. However, a proportion of hostile cyberoperations remain unattributed (and/or attribution lacks evidence). The process is further complicated by the proliferation of proxy actors on both sides that are parties to the conflict.

PROLIFERATION OF CYBER PROXIES AND PARTICIPATION IN HOSTILITIES

The ongoing conflict in Ukraine confirms proliferation of offensive cyber capabilities and proxy groups engaged in operations on behalf of both parties to the conflict.

The Russian government has many APT groups at its disposal in the ongoing conflict, half a dozen of which have been engaged in developing sophisticated malware for cyberattacks and for cyber

[234] *The Council of Advisers' Report on the Application of the Rome Statute of the International Criminal Court to Cyberwarfare*, p. 42.

[235] Ibid.

[236] Bateman, *Russia's Wartime Cyber Operations in Ukraine: Military Impacts, Influences, and Implications*.

[237] See Starks, "The war in Ukraine tests how cyberattacks fit into rules for war crimes," *The Washington Post*, 11 January 2023, https://www.washingtonpost.com/politics/2023/01/11/war-ukraine-tests-how-cyberattacks-fit-into-rules-war-crimes/.

[238] Shchyhol, "Vladimir Putin's Ukraine invasion is the world's first full-scale cyberwar."

[239] See Watts, "Preparing for a Russian cyber offensive against Ukraine this winter." See *Microsoft Digital Defense Report 2022*, p. 41-43. See also Microsoft Digital Security Unit, *Special Report: Ukraine: An overview of Russia's cyberattack activity in Ukraine*.

espionage.[239] While Sandworm has launched destructive cyberoperations against critical infrastructures, APT 28 (also called Strontium and Unit 26165) has exploited unpatched vulnerabilities in Microsoft Exchange Servers to access strategic information at military and government agencies in central Ukraine. APT groups aligned with Russia extended their reach beyond Ukraine, intruding into defense industry-related organizations of NATO allies and government agencies in Eastern Europe Between February and June 2022, Microsoft detected Russian network intrusion efforts against 128 organizations in 42 countries.[240]

On an ad hoc basis, the Russian government also recruits **cybercriminal syndicates to conduct hostile cyberoperations**; about eight of them have been identified by the U.S. Cybersecurity and Infrastructure Security Agency (CISA) as serious threats in the context of the ongoing international conflict.[241] As Sherman explains, those cybercriminal syndicates have a form of tacit agreement with the Kremlin: while they can moonlight for personal gain (incidentally also bringing money to the Russian economy), they must also respond to the government's requests for cyber offensive support.[242] One example is WIZARD SPIDER (also called UNC2727 or Gold Ulrick), known for having deployed the Conti ransomware against hundreds of public and private organizations in the U.S. and worldwide, including healthcare and first responder networks. In February 2022, WIZARD SPIDER pledged support to the Russian government and threatened critical infrastructure organizations in countries perceived to carry out cyberattacks or war against the Russian government and the Russian people.[243] In March 2022, another cybercrime gang, SCULLY SPIDER, launched a series of DDoS attacks against multiple Ukrainian government organizations in support of Russia's military offensive.[244] Groups, such as Killnet and The XakNet Team, operate in the grey zone between cybercrime and hacktivism, pledging to work for the good of Russia, but have launched intrusive attacks in a dozen nations, including against a U.S. airport.[245]

The Russian invasion of Ukraine has also led to a **resurgence of cyber militancy or cyber hacktivism in support of Ukraine**. The international hacktivist collective, Anonymous, has claimed responsibility for a large number of cyberattacks, including on Russia's Ministry of Defense, the energy group Gazprom, and the state television station RT.[246] Part of the same effort, the Polish hacktivist movement Squad303 has designed a tool to get massive access to Russian cell phone numbers (allegedly 20 million messages) and share "real" insights about the war.[247] Another group, the Cyber Partisans, leads a digital resistance against the Belarusian government and have claimed responsibility for a high-profile operation against the Belarusian railway system that reportedly halted Russian ground artillery and troop movement into Ukraine.[248]

The most notable development in this area relates to a "direct participation in hostilities." Responding to the Ukrainian government's call for support, large

[240] See Microsoft Digital Security Unit, *Defending Ukraine: Early Lessons from the Cyber War*, June 2022, p. 2.

[241] Cybersecurity & Infrastructure Security Agency (CISA), "Russian State-Sponsored and Criminal Cyber Threats to Critical Infrastructure," 09 May 2022, https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-110a.

[242] See Sherman, *Untangling the Russian Web: Spies, Proxies, and Spectrums of Russian Cyber Behavior*.

[243] CISA, "Russian State-Sponsored and Criminal Cyber Threats to Critical Infrastructure."

[244] Ibid.

[245] Ibid.

[246] Chirinos, C., "Anonymous hacks into Russian energy companies, exposing over 1 million emails," *Fortune*, 25 April 2022, https://fortune.com/2022/04/25/anonymous-hacks-into-russian-energy-publishes-emails/.

[247] Pancevski, B., "Using a New Cyber Tool, Westerners Have Been Texting Russians About the War in Ukraine," *The Wall Street Journal*, 12 March 2022, https://www.wsj.com/articles/using-a-new-cyber-tool-westerners-have-been-texting-russians-about-the-war-in-ukraine-11647100803.

[248] Pietsch, B., "Hacking group claims control of Belarusian railroads in move to 'disrupt' Russian troops heading near Ukraine," *The Washington Post*, 25 January 2022, https://www.washingtonpost.com/world/2022/01/25/belarus-railway-hacktivist-russia-ukraine-cyberattack/.

numbers of Ukrainian and remote foreign hackers have contributed their cyber defensive and offensive skills to Ukraine's "IT Army."[249] One of the main activities within Ukraine's IT Army is to provide cyber defense for Ukrainian government and military websites, as well as to secure the personal data of Ukrainian soldiers. They also monitor social media and other online platforms for disinformation and propaganda campaigns. In addition to their defensive activities, the Ukrainian IT Army has also been involved in offensive operations against high-level Russian military and government targets, leaking sensitive information and disrupting Russia's military communication. Analysts have outlined how "although varying skill level and lack of coordination will probably make it hard to sustain a 400,000 person IT army in the long term, the smaller groups may be able to achieve more disruptive operations over the course of the war, especially if they develop closer ties with the Ukrainian government."[250] There are signals that both more sensitive targets and more damaging cyber offenses could be envisioned in this decentralized paradigm for cyberwar.

As emphasized by the CyberPeace Institute, "such engagement does not happen in a legal vacuum," and "If individuals answer the call and attack or defend military targets, they could be treated as combatants and through their involvement escalate the conflict."[251] The engagement of cyber experts to conduct military attacks (whether offensive or defensive) has implications for these individuals who may unwittingly lose their legal protection as civilians under IHL and become legitimate targets of both cyber and kinetic attacks. In such a context where cyberwarfare can be outsourced to an array of hackers' groups, including remote actors, there is a significant risk for unexpected collateral damage on dual-use targets. Moreover, these voluntary proxy groups may not have received adequate training in the application of IHL to cyberwarfare, including interpreting the "attack" and "harm" thresholds and proportionality tests. In certain circumstances, they may therefore become responsible for IHL violations or war crimes.

Modern cyberwarfare risks have become increasingly difficult to regulate and control with corrosive implications for civilian harm and related accountability. **First**, the involvement of cyber proxies that transfer skills, malware, and knowledge across proliferating networks (and across borders) complicates the ability to attribute the conduct of war and ascribe responsibility for potential IHL violations. **Second**, such involvement can lead to or can be instrumentalized towards conflict escalation. **Third**, we may increasingly face an erosion of the normative acquis and of current efforts to frame what constitutes states' responsible behavior in cyberspace. What is becoming clearer in the current war is that states are increasingly willing to sponsor and support (openly or tacitly) some form of hostile cyberoperations by non-state actors outside of traditional military agency and framework. Such an evolution of cyberwarfare may increasingly undermine both international law and the normative acquis, in particular the norm against targeting civilian critical infrastructures and the principle of due diligence.

[249] Braw, E., "Ukraine's Digital Fight Goes Global," *Foreign Affairs*, 02 May 2022, https://www.foreignaffairs.com/articles/ukraine/2022-05-02/ukraines-digital-fight-goes-global.

[250] Shore, J., "Don't Underestimate Ukraine's Volunteer Hackers," *Foreign Policy*, 11 April 2022, https://foreignpolicy.com/2022/04/11/russia-cyberwarfare-us-ukraine-volunteer-hackers-it-army/.

[251] Duguin, S., and Samani, R., "Ukraine: Cyber and participation in hostilities," CyberPeace Institute, 07 March 2022, https://cyberpeaceinstitute.org/news/cyber-and-participation-in-hostilities/.

CASE STUDY 2:

SHORT SYNOPSIS ON THE ROLE OF INFLUENCE OPERATIONS AND CYBER SURVEILLANCE IN CONFLICT

While the issue would deserve its own report, information operations have emerged as a powerful threat with the goal to undermine civilian resilience of populations in conflict situations and to manipulate public opinion within and beyond national borders. According to Microsoft, Russian military and state-run media designed information operations to spread the narrative that Ukraine was working to create chemical and biological weapons.[252] Subsequently, Russian officials made unfounded accusations that Ukraine was planning to detonate a "dirty bomb," which would involve radiological material. The topic became viral on Russian state television news with the purpose of legitimizing the launch of a pre-emptive tactical nuclear strike. It is unclear if these information operations directed at civilian populations could be in breach of IHL. As mentioned earlier, a limited range of severely harmful types of psychological cyberoperations could be under the protective reach of IHL if they are directed at civilians and if "they amount to prohibited acts or threats of violence the primary purpose of which is to spread terror among the civilian population or encourage IHL violations."[253]

Another type of harmful cyber-enabled surveillance and intrusion has increasingly targeted individuals in the civilian population during the ongoing conflict in Ukraine (and beyond in NATO-allied countries). For instance, cyber proxy groups have systematically targeted humanitarian and aid organizations. As part of their intrusion and hyper-targeted surveillance techniques carried out through computers and phones, they lured citizens with fake public safety documents (about safe conduct during artillery shelling).[254]

These practices are reminiscent of the Syrian conflict that involved several cyber proxy groups, including most prominently, the Syrian Electronic Army (SEA), that acted in support of the government and President Bashar al-Assad. A 2021 report by cybersecurity and legal experts exposes how, "in conjunction with actively monitoring their own citizens, the Syrian regime, together with third party groups, is hacking websites and individuals critical of the regime."[255] The report continues, "Through "phishing" operations, social engineering, malware downloads, and gaining access to passwords and networks through security force intimidation, the Syrian Electronic Army (SEA) and the Assad regime have used these practices to monitor and track down activists and human rights defenders in Syria, who are then tortured and killed."[256]

In 2013, SEA members extracted from a standard messaging application the personal information (phone numbers, email addresses, and contact details) of millions of people and leaked the datasets to the Syrian government.[257] Other attacks targeted at social media platforms and messaging applications led to further breaches of civilians' sensitive

[252] See Smith, B., "Defending Ukraine: Early Lessons from the Cyber War," Microsoft On The Issues, 22 June 2022, p. 12-22.

[253] ICRC, *Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts*, p. 326.

[254] Hardcastle, J. L., "US Cyber Command spots another 20 malware strains targeting Ukraine," *The Register*, 21 July 2022, https://www.theregister.com/2022/07/21/us_cyber_command_malware_ukraine/.

[255] Access Now et al., *Digital Dominion: How the Syrian Regime's Mass Digital Surveillance Violates Human Rights*."

[256] Idem, p. 1.

[257] Kastrenakes, J., "Syrian Electronic Army Alleges Stealing "Millions" of Phone Numbers from Chat App Tango," *VERGE*, 22 July 2013, https://www.theverge.com/2013/7/22/4545838/sea-giving-hacked-tango-database-government.

data, including people's birthdays, personal serial numbers, ID cards, CVs, and blood types.[258] SEA's arsenal of cybertechniques is sophisticated and has been augmented by a transfer of tools and know-how from corporate firms.[259] Investigation has shown how the SEA-designed SilverHawk, a malware that disguises itself into fake versions of Microsoft Word and YouTube, as well as fake updates of WhatsApp and Telegram, to hack personal devices.[260] Similar malware injection into Facebook and messaging applications were used to capture webcam activity, monitor keystrokes, and steal passwords.[261]

A recent report claimed that "[t]he monitoring and hacking of devices are suspected to have informed kinetic operations that have cost the lives of many and undermined the crucial work being done by doctors and human rights defenders."[262] The deceptive tactics used by the SEA include social engineering and impersonation to manipulate anti-Assad activists into revealing identities of dissidents and meeting locations. In the wake of such pervasive surveillance practices, surgeons and doctors have been advised not to provide medical mentorship over the Internet to colleagues in Syria for fear of revealing the location of sheltered and underground hospitals.[263]

Both types of cyber-enabled targeting—precise surveillance and psychological operations that target the civilian population in an armed conflict—remain difficult to qualify under IHL and the Rome Statute. They may not meet any threshold of harm even if the Council of Advisers has increasingly recognized the systematic mental harm and suffering that can be imposed on civilians by digital, cyber-enabled means.

[258] Railton, J. S., Regalado, D., and Villeneuve, N., *Behind the Syrian Conflict's Digital Frontlines*, FireEye, February 2015, p. 7-9, https://cryptome.org/2015/02/fireeye-syria-hack.pdf.

[259] Access Now et al., *Digital Dominion: How the Syrian Regime's Mass Digital Surveillance Violates Human Rights*, p. 23. See also *Freedom on the Net: Syria 2019*, Freedom House.

[260] Brewster, T., "Syrian Electronic Army Hackers Are Targeting Android Phones with Fake WhatsApp Attacks," *Forbes*, 05 December 2018, https://www.forbes.com/sites/thomasbrewster/2018/12/05/syrian-electronic-army-hackers-are-targeting-android-phones-with-fake-whatsapp-attacks/.

[261] See Railton et al., *Behind the Syrian Conflict's Digital Frontlines*, Note 248.

[262] Access Now et al., *Digital Dominion: How the Syrian Regime's Mass Digital Surveillance Violates Human Rights,* p. 21.

[263] Baraniuk, C., "Surgeon David Nott: Hack Led to Syria Air Strike," *BBC*, 21 March 2018, https://www.bbc.com/news/technology-43486131.

CASE STUDY 3:

USE OF CYBER PROXIES IN GREY ZONE RANSOMWARE OPERATIONS ON CRITICAL CIVILIAN INFRASTRUCTURES

Ransomware operations have ravaged the public sector and businesses around the world over the last five years. While the war of aggression in Ukraine shows that ransomware has integrated into cyberwarfare methods, the most harmful ransomware operations have not taken place during armed conflicts. They have occurred at times of advanced global competition, in the grey zone between peace and war.

According to industry reports, the global cost of ransomware operations reached 20 billion USD in 2021 alone and is expected to reach 265 billion USD annually by 2031, with a new attack every two seconds as ransomware perpetrators progressively refine their malware payloads and related extortion activities.[264] The *Microsoft Digital Defense Report 2022* qualifies ransomware attacks as a nation-level threat that has turned more sophisticated and impactful as "the adoption of a double extortion monetization strategy has become a standard practice."[265] The report continues, "This involves exfiltrating data from compromised devices, encrypting the data on the devices, and then posting or threatening to post the stolen data publicly to pressure victims into paying a ransom."[266] As explained in the Technical Section of this publication, under the progressive industrialization of cybercrime, different stages in ransomware operations have become available as services with separate entities that build malware, gain access to victims, deploy ransomware, and handle extortion negotiations. This evolution has complex consequences for attributing and prosecuting cybercrimes, and even more, for ascribing even partial responsibility of cyber proxies' criminal behavior to a state.

Another consequence is that adversarial ransomware techniques and practices have moved from cybercrime ecosystems to state-sponsored APT groups and have progressively integrated into nation state-level competition.

The role of cyber proxies has therefore been instrumental in merging adversarial and rogue criminal behaviors in new forms of "covert" or grey zone, low-intensity cyberwarfare that are regulated by few rules. What emerges are two powerful consequences: first, escalation in ransomware attacks with increasingly weakened and destabilized victim states (for instance, leading to national emergencies); and second, more severe civilian harm as populations' datasets are monetized and weaponized for further offenses.

ESCALATION

Some of the most notable ransomware operations have had major impact on governmental infrastructures to the point of inflicting severe paralysis on essential public services. In 2021, the Colonial Pipeline company was victim of a ransomware attack by DarkSide (also known as Gold Waterfall), a cybercrime Russian-speaking gang which operates ransomware-as-a-service and is plausibly tolerated by the Russian government.[267] Colonial Pipeline had to impose a precautionary shutdown of the entire pipeline, which resulted in temporary gas shortages across the U.S. and a state of emergency in Georgia, Louisiana, North Carolina, and Virginia. In May 2022, after weeks of suffering major ransomware attacks, the Costa Rican government was forced to declare

[264] Braue, D., "Global Ransomware Damage Costs Predicted To Exceed $265 Billion By 2031," *CyberCrime Magazine*, 02 June 2022, https://cybersecurityventures.com/global-ransomware-damage-costs-predicted-to-reach-250-billion-usd-by-2031/.

[265] *Microsoft Digital Defense 2022 Report*, p. 10.

[266] Ibid.

[267] See "Colonial Pipeline Ransomware Attack (2021)," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/Colonial_Pipeline_ransomware_attack_(2021).

a national emergency after hospitals were paralyzed and custom and tax collection disrupted.[268] This succession of ransomware-as-a-service operations was perpetrated by Conti, a Russian-based cybercrime collective which allegedly was determined to overthrow the government by means of a cyberattack and, in its communication, urged Costa Ricans to pressure their government to pay a $20 million ransom.[269] In July 2022, Albania's government agencies were impacted by a combination of ransomware and destructive malware that threatened to paralyze critical services, including airports and custom/border authorities.[270] The attack was attributed to an Iranian APT group referred to as "Homeland Justice," but up to four cyber proxy groups might have been involved.[271] In September 2022, after Albania publicly attributed the July attacks to Iran, the APT groups launched a new wave of assaults against the Albanian government, using similar malware and adversarial methods.[272] Retaliation and partial destruction of services, monetization of government and civilian datasets in illicit markets, and threats to overthrow public authorities are now also part of the harm caused by ransomware operations. Another worrisome trend is the indication that ransomware gangs are increasingly going after "softer targets"—countries with less cyber defense capacity, for instance in Latin America—posing new risks for growing economies in the Global South.[273]

CIVILIAN CRISIS IN THE MAKING

Ransomware operations directly target the most vulnerable and critical sectors of nations' civilian infrastructures, from hospitals, emergency networks, water treatment systems, airports, and banks to tax agencies, schools, and pipelines. In 2019, 764 American health care organizations were paralyzed by ransomware attacks; emergency patients were turned away from hospitals, medical records were encrypted or destroyed, surgical procedures and tests were postponed, and first-responder services were interrupted.[274] In a 2022 testimony to U.S. Congress, FBI Director Chris Wray explained how U.S. cyber defense experts helped prevent a ransomware attack on Boston Children's hospital that was attempted in August 2021 by cyber proxies acting for the Iranian government.[275] In September 2022, the second largest public school district in the U.S. was targeted and eventually saw 300,000 of its files dumped online as punishment for denying the attacker's demands.[276] Meanwhile, a series of hospitals in Europe, including in the United Kingdom, France, Czech Republic, and Germany, have become targets. The September 2020 ransomware attack on the University Hospital of Düsseldorf may be responsible for the death of a patient who could not be admitted for emergency care.[277]

[268] Reed, J., "Costa Rica State of Emergency Declared After Ransomware Attacks," *Security Intelligence*, 16 November 2022, https://securityintelligence.com/news/costa-rica-state-emergency-ransomware/; Murillo, A., "Latest cyberattack in Costa Rica targets hospital system," *Reuters*, 31 May 2022, https://www.reuters.com/world/americas/latest-cyberattack-costa-rica-targets-hospital-system-2022-05-31/.

[269] Reed, "Costa Rica State of Emergency Declared After Ransomware Attacks."

[270] Greig, "CISA: Iranian hackers spent 14 months in Albanian gov't network before launching ransomware," *The Record*, 21 September 2022, https://therecord.media/cisa-iranian-hackers-spent-14-months-in-albanian-govt-network-before-launching-ransomware/.

[271] Arghire, "CISA, FBI Detail Iranian Cyberattacks Targeting Albanian Government," *Security Week*, 22 September 2022, https://www.securityweek.com/iranian-hackers-breached-albanian-government-one-year-disruptive-attacks/.

[272] See "Homeland Justice against the Albanian Government," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/Homeland_Justice_operations_against_Albania_(2022).

[273] See *Microsoft Digital Defense Report 2022*, p. 17.

[274] Bajak, F., "Suspected ransomware attack hobbles major hospital chains U.S. facilities," *PBS News Hour*, 29 September 2020, https://www.pbs.org/newshour/nation/suspected-ransomware-attack-hobbles-major-hospital-chains-u-s-facilities.

[275] Raymond, N., "Iranian-backed hackers targeted Boston Children's Hospital, FBI chief says," *Reuters*, 01 June 2022, https://www.reuters.com/world/us/iranian-backed-hackers-targeted-boston-childrens-hospital-fbi-chief-says-2022-06-01/.

[276] Goodin, D., "Big data trove dumped after LA Unified School District says no to ransomware crooks," *Ars Technica*, 03 October 2022, https://arstechnica.com/information-technology/2022/10/ransomware-crooks-dump-big-data-trove-stolen-from-la-school-district/.

[277] Eddy, M. and Perlroth, "Cyber Attack Suspected in German Woman's Death," *The New York Times*, 18 September 2020, https://www.nytimes.com/2020/09/18/world/europe/cyber-attack-germany-ransomeware-death.html.

The human cost of ransomware attacks has become increasingly evident as attacks on the healthcare sector have grown. A 2021 ransomware attack on Ireland's healthcare services agency led to a disruption in patient services for months, including the cancellation of cancer treatment and maternity appointments and of COVID-19 vaccinations.[278] In 2022, an attack on CommonSpirit Health, one of the largest non-profit health systems in the U.S., compromised the personal data of over 600,000 patients, including electronic medical records, which allegedly caused one child to be accidentally given five times the amount of medication needed.[279] In December 2022, a hospital in the Paris region was forced to transfer neonatal and intensive care patients to other facilities after its phone and computer systems were encrypted.[280] A 2021 investigation jointly conducted by Proofpoint and the Ponemon Institute surveyed approximately 600 health care facilities, a quarter of which reported increased mortality rates following a ransomware attack.[281] Research by the CyberPeace Institute indicates that an interruption of healthcare services caused by ransomware attacks can last from two weeks to several months.[282]

Beyond immediate threats to civilians' health, there are longer term security and legal implications of ransomware attacks on the healthcare sector. In 2022, Australia's largest health insurer, Medibank, fell victim to a similar ransomware assault with 9.7 million customers' data stolen, and credit card and medical history information soon after sold on the dark web.[283] The attack was allegedly perpetrated by the Russian-based ransomware gang, REvil, which is known for its large-scale 2021 attacks on the major international meat supplier JBS Foods and the software provider Kaseya VSA. Ransomware attacks perpetrated through aggressive and deceptive techniques by cyber criminals and proxies not only produce severe civilian harm and suffering, they also put executives and authorities in the position of risking long-term harm when they refuse to pay ransom and, as a result, a trove of people's data is sold in underground markets.

INTERNATIONAL LEGAL RESPONSES TO RANSOMWARE

There are several reasons why ransomware attacks targeted at critical civilian infrastructures have prompted a number of nation states to reconsider how best to respond to such destructive attacks. **First**, it is a rising global threat, described by the *Oxford Statement on International Law Protections in Cyberspace: The Regulation of Ransomware Operations and Activities (Oxford Statement)* as "having been employed at an escalating pace by a growing number of malicious actors, including states and non-state groups for financial or political purposes, often connected to criminal and other unlawful activities such as terrorism, [...] and the proliferation of

[278] Perlroth and Satariano, A., "Irish Hospitals Are Latest to Be Hit by Ransomware Attacks," *The New York Times*, 20 May 2021, https://www.nytimes.com/2021/05/20/technology/ransomware-attack-ireland-hospitals.html.

[279] Burky, A., "CommonSpirit Health revealed that cybersecurity breach was much larger than initially thought, reaching 100 facilities in 13 states," *Fierce Healthcare*, 10 April 2023, https://www.fiercehealthcare.com/health-tech/commonspirit-health-reported-it-security-incident-affecting-facilities-wash-neb-and; McGee, M. K., "CommonSpirit's Ransomware Incident Taking Toll on Patients," *Bank Info Security*, 13 October 2022, https://www.bankinfosecurity.com/commonspirits-ransomware-incident-taking-toll-on-patients-a-20259.

[280] Black, D., "Hospital cyberattack forces transfer of critically ill patients," *cybernews*, 06 December 2022, https://cybernews.com/news/hospital-cyberattack/ ; "French hospital suspends operations after cyber attacks," *France24*, 12 May 2022, https://www.france24.com/en/france/20221205-french-hospital-suspends-operations-after-cyber-attacks.

[281] Ponemon Institute, *Cyber Insecurity in Healthcare: The Cost and Impact on Patient Safety and Care*, Proofpoint, 2022, https://www.proofpoint.com/sites/default/files/threat-reports/pfpt-us-tr-cyber-insecurity-healthcare-ponemon-report.pdf.

[282] CyberPeace Institute, *Playing with Lives : Cyberattacks on Healthcare are Attacks on People, March 2021*, https://cyberpeaceinstitute.org/report/2021-03-CyberPeaceInstitute-SAR001-Healthcare-ExecSummary.pdf.

[283] Martin, A., "Medibank says hackers had access to 'all personal data' belonging to all customers," *The Record*, 25 October 2022, https://therecord.media/medibank-says-hackers-had-access-to-all-personal-data-belonging-to-all-customers/; Martin, "Medibank says it will not pay ransom in hack that impacted 9.7 million customers," *The Record*, 06 November 2022, https://therecord.media/medibank-says-it-will-not-pay-ransom-in-hack-that-impacted-9-7-million-customers/; Taylor, J., "Medibank hacker says ransom demand was US$10m as purported abortion health records posted," *The Guardian*, 09 November 2022, https://www.theguardian.com/australia-news/2022/nov/10/medibank-hacker-says-ransom-demand-was-us10m-as-purported-abortion-health-records-posted.

weapons of mass destruction."[284] **Second**, as explained above and confirmed by the *Oxford Statement*, the increasingly varied and sophisticated forms of ransomware operations have led—in the majority of cases where these attack methods have been used—to significant and widespread harm to public and private institutions, and "the significant disruption of critical infrastructure, including healthcare and education, while posing an imminent risk of destructive harm to industrial control systems such as electrical grids, water distribution systems and nuclear power plants."[285]

The types of soft targets in the healthcare, humanitarian, transport, and infrastructure sectors impacted by offensive ransomware operations would mostly qualify as civilian targets (protected personnel and protected objects) under IHL. The *Oxford Statement* specifies the conditions under which ransomware operations could lead to IHL violations and war crimes. Still, for those hostile operations to qualify as IHL violations, they would need to meet the "attack threshold." For instance, during the ongoing conflict in Ukraine, the Prestige ransomware attack perpetrated by Sandworm against logistics, transport, and aid sectors was partially thwarted by Ukrainian and allies' cybersecurity units and therefore may not have passed the harm threshold. But under IHL, provisions focused on civilian populations and objects exist. Paradoxically, legal interpretation is less clear when it comes to grey zone operations conducted between war and peacetime.

CUSTOMARY INTERNATIONAL LAW AND NORMATIVE ACQUIS

Outside of a situation of armed conflict, there is not one precise, binding legal provision that would prohibit ransomware attacks against civilian sectors and comprehensively protect these infrastructures.[286] Most rules and principles of customary international law apply to state conduct and pertain to situations when legal responsibility for committing an international wrongful act can be ascribed to a state. This is the case, for instance, when considering the prohibition of the use of force, the prohibition on intervention, the obligation to respect the sovereignty of other states, and the use of countermeasures. Customary international law therefore provides limited leverage in practice to address attacks by a state's cyber proxies against the civilian infrastructure of another state.

It is a question of fact and degree whether an offensive ransomware operation would meet the **"use of force"** threshold and qualify as an internationally wrongful act. As specified by ICRC experts, "there is consensus among academic commentators that a State-sponsored cyber operation directly resulting in the killing of persons abroad would be covered by this prohibition (see e.g., the *Tallinn Manual 2.0*, p. 333 and some states, like Australia and Estonia, have expressed the view that such cyber operations could amount to a use of force)."[287] Under this interpretation, a ransomware attack would have to lead, for instance, to malfunctions of medical equipment and operations so serious that they can be determined, at the time of investigation, as the cause of a patient's death. Given these complexities, most ransomware operations might not clearly meet the threshold. Moreover, as mentioned above, if the ransomware attack is perpetrated by cyber proxies, with different operations outsourced to criminal gangs, further obstacles remain in tracing the relationship with the state sponsoring or passively orchestrating the attack and ascribing legal responsibility.

Under the principle of **non-intervention**, intervention in the internal affairs of other states with the

[284] The Oxford Process, *The Oxford Statement on International Law Protections in Cyberspace: The Regulation of Ransomware Operations.*

[285] Ibid.

[286] The Cyber Law Toolkit has provided an extensive expert-based scenario that explores how ransomware operations that are taking place outside of an armed conflict may be classified under international law. See "Scenario 14: Ransomware campaign," Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/Scenario_14:_Ransomware_campaign.

[287] Mačák, Rodenhäuser, T., and Gisel, L., "Cyber attacks against hospitals and the COVID-19 pandemic: How strong are international law protections?" ICRC Humanitarian Law & Policy Blog, 02 April 2020.

aim of coercion is prohibited. As specified by ICRC experts, "pursuant to the element of coercion, the act in question is prohibited only when designed to compel a targeted State to change its conduct with respect to a matter on which it may otherwise decide freely (see the International Court of Justice's Nicaragua judgment, para. 205 and the *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*, p. 317)."[289] The large-scale offensive ransomware operations that forced the Costa Rican government to declare a national emergency and paralyzed hospitals and other governmental functions could come close to unlawful intervention—or at least have had an impact—in Costa Rica's internal affairs. Yet, the element of coercion is absent and the attack was perpetrated not by a state, but by the Russian-based cybercrime collective Conti. While expert-based analysis has identified Conti as a plausible proxy of the Russian state, passing the test of "effective control" and ascribing legal responsibility for state conduct may remain a problem.

Ransomware operations that paralyze essential services (border control, emergency and health services, tax agencies) in another state's territory may constitute a violation of state **sovereignty**. Beyond legal attribution of cyber proxies' unlawful activities, another challenge would be the lack of international consensus on whether there is a standalone international legal obligation to respect the sovereignty of other states in cyberspace.

The impact of ransomware attacks has led experts to consider the legality of different strategic responses to those cyberthreats. Pursuant to Article 51 of the United Nations Charter and customary international law, only a hostile cyberoperation that would cause destructive, physical harm to the extent that it constitutes an armed attack may legitimate the use of force lawfully in **self-defense**. In the aftermath of the terrorist attacks of 9/11, NATO, the UN Security Council, and individual states have considered that the law of self-defense could apply to attacks that reach the harm threshold, even if perpetrated by non-state actors.[290] According to international law experts, it remains unlikely that ransomware operations, while still causing severe harm and disruptions, would qualify as an armed attack.

Customary international law also imposes limits on the use of **countermeasures**. To engage in countermeasures, there must be an international wrongful act, and the object of countermeasures must be a state. Both conditions are unlikely to be met in most major ransomware attacks encountered previously, even those that were passively orchestrated by a state. The Cyber Law Toolkit has provided some guidance through a scenario in which "a state-sponsored ransomware campaign severely interferes with the functioning of various public institutions in another State."[291] While the impacted state and its allies could issue collective public attribution and sanctions, the expert-led scenario concludes that other types of **collective countermeasures** with a more offensive character would not be allowed under international law. Incidentally, it is useful to note that countermeasures are not anticipatory; they would intervene only after the hostile acts—as a reaction to the ransomware attack—and would therefore offer limited means of prevention. A second issue is the one of collective countermeasures, which lacks consensus among the expert community and states.[292] At the center of division is whether states, that have not been injured, may engage in coordinated responses to offensive cyberoperations in support of an impacted state. The collective approach is strategic for two reasons: the ransomware threat is global, yet few states (except leading tech-nations) would have the cyber defense skills necessary to mount proportionate, tailored

[288] See "Case Concerning Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America)," Judgment of 27 June 1986, *International Court of Justice Reports of Judgments, Advisory Opinions and Orders*, p. 107-108, para 205, https://www.icj-cij.org/sites/default/files/case-related/70/070-19860627-JUD-01-00-EN.pdf.

[289] Mačák, Rodenhäuser, T., and Gisel, L., "Cyber attacks against hospitals and the COVID-19 pandemic: How strong are international law protections?", ICRC Humanitarian Law & Policy Blog, 02 April 2020.

[290] See Arimatsu, L. and Schmitt, M. N., "The plea of necessity: an oft overlooked response option to hostile cyber operations," *International Law Studies* Vol. 97 (2021), p. 1177.

[291] "Scenario 14: Ransomware campaign," Cyber Law Toolkit.

[292] See Corn.

© iStock/cofotoisme

countermeasures. The cybersecurity divide is crucial for countries in the Global South.

Two other principles, the **plea of necessity** and the **duty of due diligence**, may be relevant and would not require harmful ransomware operations to be attributable to a state. In certain circumstances, the **plea of necessity** allows a state in a situation of grave and imminent peril to take action against non-state cyberthreat actors if the action in question is the sole means of safeguarding that interest.[293] In this instance, there would be a need to demonstrate that the state's essential interest is under "grave and imminent peril." Would the series of hostile ransomware operations against the Costa Rican government (including its healthcare and tax agencies) qualify this notion? Even if it does, it leaves an incapacitated state in the position of engaging in appropriate cyber defense, regardless of its ability and capacity to do so.

While the customary rule of **due diligence** remains controversial, it would help circumvent the attribution problem. The *Tallin Manual* specifies that states should take all measures that are feasible in the circumstances to put an end to cyberoperations that affect the right of, and produce serious adverse consequences for, other states.[294] The test in this case is to prove that the state has actual knowledge, or should have known, that its territory is being used for offensive ransomware operations that target another state's critical infrastructure and produce severe transboundary harm. The **normative acquis** reinforces the notion of due diligence. In a nutshell, states should not support, tolerate, or harbor cyber proxies that are known to have contravened responsible behavior in cyberspace, nor should they allow infrastructure residing within their sovereign territory to be utilized by cyber proxies for such acts. However, these rules provide only positive guidance but no binding obligations to cope with the scourge of ransomware attacks.

Arimatsu and Schmitt explain how "in certain limited circumstances, the due diligence rule may open the door to countermeasures taking the form of action against non-state actors whose hostile cyber operations are not attributable to a state on the basis that the territorial state has breached its obligation

[293] See Arimatsu and Schmitt.

[294] See Schmitt, "Three International Law Rules for Responding Effectively to Hostile Cyber Operations," *Just Security*, 13 July 2021.

of due diligence."[295] Other cyber defense experts have contested such "security benefits" in the form of "ready access to countermeasures," underlining how complex it would be to prove a breach of due diligence. They also point to other negative effects: "It would further incentivize bad actors to forum shop the situs of their infrastructure and operations to states actually or ostensibly lacking the capability or capacity to effectively disrupt them while opening law abiding states like the United States to another line of lawfare attack."[296]

In 2021, diplomatic discussions between the U.S. and Russia focused on offensive ransomware operations. President Joseph Biden gave President Vladimir Putin a list of 16 critical infrastructure entities that Russia could not attack without consequences, and warned him against allowing further malicious cyber activities against the United States. According to White House documents, President Biden warned that the U.S. would "take any necessary action to defend its people and its critical infrastructure in the face of this continuing challenge."[297] Such events have prompted cyber defense expert to raise interesting questions on the evolving understanding of customary international principles:

'This leaves the issue of the United States' obligations under international law and its views on how they would apply to the cyber operations implicated in Biden's threat. What can we draw from his apparent decision to embrace, at least as an option, cyber operations targeting for disruption the overseas infrastructure of non-state criminal organizations? Does it signal an evolving U.S. position on the unsettled debate over the purported rule of cyber due diligence? That is, does the United States now consider Russia (and presumably other states in whose territory ransomware infrastructure resides) legally accountable, directly or indirectly, for the actions of non-state criminal organizations? Or does it reflect an assessment, consistent with at least the United Kingdom's view (and that of the U.S. Department of Defense), that the principle of sovereignty does not present a legal barrier to conducting certain counter-ransomware operations?"[298]

INTERNATIONAL HUMAN RIGHTS LAW

Another important question is whether offensive ransomware campaigns against essential services or critical sectors of another state could violate IHRL. The type of destructive ransomware attacks witnessed in the last years have not only paralyzed essential computer functions, but also leaked people's sensitive and personal information to malicious actors, likely constituting interference and even violations of fundamental rights[299] to life, health, security, and privacy. The problem of multi-jurisdictional and extraterritorial types of human rights violations has also become increasingly complex. Offensive ransomware campaigns can be launched and deployed by an array of proxy and criminal actors operating from different hot spots. For instance, the 2022 Medibank attack deployed by the ransomware gang REvil potentially compromised the privacy and security of millions of Australians. Yet, there is a need to trace back the different sequences of ransomware campaigns across jurisdictions and confront the Russian government with breach of human rights that took place extraterritorially. The same evidentiary process is necessary when tracing back Conti's ransomware attacks that interfered and may have violated the right to life and the right to health of the Costa Rican population.

While there is no international consensus on whether human rights obligations apply to a state's extraterritorial cyberactivity, the UN Human Rights Committee has clarified that a state's obligations to respect and ensure the right to life extend to "per-

[295] See Arimatsu and Schmitt, p. 1180.

[296] See Corn, "International Law's Role in Combating Ransomware."

[297] Chalfant, M. and Miller, M., "Biden warns Putin on Russian ransomware attacks," *The Hill*, 09 July 2021, https://thehill.com/homenews/administration/562282-biden-presses-putin-to-take-action-on-russian-ransomware-attacks/.

[298] See Corn, "International Law's Role in Combating Ransomware."

[299] The right to health enshrined in Article 12 of the International Covenant on Economic, Social and Cultural Rights (ICESCR), or the right to life enshrined in Article 6 of the International Covenant on Civil and Political Rights (ICCPR).

sons located outside any territory effectively controlled by the State, whose right to life is nonetheless impacted by its military or other activities in a direct and reasonably foreseeable manner."[300] Importantly, the UN Committee on Economic, Social, and Cultural Rights has argued that "States parties have to respect the enjoyment of the right to health in other countries."[301] Overall, legal ambiguities, as well as challenges in determining compliance and accountability, continue to pose obstacles in the application of IHRL to harmful cyberoperations, such as ransomware campaigns.

Oxford Statement on International Law Protections in Cyberspace: The Regulation of Ransomware Operations

(Source: The Oxford Process[302])

Given the ambiguities as to how international law applies to offensive cyberoperations, such as ransomware attacks, the *Oxford Statement* has gathered international legal expertise and provided critical guidance. Among others, the below principles (2, 3, 4 and 5) are particularly relevant to the present Case Study:

2. States must refrain from conducting, directing, authorising or aiding and assisting ransomware operations which violate the principles of sovereignty or non-intervention in a state's internal or external affairs, or amount to a prohibited threat or use of force within the meaning of the Charter of the United Nations. In particular, states must refrain from ransomware operations which are aimed at or result in disruption to electoral systems, healthcare, electric grids, water distribution systems, and nuclear power plants.

3. States must refrain from conducting, directing, authorising or aiding and assisting ransomware operations that result in violations of the human rights of individuals within their jurisdiction, such as the right to life, health, private life, education, property, freedoms of thought and opinion, freedom of expression, including the freedom to seek, receive and impart information and ideas of all kinds.

4. a) States must not allow their territory or infrastructure under their jurisdiction or control to be used by states or non-state actors for ransomware operations that are contrary to the rights of other states, when the former states know or should know of such operations.

   b) To discharge those duties, states from which ransomware operation emanates, in full or in part, must take feasible measures to stop such operations and otherwise address the situation. Such measures may include the conduct of investigations, the adoption of legal and technical measures, as well as cooperation with other states. Any measures taken in this regard must be compliant with applicable obligations under international law, including international human rights law.

[300] Human Rights Committee, General Comment No. 36 (2018) on Article 6 of the International Covenant on Civil and Political Rights on the right to life, UN Doc. CCPR/C/GC/36, 30 October 2018. See Idem at para. 63, para. 66.

[301] Office of the High Commissioner for Human Rights, Committee on Economic, Social and Cultural Rights (CESCR) General Comment No. 14: The Right to the Highest Attainable Standard of Health (Article 12), p. 14, https://www.refworld.org/pdfid/4538838d0.pdf#page=14.

[302] The Oxford Process on International Law Protections in Cyberspace is an initiative of the Oxford Institute for Ethics, Law and Armed Conflict (ELAC) at the Blavatnik School of Government that was set in motion in May 2020 in partnership with Microsoft. At the core of this initiative lies a collaborative effort between international legal experts from across the Globe aimed at the identification and clarification of rules of international law applicable to cyber operations across a variety of contexts. The Oxford Process complements other efforts and initiatives related to cyberspace governance, including those undertaken by the United Nations and its various bodies. While the Oxford Process is an informal initiative, its findings and discussions can potentially feed into and inform formal processes within the UN and other international fora. The process is consistent with the broader international efforts to promote responsible state behavior in cyberspace, respect for human rights online, and the development of norms to prevent cyber conflicts and cyber warfare.

5. States must take measures to protect the human rights of individuals within their jurisdiction from harmful ransomware operations, including when such operations are carried out by other states and non-state actors. To discharge this obligation, states may, among other measures, prohibit ransomware by law, take feasible steps to stop ransomware operations, mitigate their effects, investigate and punish those responsible, as well as prevent and suppress ransom payments to the extent possible. Where such protective measures interfere with other human rights, they must conform with applicable legal requirements, such as legitimate purpose, legality, necessity, proportionality and non-discrimination.

INTERNATIONAL COUNTER RANSOMWARE TASKFORCE

At the multilateral level, efforts to strengthen international legal frameworks and hold states accountable for the offensive ransomware campaigns perpetrated by proxies, and often with their tacit support, have been limited. However, a growing number of states—under the leadership of the United States and Australia, and to some extent, Estonia and Germany—have turned to renewed international cooperation on cybercrime and cyber defense. Several White House summits and joint statements have led to the International Counter Ransomware Initiative 2022 and the International Counter Ransomware Task Force (ICRTF).[303] Since January 2023, Australia has taken the lead of ICRTF to drive collaboration among a coalition of 36 member states and the European Union to counter the spread and impact of ransomware by: (1) holding ransomware actors accountable for their crimes and not provide them safe haven; (2) combating ransomware actors' ability to profit from illicit proceeds by implementing and enforcing anti-money laundering and countering the financing of terrorism measures; (3) disrupting and bringing to justice ransomware actors and their enablers to the fullest extent permitted under each partner's applicable laws and relevant authorities; and (4) collaborating in disrupting ransomware by sharing information about the misuse of infrastructure to launch ransomware attacks.

Interestingly, the new Task Force shows a substantial focus on tackling the existing challenges described in the Technical Section regarding tracing back and attributing cyber proxy offensive activity globally and across jurisdictions:

"We intend to improve our comprehensive and holistic understanding of the strategies used by these criminal actors and the means by which their malicious activity can be identified and addressed in respective jurisdictions to improve our tools, relevant authorities, and capabilities to disrupt. We commit to work together to prioritize disruption targets to leverage the breadth of authorities and tools available to pursue hard and complex targets more effectively. We intend to increase the number and impact of our disruption actions so that ransomware actors are stopped in their tracks. The Counter Ransomware Initiative is committed not only to protecting ourselves and each other from ransomware, but also to helping other countries protect and disrupt so that ransomware is unable to gain traction worldwide."[304]

It is important to note that, while such cooperation is initially rooted in the cybercrime and law enforcement domain, it also harbors a strong "active cyber defense" component, which, in some instances, might connect to military and defense aspects.

[303] "FACT SHEET: The Second International Counter Ransomware Initiative Summit," The White House, 01 November 2022, https://www.whitehouse.gov/briefing-room/statements-releases/2022/11/01/fact-sheet-the-second-international-counter-ransomware-initiative-summit/. "International Counter Ransomware Initiative 2022 Joint Statement," The White House, 01 November 2022, https://www.whitehouse.gov/briefing-room/statements-releases/2022/11/01/international-counter-ransomware-initiative-2022-joint-statement/.

[304] "International Counter Ransomware Initiative 2022 Joint Statement."

Cooperation through "active cyber defense"—operations aimed at disrupting the conduct of cyberthreat actors—is emerging as a new strategy in cyberspace, and members of the Task Force have committed to use all appropriate tools of national power to disrupt offensive ransomware operations. The Biden administration made clear that the United States will no longer view the surge of destructive ransomware attacks solely through the limited lens of law enforcement, but will also engage the expertise of the U.S. Cyber Command and National Security Agency.[305]

In a similar move, the Australian government announced a standing operation to hunt down ransomware threat actors and an effort that would join forces between the Australian Signals Directorate—its defense arm—and the Australian Federal Police.[306] An official statement described this effort as a way to "collect intelligence and identify ringleaders, networks and infrastructure in order to disrupt and stop their operations – regardless of where they are."[307] In November 2022, the Australian Cyber Security Centre stated, "We are currently witnessing deteriorating strategic circumstances in our region and globally, including a military build-up unseen since World War II, and expanding cyber and gray zone capabilities are of particular concern."[308] The Australian agency also warned that the regional dynamics in the Indo-Pacific were "increasing the risk of crisis" and cautioned that "cyber operations are likely to be used by states to challenge the sovereignty of others."[309]

As Maurer noted in 2018, "For offensive cyber actions, whose effects remain well below the threshold of what constitutes use of force – i.e., the vast majority of malicious cyber activity – the phenomenon of non-state actors operating extraterritorially points to increasing challenges for international law enforcement cooperation, including pressure on the existing extradition treaty regime and other jurisdictional nightmares."[310] Up to now, operational cooperation that achieves prosecution has been rare, but the ICRTF promises more momentum and hopefully results. It remains to be seen how ICRTF will support and accelerate international active cyber defense and law enforcement cooperation and how it will address complex evidentiary processes across jurisdiction, prosecution, and extradition.

[305] "FACT SHEET: Ongoing Public U.S. Efforts to Counter Ransomware," The White House, 13 October 2021, https://www.whitehouse.gov/briefing-room/statements-releases/2022/11/01/international-counter-ransomware-initiative-2022-joint-statement/.

[306] "Joint standing operation against cyber criminal syndicates," Australian Government Attorney-General's Portfolio, 12 November 2022, https://ministers.ag.gov.au/media-centre/joint-standing-operation-against-cyber-criminal-syndicates-12-11-2022.

[307] Ibid.

[308] *ACSC Annual Cyber Threat Report, July 2021 to June 2022*, Australian Government and Australian Cyber Security Centre, 04 November 2022, https://www.cyber.gov.au/about-us/reports-and-statistics/acsc-annual-cyber-threat-report-july-2021-june-2022.

[309] Martin, A., "'Cyberspace has become a battleground,' warns Australian Cyber Security Centre," *The Record*, 03 November 2022, https://therecord.media/cyberspace-has-become-a-battleground-warns-australian-cyber-security-centre/.

[310] Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 134.

GOVERNANCE SECTION: THE ACCOUNTABILITY CHALLENGE

The world has entered a complex and dangerous decade. As new and old threats converge and challenge the multilateral order, the most seismic shift is taking place at the intersection of war, technology, and cyberspace. **Modern conflicts—whether declared, contested, or waged in the grey zone—are born out of a technical revolution that is inherently dual-use**. These conflicts constantly merge physical and digital fronts, invading cities and factories, homes, and everyday devices, and producing new targets and victims in their wake. Lines between peace and war, offense and defense, civilian and military technologies, and patriots, combatants, and criminal groups are fading.

> Modern conflicts—whether declared, contested, or waged in the grey zone—are born out of a technical revolution that is inherently dual-use.

We face a privatization of cyber offense, which could lead to a privatization of cyberwarfare. Some might argue that surrogate warfare is not a new phenomenon and that the world has been forced to cope with mercenaries in the past. The difference is that, while in the past outsourcing war depended largely on arms trade and trafficking, **cyber proxies today thrive on the intangible transfer of dual-use knowledge and techniques**. Bringing transparency to such polymorphous ecosystems is not only hard; it also undermines what powerful states still see as the advantage of strategic ambiguity.

In this new strategic environment, hostile states gain an immense competitive advantage from cultivating and tapping into an ecosystem of cyber proxies that pervasively use obfuscation strategies to make attribution more difficult and obscure their relationships with supporting states. **As a result, there are less and less clear patterns of deputization, delegation, subordination, and control between nation states and their proxies and minimal evidence to ascribe responsibility for international wrongful acts and hostile behavior in cyberspace**.

There has also been a sharp erosion in governance and control—a challenge to which cyber proxies have greatly contributed, along with the merger of the underground cyber-arms and cybercrime industries. Disarmament efforts and non-proliferation architectures, formerly based on controlling physical weapons, need a fundamental rethinking and new anticipatory and strategic approach in the face of the rapid transfer of intangible dual-use knowledge between threat actors in the cyber and AI domains. This is even more true given that, in an increasingly conflict-ridden environment, there is no strategic reason for hostile states to "cyber disarm." Prominent experts have voiced another inconvenient truth that "in this contest, democratic states should not renounce cyberweapons and their use unless they are suicidal," and "States must defend themselves and their citizens."[311] Therefore, **imposing restraint, limiting the proliferation of offensive cyberservices, solidifying norms of responsible cyber behavior, and ensuring accountability and remedy to civilian populations constitute a set of pressing and unprecedented challenges for this century**.

[310] Lewis, "Private Actors' Roles in International Cybersecurity Agreements – Unlearned Lessons," *The Cyber Defense Review*, WINTER 2022, Vol. 7, N° 1, SPECIAL EDITION: Unlearned Lessons from the First Cybered Conflict Decade, 2010-2020 (Winter 2022), p. 34.

© UN Photo/Eskinder Debebe

> Imposing restraint, limiting the proliferation of offensive cyberservices, solidifying norms of responsible cyber behavior, and ensuring accountability and remedy to civilian populations constitute a set of pressing and unprecedented challenges for this century.

In this context, we should assess with some lucidity the progress made in clarifying how international law and the normative acquis can regulate and shape states' cyber conduct and cyber proxies' behavior. It has been almost a decade since inter-state negotiations on governance in cyberspace began, yet there is still no international consensus on **how to qualify and regulate the use of cyberattacks in wartime**.[312] Legal ambiguities remain in how to define "cyber weapon" and "cyber war crime," as well as how to determine the harm threshold that would qualify cyberattacks as "use of force" or "armed attacks."

Yet, the ongoing war of aggression in Ukraine and its cyber component has also prompted momentous legal and policy discussions. The ongoing conflict is the first instance of integrating cyberwarfare into an armed conflict, and there will be a "before and after" the war of aggression in Ukraine. As shown in Case Study 1, there are a wealth of initiatives to monitor and document the use of offensive cyberoperations in hostilities, attribute those operations to state-supported proxies, and collect evidence in order to determine whether IHL violations or cyber war crimes have been committed. Discussions among international organizations have also focused on what direct participation in hostilities means in cyberspace and what the implications could be for those who lose their protected status. Such efforts are critical to modernize the application of international law to cyberwarfare, along with important initiatives by ICRC on the development of a "humanitarian cyberspace" and the use of "digital emblems" for protected personnel and objects.[313]

Lessons learnt from the ongoing conflict in Ukraine will presumably contribute to current efforts to address hostile operations by cyber proxies. Among other deeply worrisome trends, it will remain difficult to map the increased use by cyber proxies of **cyber surveillance, information, and psychological operations directly targeted at civilian populations**. Such types of harmful and often aggressive targeting do not clearly meet existing thresholds within the international legal framework. Targeted civilians are unlikely to be protected by IHL, and perpetrators are unlikely to be prosecuted under the Rome Statute. International human rights law applies, but problems of compliance and accountability abound.

Global agreement in 2021 on **the normative acquis and a framework of responsible state behavior** was a substantial achievement to delineate a first set of primary rules for cyber governance. While the non-binding character of the normative acquis limits its observance by states, it nevertheless provides what some governments have termed rules of engagement or "rules of the road." In this instance, cyber norms can help clarify what constitutes wrong-

311 Lewis, "Private Actors' Roles in International Cybersecurity Agreements – Unlearned Lessons," *The Cyber Defense Review*, WINTER 2022, Vol. 7, N° 1, SPECIAL EDITION: Unlearned Lessons from the First Cybered Conflict Decade, 2010-2020 (Winter 2022), p. 34.

312 Idem, p. 38.

313 For "humanitarian cyberspace," see Marelli, M., "Hacking Humanitarians: Defining the Cyber Perimeter and Developing a Cyber Security Strategy for International Humanitarian Organizations in Digital Transformation," *International Review of the Red Cross* (2020), 102 (913), p. 367–387, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3969883. For "digital emblems," see ICRC, *Digitalizing the Red Cross, Red Crescent and Red Crystal Emblems: Benefits, Risks, and Possible Solutions*, 2022, https://shop.icrc.org/digitalizing-the-red-cross-red-crescent-and-red-crystal-emblems-pdf-en.html.

ful cyber behavior and legitimize counteraction against those states who do not observe them.[314] Case Study 2 on grey zone ransomware operations illustrates how UN cyber norms are used as a normative reference in the International Counter Ransomware Task Force, in inter-state discussions (that could later form *opinio juris*), and in international expert assessment such as the Oxford Process and the Cyber Law Toolkit. The most strategic contribution of this non-binding "normative jurisprudence" might therefore be to draw red lines that can be used to determine and legitimize a range of internationally lawful responses by injured states.

This is essentially a work in progress, and as shown in both Case Studies, there remain several challenges such as, *inter alia*:

- What evidentiary standards in attribution have to be met to hold states accountable when cyber proxies—tolerated or loosely affiliated with those states—are breaching UN cyber norms?
- What types of collective or coordinated counteractions can be taken by targeted states within the spectrum of active cyber defense (for example, self-help, disruption, or hunting down, which are all notions that have no formal definitions)? How to harmonize such coordinated counteractions with international law and existing rules on collective self-defense? And how to frame roles and responsibilities for private sector actors that would be needed to support such counteractions? To what extent could rising forms of active cyber defense collaborations between states and private sector actors erode the normative acquis?

This normative "work in progress" is getting more critical given that a number of Western states, including the United States, the United Kingdom, and Australia, have intensified efforts to attribute and respond to hostile cyber proxy activity, including by framing the types of countermeasures that constitute persistent engagement, active cyber defense, or what is also called "defense forward." Since 2018-2019, the U.S. Department of Defense and U.S. Congress have supported the conduct of military cyberoperations to defend the nation against active and systematic attacks by cyber proxies acting in the strategic interest of Russia, China, North Korea, and Iran.[315]

Following the war of aggression in Ukraine, similar positioning has taken place within NATO and the European Union. In June 2022, NATO's Strategic Concept pointed to the risk of hybrid operations conducted by states and their proxies and clarified application of Article 5 of the North Atlantic Treaty: "A single or cumulative set of malicious cyber activities; or hostile operations to, from, or within space; could reach the level of armed attack and could lead the North Atlantic Council to invoke Article 5 of the North Atlantic Treaty."[316] NATO's Strategic Concept added that hybrid operations—"coercive use of political, economic, energy, information and other hybrid tactics by states and non-state actors"—could equally reach the level of armed attack and could lead the North Atlantic Council to invoke Article 5 of the North Atlantic Treaty.[317]

In November 2022, the European Commission published the EU Cyber Defense Policy, inciting member states to "increase investments in full-spectrum cyber defense capabilities, including active defense capabilities" and emphasizing the need for a key partnership with NATO.[318] The EU Cyber Defense Policy states that "Whilst remaining fully committed to international law and norms in cyberspace, the EU should signal its willingness to use these [active

[314] See Lewis, "Private Actors' Roles in International Cybersecurity Agreements – Unlearned Lessons," p. 37.

[315] See DOD General Counsel, "Remarks at U.S. Cyber Command Legal Conference [Remarks By Hon. Paul C. Ney, Jr.]," U.S. Department of Defense, 02 March 2020, https://www.defense.gov/News/Speeches/speech/article/2099378/dod-general-counsel-remarks-at-us-cyber-command-legal-conference/. See also *John S. McCain National Defense Authorization Act for Fiscal Year 2019*, Public Law 115–232, 13 August 2018, https://www.congress.gov/115/plaws/publ232/PLAW-115publ232.pdf.

[316] *NATO 2022 Strategic Concept*, adopted by Heads of State and Government at the NATO Summit in Madrid, 29 June 2022, p. 7, para. 25, https://www.nato.int/strategic-concept/.

[317] Idem, p. 7, para 27.

[318] European Commission, Joint Communication to the European Parliament and the Council, *EU Policy on Cyber Defense*, 10 November 2022, p. 1, https://www.eeas.europa.eu/sites/default/files/documents/Comm_cyber%20defence.pdf.

defense] capabilities in a coordinated way in case of a cyberattack on a Member State."[319]

Persistent engagement has been relevant to the conflict situation in Ukraine. The U.S. Cyber Command has relied on its cyber defense personnel (in use already in 2021 as tensions were rising) and on private sector expertise to conduct a range of defensive, disruptive, and offensive cyberoperations to help protect Ukraine's infrastructures and reinforce cybersecurity in other allied countries.[320] Counteractions to information operations have also been deployed. Increasingly, in cyberconflict, offenses invade homes, cities, and factories, and concepts of cyber resilience and cyber defense therefore involve strengthened collaboration with the private sector (and, even, to some extent, civil society).[321] In this conflict situation, operational reality is evolving more rapidly than the clarification of international law, and Member States that aim to support the implementation of UN cyber norms may need to take all precautions not to undermine the normative acquis. As Beecroft notes, "Collective defense is not only demonstrating its operational potential in Ukraine but also revealing strategic tensions that would have to be addressed in any more enduring arrangements."[322] Beecroft adds, "At the heart of the challenge for democracies are the integration of commercial actors as agents of foreign and defense policies and the reality that a handful of American companies are indispensable to large-scale cyber defense."[323] Building mechanisms for coordinated responses between states and across sectors will raise important issues about limits and rules of engagement, accountability, and harmonization with international law and the normative acquis.

Outside of a declared conflict, active cyber defense is also important to counter offensive cyberoperations in the grey zone. It is at the core of the International Counter Ransomware Task Force, in particular the working group led by Australia on disruption operations. But international cooperation and joint efforts by law enforcement and cyber military authorities also brings into question the blurring of criminal and national security matters, and relevance of multilateral norms. At the same time, in situations that fall short of outright war, enhanced international cooperation in cybercrime prevention might remain the only practical governance avenue to tackle the global security implications of the merger of the cybercrime and cyberarms industries.

The present report closes with a reflection on the accountability challenge in cyberspace and cyber proxy relationships, including a set of recommendations for increased normative cooperation grounded in international law and for collaboration in non-proliferation and cybercrime prevention. The report's last section below takes stock of a wealth of insightful recommendations from interviews and discussions with experts, as well as from the UN Working Group on the use of mercenaries, the CyberPeace Institute, and the *Microsoft Digital Defense Report 2022*.

[319] Idem, p. 1.

[320] Cyber National Mission Force Public Affairs, *Before the Invasion: Hunt Forward Operations in Ukraine*, U.S. Cyber Command, 28 November 2022, https://www.cybercom.mil/Media/News/Article/3229136/before-the-invasion-hunt-forward-operations-in-ukraine/. See also Schmitt, M., "Ukraine Symposium—U.S. Offensive Cyber Operations in Support of Ukraine," *Articles of War*, 06 June 2022, https://lieber.westpoint.edu/articles-of-war/.

[321] See ICRC, *Avoiding Civilian Harm from Military Cyberoperations during Armed Conflict*.

[322] See Beecroft, *Evaluating the International Support to Ukrainian Cyber Defense*.

[323] Ibid.

RECOMMENDATIONS: REGULATING THE ROLE AND INVOLVEMENT OF OFFENSIVE PROXY ACTORS IN CYBERCONFLICT

I. STRENGTHEN NORMATIVE AND POLICY ENGAGEMENT AT THE MULTILATERAL LEVEL

1. **To provide accountability in cyberspace, states and non-state actors should attribute offensive operations by cyber proxies. To the extent possible, governments should coordinate and collaborate on the evidentiary process necessary for attribution, investigation, and prosecution and engage in a proportionate collective response**.[324] As indicated by the CyberPeace Institute, "Attribution can and should be done at the technical, legal and political level in order to provide all of the necessary information to provide evidence and create methodology for public awareness and effective judicial recourse." The UN Working Group on the use of mercenaries emphasizes that **"States must investigate, prosecute and sanction alleged violations of international humanitarian law and human rights abuses by mercenaries, mercenary-related actors and private military and security companies and provide effective remedies to victims."**[325] The *Microsoft Digital Defense Report 2022* also underlines that governments should "cite norms, laws and consequences in attribution" and "highlight what manner of consequences will be imposed to help strengthen recognition of international expectations."[326] According to cybersecurity and legal experts, referring to the international and normative framework in attribution is important because it **justifies the type of internationally lawful response, sanctions, and counteractions** that can be undertaken against those that do not observe laws and norms.[327]

   The challenge of collective response is not new but rising geopolitical tensions bring to the front important issues like accountability and burden-sharing for cyber defense and cyber governance. There is a need for like-minded states to discuss measures and mechanisms for the coordination of collective action, partnerships with existing alliances (e.g., NATO, EU Cyber Defense Policy, EU Cyber Diplomacy Toolbox), and harmonization with international law (collective self-defense), and the normative acquis. To this dialogue on collective response, states should integrate meaningful consultations with multistakeholder communities, in particular the cybersecurity industry.

2. **To clarify how international legal frameworks apply to cyberspace, states should explain how they understand their obligations under international law.** In the case of offensive operations by cyber proxies, it may be relevant for states to clarify their position on rules and principles of customary international law (including, sovereignty and due diligence), as well as Article 8 of the International Law Commission's Articles on State Responsibility.

   This effort of legal interpretation is crucial in situations below the threshold of war, for

[324] See CyberPeace Institute, "Mercenary-Related Activities in Cyberspace."

[325] A/76/151, p. 19, para. 75.

[326] See *Microsoft Digital Defense 2022 Report*, p. 53.

[327] Interviews with cybersecurity and legal experts. See also Lewis, "Private Actors' Roles in International Cybersecurity Agreements – Unlearned Lessons," p. 37.

instance when offensive cyberoperations target and cause transboundary harm to civilian populations and industries and constitute a **direct breach of the normative acquis** (cf. Case Study 3). As Lewis mentions, "Ultimately, the agreed norms can reinforce international law to reduce cyberconflicts."[328] But this effort is also extremely important in the current war of aggression against Ukraine as it may contribute to developing a **clearer understanding of how IHL and the Rome Statute of the ICC apply in the cyber context**. The Permanent Representative of Liechtenstein to the UN notes, "Such clarity is necessary for the Court's own work, but it can also help inform the work of the United Nations Security Council, in particular regarding how it uses its power to refer to situations involving acts of aggression to the ICC – a referral that provides an important enforcement mechanism in support of the UN Charter's prohibition on the use of force."[329] The Permanent Representative of Liechtenstein to the UN pursues, "We should be ready for the potential wars of the 21st century by deterring malicious cyberoperations through establishing the necessary means for accountability."[330]

3. **To better anticipate evolving threats from cyber proxies and to proactively pursue accountability, states should engage in multilateral dialogues that can help them recognize the strategies, behaviors, and modus operandi of cyber proxies**. Such dialogues could be instrumental to better understand the different ways that offensive operations by cyber proxies **may violate IHL and IHRL** and identify new norms, remedy, and reparation mechanisms. Case Studies 1 and 2 demonstrate the importance of including the field expertise and capacity of civil society organizations and human rights clinics in this monitoring effort.

As shown in the Technical Section of this report, potential beneficiaries of the underground cyberarms and cybercrime industries may increasingly include private mercenary groups, terrorist groups, transnational illicit networks, and proxy forces involved in conflict. Such "diffusion of cyber power" may rapidly reach a growing number of private sector offensive actors and private groups associated with mercenary activity. As such, there might be an **increasing correlation between criminal accessibility and mercenary and terrorist capability**. In this context, dialogues between states need to focus on the implications that such diffusion of cyberthreats poses to the operational arena of **non-international armed conflicts**. Furthermore, as stated by the UN Working Group on the use of mercenaries, "States should agree on and support international processes to identify, assess and further develop mechanisms to more clearly and formally recognize the **international human rights obligations of armed non-State actors**, including criteria to determine the latter's capacity to hold human rights obligations."[331]

RELEVANT SYNERGIES: The three recommendations above are also relevant to the ongoing effort of the OEWG on ICT in the context of international security; ICRC's Support Relationships in Armed Conflicts effort; UN disarmament and non-proliferation agencies; regional security organizations; the UN Working Group on the use of mercenaries; and the Open-ended intergovernmental working group to elaborate the content of an international regulatory framework on the regulation, monitoring, and oversight of the activities of private military and security companies.

[328] Idem, p. 38.

[329] The Permanent Mission of Liechtenstein to the United Nations, *The Council of Advisers' Report on the Application of the Rome Statute of the International Criminal Court to Cyberwarfare*, Foreword.

[330] Idem.

[331] A/76/151, p. 20, para. 79.

II. REINFORCE INTER-STATE COOPERATION AND TRANSLATE NORMATIVE EFFORTS INTO ACTION AT THE NATIONAL LEVEL

For decades, technical challenges and a lack of capacity-building have hindered effective criminal justice responses to global cybercrime. In particular, a lack of expertise in applying cyber forensics and in documenting digital evidence has impaired the capabilities of law enforcement in individual nations, which in turn has slowed down transnational investigations. Developing sufficient skills, as well as strong and coordinated legal instruments across jurisdictions, remains a challenge to effectively investigate and prosecute cybercrime. Enhanced international cooperation in cybercrime prevention is even more pressing today to tackle the global security implications of the merger of the cybercrime and cyberarms industries. **It might also constitute an increasingly powerful way to shape, influence, and disrupt the ecosystems of offensive cyber proxies cultivated by some nation states**.

4. **Within the International Counter Ransomware Initiative and Task Force, states should build collective resilience to offensive ransomware operations, counter illicit finance that underpins the ransomware ecosystem, work with the private sector to defend against ransomware attacks, cooperate to disrupt ransomware, and pursue the actors responsible to the full extent permitted under each partner country's applicable laws and relevant authorities**. Importantly, experts indicate, "Strong capacity building efforts and coordination with those countries that struggle most with technical and policy cyber capabilities could help stop them from becoming the focus of attention for cybercriminals looking for easy targets"[332] and safe havens.

5. **Building on successes within the International Counter Ransomware Initiative and Task Force, states should address the industrialization of cybercrime across appropriate multilateral formats to establish broader-based practices, actions, and norms and cooperate internationally across all aspects of the cybercrime threat**. Such cooperation could help law enforcement agencies, often in partnership with diplomats and the private sector, to build and develop the capability and technical expertise to attribute, investigate, and prosecute cybercriminals, including across multiple legal jurisdictions.[333]

III. DEFINING GOVERNMENTAL AND CORPORATE RESPONSIBILITY

6. **States should implement a transparent and operational framework to determine what element of offensive cyber activity constitutes *inherently governmental functions* and what element constitutes *closely associated functions* that can be performed by private sector actors**.[334] Distinguishing what offensive cyberoperation functions can be outsourced to a third party (and increasingly, automated systems) has legal consequences for determining direct participation in hostility and lawful targets for counterattacks. Both the UN Working Group on the use of mercenaries and the CyberPeace Institute have emphasized the importance of implementing a transparent and operational framework through which states should be transparent about the outsourcing and contracting of military services that support offensive cyberoperations, including "the nature of services, procurement procedures, the terms of contracts and the names of services providers in a sufficiently detailed and timely manner."[335] The CyberPeace Institute insists on the need to distinguish more clearly between offensive and defensive services in contracts and procurement procedures, and to clarify the legal status of military and security services provided in cyberspace

[332] Lostri, E., "Keeping up with Ransomware," *Lawfare*, 18 November 2022, https://www.lawfareblog.com/keeping-ransomware.

[333] See excellent analysis by Vignard, K. and Hakmeh, J., *ICTs, International Security, and Cybercrime*, United Nations Institute of Disarmament Research, 11 October 2021, https://unidir.org/publication/icts-international-security-and-cybercrime.

[334] See Maurer, *Cyber Mercenaries: The State, Hackers, and Power*, p. 142-143.

[335] A/76/151, p. 19, para. 74.

[336] CyberPeace Institute, "Mercenary-related Activities in Cyberspace."


© iStock/sdecoret

as a way to ensure respect of human rights obligations and accountability.[336]

7. **When it comes to private sector offensive actors**, the UN Working Group reiterates that **"States should refrain from recruiting, using, financing and training mercenaries and should prohibit such conduct in domestic law and effectively regulate private military and security companies."**[337] Among the services offered by private sector offensive actors are targeted cyber surveillance, cyber intrusion, and information and influence operations. As mentioned throughout this report, limited legal measures exist to effectively reign in such an underground market and the hostile services it provides during and outside armed conflict. The 2019 report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression already described a market shrouded in secrecy, functioning close to impunity, with corrosive implications for the protection of human rights globally. The Special Rapporteur called for **"tighter regulation of surveillance exports and restrictions on their use**, as well as a call for an immediate moratorium on the global sale and transfer of the tools of the private surveillance industry until rigorous human rights safeguards are put in place to regulate such practices and guarantee that Governments and non-State actors use the tools in legitimate ways."[338]

Targeting and neutralizing the underground cyberarms industry and its exchanges with hostile states and non-state actors is a massive and complex challenge for which traditional non-proliferation approaches are inadequate. Offensive surveillance and cyber capabilities often depend on the intangible transfer of dual-use knowledge between malicious actors and are commodified based on civilian technologies outside of highly classified settings such as military, defense, and intelligence agencies. Outsourcing offensive cyberoperations on behalf of state actors in a self-

[337] A/76/151, p. 19, para. 73.

[338] Human Rights Council (41st Session), *Surveillance and human rights - Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, p. 1.

regulated market therefore tends to "provide a stress test to states' will and capacity to monitor and enforce its obligations to respect, protect and fulfil human rights; some deliberately choose to use mercenaries in an attempt to escape accountability."[339] In this context, **any regulatory regime for offensive cyber capabilities needs to move beyond export control, severely target and reign in private sector offensive actors, and oversee innovation in civilian sectors**.

8. **The sensitive nature of cyber defense and cyber offense services—increasingly converging with AI technologies—requires a rigorous set of actions across the civilian sector for companies to meet their responsibilities to be in full compliance with international human rights law, international humanitarian law, and international criminal law**. Using a "human rights by design" approach, companies in the civilian sector should:

   - Anticipate the illicit use of their technology and begin engineering solutions for the inevitable impacts.
   - Rely on contractual, technical, and organizational mechanisms to ensure that sensitive dual-use technology remains in the hands of their legitimate customers and does not spread to organizations that fall outside the scope of due diligence.
   - Formulate operational policy safeguards that ensure that the use of (active) cyber defense services is in full compliance with international law.
   - Assess the remediability of potential harms related to technological misuses and prioritize due diligence and transparency for those that are more challenging to rectify.
   - Develop, implement, and support effective grievance mechanisms for vulnerable populations whose rights may be harmed by technological misuses and offensive uses, including national civil and military justice systems and international or regional human rights mechanisms and courts.

A UN PERMANENT ACCOUNTABILITY MECHANISM FOR CYBERSPACE

A growing number of experts have argued in favor of the development of a standing accountability body to support responsible state behavior in cyberspace, a permanent UN mechanism to deal with cyberspace as a domain of conflict.[340] In the *New Agenda for Peace* published in July 2022, the UN Secretary-General calls for establishing an independent multilateral accountability mechanism for the malicious use of cyberspace by States to reduce incentives for such conduct.[341] In the past, UN working groups have provided strategic forums for policy and normative discussions but have been limited in their capacity to ensure accountability in cyberspace.

As Lewis explains, legal and political attribution of offensive cyberoperations should remain a sovereign responsibility and a state should ultimately remain in charge of the evidentiary process and the political analysis (trade-off) that comes with attribution.[342] Yet, the multilateral dimension can offer strategic support in "developing common evidentiary standards and information-sharing mechanisms for coordination of collective attribution."[343] As Lewis adds, "Coordinated attribution of malicious activity will require better information

[339] CyberPeace Institute, "Mercenary-related Activities in Cyberspace."

[340] See Lewis, *Creating Accountability for Global Cyber Norms*. See CyberPeace Institute and Moriani, "Untangling Accountability in Cyberspace." See *Microsoft Digital Defense Report 2022*, p. 53.

[341] See *Our Common Agenda: Policy Brief 9—A New Agenda for Peace*, p. 27.

[342] Lewis, *Creating Accountability for Global Cyber Norms*, p. 4-7.

[343] Idem, p. 9.

sharing between partners, and perhaps new mechanisms for sharing and harmonization, but will greatly strengthen the political effect of any accusation."[344] To increase accountability, states will also need to collaborate on "a broadly accepted menu of possible consequences and an ability to ensure that any consequences imposed are both proportional to the initial incident and consistent with international law and practice."[345]

This is a crucial collaborative endeavor, which could help address some of the most worrisome trends illustrated in this report, in particular the merger and exploitation of the cyberarms and cybercrime industries by nation states. Such a multilateral accountability mechanism would be instrumental:

- To build capacity and strengthen methods and practices to monitor offensive activity by cyber proxies, and support the evidentiary process required for coordinated attribution, investigation, and prosecution efforts, both in situations of peace and conflict; this could lead to a coordinated capacity for technical, legal, and political attribution that could benefit states with less expertise and capabilities.
- To collaborate on a range of internationally lawful responses with important implications to hold states and non-state actors accountable for hostile behaviors in cyberspace.
- To better analyze and anticipate current and evolving forms of offensive cyberoperations and the modus operandi, strategies, and behaviors of cyber proxies; such anticipatory analysis and foresight capacity could support prevention and mitigation of civilian harm and would progressively constitute an "institutional memory" of evolving threats in cyberspace as a domain of conflict.
- To develop understanding of the evolving forms of dual-use technologies in cyberspace and related technological and knowledge transfer between actors; such interest in adaptive governance and responsible innovation would help modernize disarmament and non-proliferation efforts.
- To support ongoing normative efforts that aim to clarify how international law applies to cyberspace, in particular discussions to reaffirm states' obligations and responsibilities (including in their relationships with proxies) and to clarify the under-conceptualized, under-regulated zone that non-state actors occupy in cyberspace (for example, the legal definition of cyber proxies or cyber mercenaries).
- To support capacity-building efforts that involve countries most impacted by the digital and cybersecurity divides.

[344] Ibid.
[345] Ibid.

ACRONYMS

AI	Artificial intelligence
ANSA	Armed non-state violent actors
APT	Advanced persistent threat
CAC	Cyberspace Administration of China
CISA	U.S. Cybersecurity and Infrastructure Security Agency
DDoS	Distributed denial of service
EU	European Union
FBI	Federal Bureau of Investigation of the United States
FSB	Federal Security Service of the Russian Federation
GGE	Group of Governmental Experts
GRU	Main Directorate of the General Staff of the Armed Forces of the Russian Federation
ICC	International Criminal Court
ICRC	International Committee of the Red Cross
ICRTF	International Counter Ransomware Task Force
ICT	Information and communication technology
IHL	International humanitarian law
IHRL	International human rights law
ILC	International Law Commission
INTERPOL	International Criminal Police Organization
IS	Islamic State
IT	Information technology
NATO	North Atlantic Treaty Organization
NSA	U.S. National Security Agency
OEWG	Open-ended Working Group
PMSC	Private military and security companies
PSOA	Private sector offensive actors
SAIC	Science Applications International Corporation
SEA	Syrian Electronic Army
TTP	Tactics, techniques, and procedures
UK	United Kingdom
UN	United Nations
UNODC	United Nations Office of Drugs and Crime
U.S.	United States

ABOUT THE AUTHOR

Eleonore Pauwels is an international expert in the security, societal, and governance implications generated by the convergence of artificial intelligence with other dual-use technologies, including cybersecurity, genomics, and genome-editing. Pauwels provides expertise to the World Bank, the United Nations, and the Global Center on Cooperative Security in New York. She also works closely with governments and private sector actors on the changing nature of conflict, foresight, and global security, as well as responsible innovation related to AI, biotechnologies, and cybersecurity.

In 2018 and 2019, Pauwels served as Research Fellow on Emerging Cybertechnologies for the United Nations University's Centre for Policy Research. At the Woodrow Wilson International Center for Scholars, she spent ten years within the Science and Technology Innovation Program, leading the Anticipatory Intelligence Lab. She is also part of the Scientific Committee of the International Association for Responsible Research and Innovation in Genome-Editing (ARRIGE). Pauwels is a former official of the European Commission's Directorate on Science, Economy and Society.

Pauwels regularly testifies before U.S. and European authorities including the U.S. Department of State, NAS, NIH, NCI, FDA, the National Intelligence Council, the European Commission, and the UN. She writes for *Nature*, *The New York Times*, *The Guardian*, *Scientific American*, *Le Monde*, *Slate*, *UN News*, *The UN Chronicle*, and The World Economic Forum.

Konrad Adenauer Foundation
New York Office

220 E. 42nd Street, Suite 3300
New York, NY 10017
www.kas.de/newyork

newyork@kas.de