ISSUE BRIEF
JUNE 2024

CYBER MERCENARIES DECODED

Perspectives on Countering Proliferation of Cyber Intrusion Capabilities

Pablo Rice
Cyberspace Governance Programme

PARIS PEACE FORUM
de PARIS sur la PAIX

NOTE

This issue brief is derived from a course delivered at the South Africa - Netherlands Cyber Security School 2024. The views expressed in the publication are the sole responsibility of the individual author. They do not necessarily align with those of the Paris Call for Trust and Security in Cyberspace community, the Paris Peace Forum, its staff members or partners.

ABOUT THE PPF CYBERSPACE GOVERNANCE PROGRAMME

The digital transition and accelerating technological shifts mean that our safety, prosperity and access to basic services as well as to the public sphere increasingly relies on access to a functioning, free and open cyberspace. The rapidly evolving international landscape and rising tensions across the globe as well as the emergence of new kinds of threats poses a risk to the openness and stability of cyberspace and requires a strengthening of global governance for the three layers of the Internet.

The Paris Peace Forum's Cyberspace Governance Programme is committed to address this issue by coordinating the Paris Call for Trust and Security in Cyberspace. Launched at the 2018 edition of the Forum, The Paris Call provides a platform to advance comprehensive norms for human security in cyberspace and to ensure better collaboration among all actors, whether public and private in this regard. It currently gathers 1200 supporters, including 80 States, 2 international organizations, 700 companies and 350 organizations from the civil society from across the world and the digital ecosystem, around 9 core principles to secure the open, free and stable cyberspace.

The work within Paris Call is currently organized around two long-term objectives:

- Countering the Proliferation of Commercial Cyber Intrusion Capabilities
- Protecting Transnational Critical Infrastructures - the Public Core of the Internet

In line with its mandate to support, improve, and complement institutional multilateralism, the Forum is articulating the Paris Call's efforts with current developments within key multilateral processes on the use of ICTs. As such, the Forum contributes in particular to the United Nations' Open-ended Working Group on security of and in the use of information and communications technologies.

CONTENTS

Introduction

Since 2021, the notion of the "cyber-mercenary" has gained a great deal of attention in the international public sphere, echoed in numerous press headlines but also within industry circles[1] and civil society[2]. In June this year, however, the dissemination of the concept shifted gears as the United Nations Secretary-General took up the notion in his address to the Security Council during a high-level debate on evolving threats in cyberspace. On this occasion, António Guterres warned of illicit activities by the "*so-called cyber-mercenaries*" and stressed that "*software vulnerabilities are being exploited and cyber-intrusion capabilities are even sold over the Internet*" - thus contributing to the weaponization of digital tools worldwide[3].

One might assume that a so commonly used term would enjoy a fairly clear consensus on its very meaning and scope. Yet, the definitional aspect remains the most challenging one given the opacity of the ecosystem to which the concept seems to refer. It may also be the most crucial, as it provides the foundation for any comprehensive policy action to address this issue.

I - From mercenary spyware to cyber-mercenaries: a mere extension of mercenarism's conventional scope?

Bringing up "mercenaries" in the cyber context started to gain traction in 2016, alongside the revelations made to the public at large of massive governments' use across the world of the sophisticated spyware (for spying software) Pegasus, made commercially available by the Israeli firm NSO. This software bypasses most smartphones' security using zero-day vulnerabilities acquired or discovered by NSO, granting remote access to files, messages, photos, passwords, enabling audio, video recording and geolocation tracking discreetly from a workstation installed at the client's site. NSO asserts that its product stands out from any typical malware by being more "transparent", evading antivirus detection while not requiring any kind of involvement from the targets in order to be installed on their device[4].

Presented as a solution for fighting terrorism and organized crime, Pegasus has been revealed by several civil society organizations and investigative journalists to be used against political opponents, human rights activists, journalists, executives, and officials within and beyond the borders of the client State. The echo given to these disclosures was largely influenced by the number and standing of identified victims.

[1] Cybersecurity Tech Accord, "Cyber mercenaries: An old business model, a modern threat", March 2023
[2] See, in this regard: Tim Maurer, The State, Hackers, and Power, Cambridge University Press, 2018
[3] Secretary-General's remarks to the Security Council's High-Level Debate on "Maintenance of International Peace and Security: Addressing Evolving Threats in Cyberspace", UN Statements, June 2024
[4] NSO, Pegasus Product Description, 2019

In a 2020 report grounded in forensic analysis, two United Nations special rapporteurs implicated Saudi Arabia in infiltrating Jeff Bezos' phone with such a tool. Then, in 2021, an investigative journalist consortium coordinated by the French NGO Forbidden Stories suggested that Pegasus might have played a role in the assassination of Saudi journalist Jamal Khashoggi at the Saudi embassy in Istanbul in 2018, relying on a listing recovered by Amnesty International[5]. The same consortium also revealed that the target list included telephone numbers of at least six current heads of state, among them French President Emmanuel Macron.

At the forefront of this global effort in unveiling spyware misuse is the Citizen Lab at the University of Toronto, an interdisciplinary laboratory on tech policy whose director Ronald Deibert first brought the concept of "mercenary spyware" to public awareness[6]. The term "mercenary" was likely chosen opportunistically back then to resonate with the general public. It highlights that this intrusive software was neither developed internally by states nor the work of state-sponsored actors per se, but rather part of a commercial service offered by economic operators, some of them - such as NSO - maintaining legal existence as being regularly established within a State's jurisdiction.

More than just a discursive element, "mercenary" also carries significance in international instruments, denoting specific actors in the physical world. This prompts consideration of its relevance in the cyber domain - in light of agreed international norms.

- In its article 47, the **first Protocol Additional to the 1949 Geneva Convention** defines mercenaries as as any "*person who is specially recruited locally or abroad in order to fight in an armed conflict, taking actually a direct part to the hostilities*". In this context, "*he has to be motivated essentially by the desire for private gain that in fact, is promised, by or on behalf of a Party to the conflict, material compensation substantially exceed the one promised or paid to combatants of similar ranks and functions in the armed forces of that Party*". Lastly, he is "*neither a national, a resident of a party, nor is a member of a Party's armed forces and has not been sent by a third-party state on official duties*"[7].

In the same vein as for all groups and individuals designated under international humanitarian law, a mercenarism exclusively tied to armed conflict contexts challenges its application to cyber operations. The controversy lies in defining the threshold of violence intensity necessary to qualify such situations - which remains highly contentious in this environment.

[5] Forbidden Stories, "The Rise and Fall of NSO Group", Pegasus Project, 2021
[6] See, in this regard: Ronal Deibert, "The Autocrat in Your iPhone: How Mercenary Spyware Threatens Democracy", Foreign Affairs, January 2023
[7] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977, Article 47 "Mercenaries"

- Asserting that the reality of mercenaryism can't be exclusively understood from the perspective of participation to armed conflicts' hostilities, the **1989 International Convention against the Recruitment, Use, Financing and Training of Mercenaries** proposes retaining the definition in First Protocol Additional while adding certain mentions aimed at covering other contexts. In the latter case, a mercenary is described in the article 1 as "*any person who Is specially recruited locally or abroad for the purpose of participating in a concerted act of violence aimed at, overthrowing a government or otherwise undermining the constitutional order of a State (...) or undermining the territorial integrity of a State*" - provided that the conditions of compensation and independence are also met[8].

Such definition aligns better with cyberspace parameters since the concept of territorial integrity infringement can be broadly interpreted. However, defining a "concerted act of violence" in the context of ICT use raises significant socio-technical challenges. In particular, unpacking intents and the organizational degree of a virtually constituted group adds a layer of complexity to the already intricate attribution issue.

- The mercenary concept additionally bears similarities with the notion of "Private Military and Security Companies" (PMSCs), which emerged later to accommodate the evolving landscape of conflicts marked by blurred lines as well as the growing privatization and outsourcing of defense functions. The **Montreux Document of 2008** stands out as a significant intergovernmental policy instrument, a product of international collaboration spearheaded by Switzerland and the International Committee of the Red Cross. This document, though not legally binding, carries a certain political weight and has garnered endorsement from 17 states. It defines PMSCs as "*private business entities that provide military and/or security services, irrespective of how they describe themselves*"[9].

While this definition captures activities akin to those of notable entities such as NSO, it faces two important challenges. Firstly, its breadth may encompass a wide array of activities, potentially diluting its specificity. By focusing on "business entities," it also risks excluding various non-corporate actors, including underground entities, groups, or individuals operating beyond the purview of traditional market structures.

Formulating a clear and comprehensive definition that would fit in the cyber realm remains an ongoing hurdle, as boundaries in for this environment tend to blur, intentions resist clear categorization, conflicts take on hybrid forms, and proxy use becomes increasingly prevalent.

[8] International Convention against the Recruitment, Use, Financing and Training of Mercenaries General Assembly resolution 44/34, 4 December 1989, Article 1
[9] The Montreux Document On pertinent international legal obligations and good practices for States related to operations of private military and security companies during armed conflict, 17 September 2008, Preface

This state of play is precisely acknowledged in the 2021 thematic report from United Nations Working Group on the use of mercenaries, an expert body established by the Office of the High Commissioner for Human Rights. It nevertheless sought to put forward a new proposal in this regard, designating cyber-mercenaries as "*these actors who benefit from developing, maintaining and operating cybercapabilities and cyberservices, which might be used in the conduct of hostilities, in conflict and in non-conflict settings*". As for the core material element of this definition, the experts elaborate that cyber services encompass both providing support services to States concerning their existing cyber capabilities, and supplying cyber products that States can then use. Adding a functional component, the Group specifies that the end use of such supply involves, inter alia, sabotage via malware and ransomware, espionage and subversion including by spreading misinformation and disinformation[10]. In light of the numerous scenarios envisaged by the report, such proposal gains in precision yet risks targeting too broadly, so as to cover all cyber tool and service providers - even those strictly focused on defensive purposes such as penetration testing, honeypots, or tarpits.

II - Cyber-Mercenaries as a commercially-driven ecosystem with potential for proliferation

To prevent any initiatives in this field to be hindered by lingering definitional issues, one could navigate by shifting from an actor-centric focus to an approach that seeks to capture a commercial ecosystem as a whole, while highlighting the dynamics driving its growth. Several unpacking attempts have been undertaken in recent years, with the most convincing ones achieved under the Atlantic Council's Cyber Statecraft Initiative[11]. But building operational categories for this environment is by no means easy, as the capabilities and services blamed for violating the rights and interests of individuals and States may intersect with various markets that was identified for other purposes - such as financial analysis. The surveillance technology market or the so-called "lawful interception" market could therefore easily be invoked. The value of the last, estimated at $12 billion by Moody's a few years ago, is frequently highlighted in both academic research and political efforts on cyber-mercenaries and commercial spyware, underscoring the persistent confusion surrounding these ecosystems.

For the sake of clarity and with these limitations in mind, this paper focuses on three key segments of the cyber-mercenary supply chain deployed in both the semi-regulated and underground market, drawing empirically on the most notable case analyses. It should be

[10] Report of the Working Group on the use of mercenaries, A/76/151: "The human rights impacts of mercenaries, mercenary-related actors and private military and security companies engaging in cyberactivities", July 2021, pp. 4-7
[11] See, in particular : Winnona DeSombre & al., "Countering cyber proliferation: Zeroing in on Access-as-a-Service", Cyber Statecraft Initiative Report, March 2021; Jen Roberts, Trey Herr, Emma Taylor, and Nitansha Bansal, "Market Matters: A Glance into the Spyware Industry", Cyber Statecraft Initiative Report, April 2024

noted, however, that these segments and related functions can often be intertwined or overlapping, with the same actor undertaking in-house multiple tasks falling to other "nodes".

A) Segment #1 - Vulnerability research and exploit development

Researching systems' vulnerabilities, especially the so called zero-day vulnerabilities, is not inherently detrimental. It can occur in academic setting, or under corporate or State-led frameworks - including bug bounty programs and coordinated disclosure processes. When conducted within well-designed frameworks that should increase legal certainty, "good faith vulnerability research"[12] can significantly bolster overall cyber security and stability[13]. Looking at the last 15 years however, one can't ignore the discernible shift from a genuine research "community" to something closer, at least in certain spheres, to a vulnerability "industry".

The rise of vulnerability brokers as trusted intermediaries enabling increased anonymity of transactions has been instrumental in this respect. The emergence and growing competition from "clearing houses" has further boosted the commercial impetus of vulnerability research. These new players operate with an in-house research team, while securing exclusivity for certain vulnerabilities acquired directly on the market - refining them into "production-ready" exploits[14].

Should profit-driven motives override due diligence in transactions, there's a significant risk that these vulnerabilities and related exploits may fall into the hands of malicious entities, such as cybercriminals, adversarial State actors, as well as other doubtful private operators acting in other segments of the cyber-mercenary supply chain. Earlier this year, Google's Threat Analysis Group thus attributed 60% of the 97 zero-day vulnerabilities observed in 2023 as being exploited in the wild for threat actors' motivations, that especially include espionage and financially motivated hacking[15].

B) Segment #2 - Monetized intrusive capabilities development

Exploits purchased directly from a third party or refined after the acquisition or in-house discovery of a vulnerability can then be integrated in the development of an intrusive capability, often in a combined fashion to reach the three functions they can separately

[12] For a definition of good faith researcher, see: Tarah Wheeler, "How to recognize a good faith cybersecurity researcher as opposed to a computer criminal", GFCRC, 2023
[13] See, in this regard : Organization for Economic Co-operation and Development, Working Party on Security in the Digital Economy, "Encouraging Vulnerability Treatment Overview for policy makers", February 2021
[14] Maor Shwartz, "The boom, the bust and the adjust: The offensive cybersecurity industry — trends and updates", Medium, June 2023
[15] Google Threat Analysis Group & Mandiant, A Year in Review of Zero-Days Exploited In-the-Wild in 2023, March 2024

enable - access, escalation of privilege and code execution.

Exploits must therefore be seen as a crucial gate opener for the two other components of an intrusion capability, the payload and the propagation method, to be implemented. As an immediate purpose of the capability, the payload is written to achieve specific effects in the targeted systems, such as data theft, alteration, or remote activation of functionalities. The propagation method is the delivery technique for the execution of the capability in the targeted system. Available options in this regard, including compromised website or email attachment, are selected based on the characteristics and scale of targets at stake.[16]

Ready-to-use capabilities are offered either as standalone product or in relation with a broader package of services that typically include customer support for installation, operation, maintenance, and upgrades. Using some of them can also follow a license model, with a limited number of targets for each license. As such, the transaction is not necessarily extinguished with the delivery of the capability. Ties between suppliers and customers can be made even more permanent when the capability at stake rely on a command-and-control infrastructure that is still owned and operated by the supplier itself. NSO's Pegasus thus depends of several proprietary networks, known as the Pegasus Anonymizing Transmission Network and made of hundreds of domain names, DNS servers and other network infrastructures[17].

C) Segment #3 - Operational delivery

In this third segment, the transaction's focus is not on the intrusion capability itself, but on the actual use of such tools. Individuals, groups, or organizations operating either covertly or within legally established entities thus execute intrusive and/or offensive operations using a capability they have either purchased or developed in-house. The most recent political works tend to identify two types of actors in this space, distinguishing hacking as-a-service companies from hackers-for-hire[18].

Hacking-as-a-service companies are formally established entities that offer unauthorized computer system penetration as a service. These companies operate like regular businesses, providing clients with the means to infiltrate third-party systems without consent. Customers specify their requirements, such as target selection, and use the resulting information. On the other hand, hackers-for-hire are unaffiliated individuals or groups who operate on a more freelance basis. They are contracted to perform specific tasks that involve penetrating computer systems to fulfill the clients' requirements. The opacity surrounding their

[16] For a detailed explanation of malware constitutive elements: Trey Herr, "PrEP: A Framework for Malware & Cyber Weapons", Journal of Information Warfare, Vol. 13, No. 1, 2014
[17] For further description of the Pegasus infrastructure: Bill Marczak, John Scott-Railton, and Ron Deibert, "Infrastructure Linked to Targeting of Amnesty International and Saudi Dissident", The Citizen Lab Research Report #110, July 2018
[18] See, in this regard : Pall Mall Process Foundational Declaration, Annex A, February 2024

contractual ties are likely even greater than that of hacking-as-a-service companies, which however often operate under storefronts more or less distant from their core activities.

III - Policy and legal options to curb harmful proliferation and irresponsible use of cyber intrusion capabilities

Understanding the cyber-mercenary phenomenon is further complicated by the dual nature of the segments detailed above. Cyber intrusion capabilities, spanning vulnerability research, exploit development and offensive security services, can serve ethical and legitimate purposes in strengthening clients' overall cybersecurity or enabling digital forensics for national security interests - bearing in mind, however, the variable and potentially opportunistic nature of the latter. Using these tools and means also risks enabling harmful activities that undermine cyber peace and stability - including through a potential backlash effect where a priori legitimate customers sustain the growth of market players which subsequently offer their services to adversarial threat actors, thereby undermining the very security they once aimed to fortify. Such intricacies obfuscate clear policy boundaries and has led to favor case-by-case approaches to disrupt specific actors and tools over the past decade, rather than attempting to tackle the phenomenon as a whole. The various policy tools and strategies implemented thus far can be broadly classified based on the proliferation vector they aim to target: suppliers, customers, or compromised service/device providers (the "endpoints")[19].

A) Targeting the supplier

i. Export control measures

a) Targeted export restrictions

These restrictions are company-specific, imposing requirements such as export or transfer licenses, or outright denial of such licenses. In 2021, the US Commerce Department added four major spyware firms, including NSO Group and Candiru, to its Entity List[20]. US companies are thus prohibited from exporting, reexporting, or transferring items to listed entities without obtaining a special license. Similarly, in early 2022, the Israeli Defense Exports Control Agency froze the granting of export licenses to cyber intrusion companies based in Israel, later allowing exports only to democratic countries[21].

[19] For a broader overview of the policy toolbox implemented to date to specifically counter commercial spyware abuse, see: Freedman Consulting LLC, "Spyware Accountability Mechanisms: A Guide to Support Discussions on Spyware Accountability", September 2023

[20] US Department of Commerce, Bureau of Industry and Security, 86 FR 60759, Addition of Certain Entities to the Entity List, 14 April 2021

[21] Assaf Gilead, « Export controls strangling Israel's cyberattack industry », Globes, 25 April 2022

Such measures have proven effective, leading to the closure or takeover of major spyware companies. However, this impact cannot be separated from the characteristics of the implementing States. Israel, as a global hub for the cyber intrusion capabilities industry, and the United States, with its essential components and services, play crucial roles in these companies' operations. Furthermore, the evolving nature of the cyber intrusion ecosystem reveals that the targeted companies, or at least the individuals behind them, frequently reestablish operations under new names in favorable jurisdictions. This continual rebranding, coupled with the reliance on intermediaries, makes it challenging to halt industry growth through national entity-specific measures alone[22].

b) International export control frameworks

Adopting international export control frameworks that target specific items aims to ensure that supply companies cannot evade restrictions. The main instrument in this regard is the **Wassenaar Arrangement**, a multilateral agreement established in 1996 to control exports of conventional arms and dual-use goods and technologies, now formally endorsed by 41 States[23]. Since 2013, intrusion software and IP network communications surveillance systems have been added to the Wassenaar's list of controlled items following a proposal from France and the United Kingdom. Instead of directly controlling intrusion software, the arrangement has since then imposed export controls on the software, systems, or equipment that interact with it.[24]

This approach endeavors to broaden the geographical scope of export controls by fostering a common understanding among participating States. Nevertheless, the framework's effectiveness is hampered by its non-legally binding nature, divergences in interpretation and implementation, and the limited scope of participating States - that does not include certain key players. Universalization, actual implementation, and adaptation to market realities persist as significant hurdles.

ii. Targeted sanctions against identified firms

This second tier of sanctions was recently deployed by the U.S. Treasury Department's Office of Foreign Assets Control, that targeted in March 2024 two individuals and five entities associated with the European-based spyware consortium Intellexa, blocking any assets owned by the designated individuals and entities within the US - including properties and majority-owned entities[25]. Additionally, they prohibit all transactions by U.S. persons,

[22] See, in this regard: Winnona DeSombre, "Export Control is Not a Magic Bullet for Cyber Mercenaries", Lawfare, March 2023
[23] Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies, Founding Documents, WA-DOC (19) PUB 007
[24] See, in this regard : Collin Anderson, "Considerations on Wassenaar Arrangement control list additions for surveillance technologies", Access Now, 2015
[25] US Treasury Department, Press Release, Treasury Sanctions Members of the Intellexa Commercial Spyware Consortium, 5 March 2024

or those within or transiting through the United States, that involve any property or interests of designated or otherwise blocked persons. A month before, the US Department of State announced it would impose global visa restrictions on individuals involved in the misuse of commercial spyware, as well as those who facilitate or profit from this misuse.

While individual sanctions may deter other suppliers and mitigate certain evasion tactics employed by targeted companies, they remain constrained by the inherent limitations of a case-by-case approach within this rapidly evolving domain.

iii. Judicial and law enforcement means

a) Legal proceedings

Although legal actions have occasionally been initiated by public prosecutor's offices[26], the majority of recent proceedings stem from complaints filed by non-governmental stakeholders with a legally recognized interest in doing so. Judicial remedies for such cases are more accessible when host jurisdictions have extraterritorial laws relevant to cyber-related incidents, while enabling rulings to have international impact. Consequently, it is likely no coincidence that the majority of these cases were held before British and American courts.

The most notable procedures to date are those involving suppliers of software, services and devices whose security has been bypassed by intrusion capabilities. META's WhatsApp was the first to launch a battle against NSO in 2019, followed Apple in 2021, both on the grounds of violation of US laws on computer fraud. To date, all of NSO's attempts to avoid trial in the US have been dismissed by district courts, federal courts, and even the Supreme Court. It follows that the Israeli firm is not entitled to immunity from legal challenges as a foreign government's agent[27], nor should the complaint be processed in Israel instead of American courts[28]. Judges also ordered for Pegasus' code to be disclosed to WhatsApp, but upheld NSO's request to keep its customers and servers' architecture confidential[29].

However, individual victims have faced greater challenges pursuing legal avenues to date. In October 2023, a federal court in Virginia dismissed the complaint filed against NSO Group by the wife of murdered Saudi journalist Jamal Khashoggi, citing a lack of jurisdiction over

[26] See, in this regard: "German prosecutors charge four over violating trade act to sell spyware to Turkey", AP, 22 May 2023 ; "Polish district prosecutor launches proceedings in spyware-hack case", Euractiv, 19 January 2022

[27] Alexander Martin, "Supreme Court dismisses spyware company NSO Group's claim of immunity", The Record, 9 January 2023

[28] Jessica Lyons, "US judge rejects spyware slinger NSO's attempt to bin Apple lawsuit", The Register, 24 January 2024

[29] Stephanie Kirchgaessner, "Court orders maker of Pegasus spyware to hand over code to WhatsApp", The Guardian, 29 February 2024

the circumstances at stake[30].

b) Law enforcement operations

Recent years have seen a surge in transnational police operations targeting hacker-for-hire groups and underground marketplaces for vulnerabilities or intrusion capabilities. INTERPOL's Cybercrime Directorate has been at the forefront of this endeavor, achieving notable successes across regions with assistance from the private sector. In partnership with AFRIPOL, the 2022 Africa Cyber Surge Operation effectively shut down a darknet marketplace operated from Eritrea where hacking tools and cybercrime services were exchanged[31]. In the same vein, the 2023 16shop operation resulted in the arrest of some of the operators of a "phishing kit" platform[32]. Europol's European Cybercrime Centre boast similar successes, including the 2021 closure of the DarkMarket platform that was used in particular for selling malware[33].

iv. Name and shame

Strategies to publicly identify and denounce irresponsible actors and practices are available not only to States' authorities but also the broader stakeholder community. Civil society organizations have been instrumental in addressing commercial spyware misuse from the outset by effectively employing such strategies. Investigative reports from groups such as Citizen Lab, Amnesty Tech, and Access Now, coupled with statements and actions from key public authorities like the U.S. government and the European Parliament, have significantly increased public and political scrutiny over companies, individuals, and groups providing cyber intrusion capabilities globally. This heightened transparency has certainly driven changes in corporate and States' practices, especially motivated by fears of reputational, regulatory and diplomatic backlash.

B) Targeting the customer

i. Restriction on use

States' executive or legislative branches can impose bans or restrictions on the use of certain cyber intrusion tools or services by government agencies, including intelligence and law enforcement entities which are the main customers of cyber intrusion capabilities in

[30] Salvador Rizzo, "Judge dismisses Khashoggi widow's suit against spyware maker NSO Group", Washington Post, 26 October 2023
[31] INTERPOL Press Release, "Operation across Africa identifies cyber-criminals and at-risk online infrastructure", 25 November 2022
[32] INTERPOL, Press Release, "Notorious phishing platform shut down, arrests in international police operation", 8 August 2023
[33] Europol, Press Release, "DarkMarket: world's largest illegal dark web marketplace taken down", 12 January 2021

terms of transaction value. In March 2023, the U.S. White House thus released an **Executive Order** prohibiting federal agencies from using commercial spyware that poses threats to national security, counterintelligence efforts, enables human rights violations by foreign governments or targets American citizens[34]. Under this order, any agency seeking to acquire spyware technology must first conduct an assessment to determine compliance with the enumerated requirements. While widely welcomed, critics have highlighted the presence of some loopholes benefiting to the U.S. industry and allowing use in "extraordinary circumstances"[35].

ii. Towards a global moratorium?

Responding to longstanding advocacy from civil society organizations and calls from several United Nations Special Rapporteurs[1], Costa Rica took a notable political stance in 2022 by becoming the first State to pledge for a global moratorium on the sale, transfer, and use of spyware technology "until a regulatory framework that protects human rights is implemented"[36].

Achieving such far-reaching prohibition clashes with the growing reliance of States on surveillance technologies. A recent study from the Carnegie Endowment for International Peace revealed that 53 States across the world purchased spyware between 2011 and 2023, including 44 closed or electoral autocracies and 13 electoral or liberal democracies [38]. This certainly raises doubts about any wide-ranging political will to support this initiative, even with a narrow scope.

C) Targeting the “endpoints” - penetrated services and devices’ providers

i. Industry unilateral initiatives

Akin to any cybersecurity issue, technology suppliers face the dual challenge of keeping pace with rapid innovation while dedicating sufficient resources and time to diligently track and correct potential vulnerabilities in their commercially available products. Protecting

[34] Executive Order on Prohibition on Use by the United States Government of Commercial Spyware that Poses Risks to National Security, US White House, 27 March 2023
[35] Omer Benjakob, « Biden’s Spyware ‘Ban’ Proves: When America Hacks Your Phone, It’s Not Illegal », Haaretz, 29 March 2023
[36] UN OHCHR, Press Release, “UN expert calls for immediate moratorium on the sale, transfer and use of surveillance tools”, 25 June 2019 ; UN OHCHR, Press Release, “Spyware scandal: UN experts call for moratorium on sale of ‘life threatening’ surveillance tech”, 12 August 2021
[37] « Stop Pegasus: Costa Rica is the first country to call for a moratorium on spyware technology », Access Now, April 2022
[38] Steven Feldstein and Brian Kot, “Why Does the Global Spyware Industry Continue to Thrive? Trends, Explanations, and Responses”, Working Paper, Carnegie Endowment for International Peace, March 2023

their products and services from cyber-mercenary attacks can therefore be viewed as both a reputational safeguard and as a strategy to avoid stricter security-by-design regulations from public authorities.

This approach can primarily be deployed on a technical level. Apple's 2022 release of iPhone's Lockdown Mode is one such initiative, aimed at reducing the attack surface that could be exploited by mercenary spyware. The mode works by degrading the user experience and restricting smartphone functionality when activated. Some attempted hacks using Pegasus and Predator spyware have reportedly been blocked since the launch of this new protection layer[39].

A more widespread corporate strategy in the industry involves establishing coordinated vulnerability disclosure and treatment policies and launching bug bounty programs. These measures, which are not limited to intrusion capabilities and are rather a matter of corporate strategy than technical solutions per se, aim to enhance multi-stakeholder and ethical cooperation in identifying vulnerabilities - particularly between the research community and vendors. However, critics argue that these approaches are often company-specific, vary in effectiveness, do not always shield researchers from legal risks and can have unintended consequences. Bug bounty programs may thus inadvertently increase financial incentives for identifying vulnerabilities.

ii. Top-down regulations

When governments deem private sector initiatives inadequate, insufficient or too volatile in securing their products, they may adopt stricter "security by design" regulations for hardware and software suppliers, encompassing both the design phase and aftermarket. These frameworks often include vulnerability disclosure requirements. The **2023 European Union's Cyber Resilience Act** represents a weighty development in this regard by significantly enhancing security standards for manufacturers within the EU and globally. One of its most contentious provisions requires device manufacturers to report any vulnerabilities exploited in their products to Member States' Computer Security Incident Response Teams within 24 hours of discovery, regardless of patch availability[40]. Across the Atlantic, the Biden administration tends to follow a similar path with the **2021 Executive Order on cybersecurity** and the **2023 National Cybersecurity Strategy**, although these do not impose direct binding rules on industry players.

[39] Lorenzo Franceschi-Bicchierai, « Apple says it is not aware anyone using Lockdown Mode got hacked », TechCrunch, 13 December 2023
[40] See, in this regard : Harley Geiger & Alex Botting, "Vulnerability Management Under The Cyber Resilience Act", Center for Cybersecurity Policy and Law, January 2024

IV - The pressing call for enhanced global teamwork

Due to the manner in which ICT tools and services are exchanged globally and the dynamics underlying the perpetuation of the cyber intrusion ecosystem, efforts to address cyber-mercenaries must extend beyond like-minded coalitions. This approach should especially involve engaging major safe havens for the intrusion industry and significant State customers – directly or indirectly. As such, any endeavors in this realm faces the universalization challenge.

International cooperation and diplomatic efforts are crucially complementary to the measures described earlier. These actions may aim to raise awareness and foster a shared understanding of the various harms at stake while creating political incentives for initially reluctant States to take a stand on the issue. Ultimately, they could result in the establishment of binding or voluntary norms and, in turn, anchoring good practices to curb the proliferation and misuse of cyber intrusion capabilities.

As the cyber-mercenary phenomenon gained momentum on the international political stage in 2022, two significant initiatives emerged in response. The United States was the first to jump in, introducing an international variation to the 2023 Executive Order on commercial spyware in the context of the second Summit for Democracy. More of a declaration of intent and a conceptual framework than a set of concrete measures, the **Joint Statement on Efforts to Counter the Proliferation and Misuse of Commercial Spyware** was endorsed on this occasion by 11 States, followed in March 2024 by 6 others[41]. Despite an initial in-person meeting at the 3rd Summit for Democracy among participating states' representatives to exchange on best practices[42], there are no indications yet that this initiative will be further structured in a manner similar, for instance, to the US-led International Counter Ransomware Initiative.

More recently, France and the United Kingdom have decided to join forces to address the broader spectrum of commercially available cyber intrusion capabilities by launching the **Pall Mall Process**. At the occasion of the inaugural conference in London on February 2024, the Pall Mall Process' foundational declaration was endorsed not only by 25 States but also by several major tech companies, leading civil society organizations and experts. The conference's inclusiveness and the declaration's explicit recognition of the role of non-governmental actors are both encouraging steps toward a much-needed platform for multi-stakeholder cooperation at the global level. Such a cooperative scheme was already urged by the Working Group on Cyber Mercenaries of the Paris Call for Trust and Security in

[41] Joint Statement on Efforts to Counter the Proliferation and Misuse of Commercial Spyware, White House, 18 March 2024
[42] Background Press Call on U.S. Efforts to Counter Misuse of Commercial Spyware and the Third Summit for Democracy, White House, 18 March 2024

Cyberspace in November 2023[43]. However, it remains to be seen how much space will concretely be given to the industry and the civil society during the policy design phase. As highlighted by Jerome Barbier of the Paris Peace Forum, leveraging relevant multistakeholder platforms in this field could prevent the risk to "*keep navigating in comfortable processual waters instead of actually solving the problems the initial declaration has laid out and (…) reinventing the wheel in Pall Mall meetings*"[44]. Acting as the Paris Call's secretariat, the Forum will foster and coordinate civil society contributions to the Process.

One of the main stumbling blocks in the Process is likely to lie in the notion of "responsible" use of intrusion capabilities - as acknowledged by States in the foundational declaration. While this reference may reflect a realistic approach to the issue based on prevailing practices globally, failing to set a clear definition and safeguards could potentially provide additional justifications for uses harmful to a wide range of stakeholders. From this year onwards, the Paris Call's Working Group on Cyber Mercenaries will therefore delve into this concept by examining practical use cases, aiming to develop a definition aligned with individual rights, international security principles and cyberspace stability.

The cooperative effort towards greater accountability across the cyber mercenary supply chain could also be extended to the traditional multilateral fora. Since the Open-ended Working Group on Information and Communication Technologies (OEWG) is still the main process for discussing cybersecurity at the United Nations and can count on the participation of almost all Member States, acting within it might partly respond to the universalization challenge. In March 2024, the OEWG's Chair has paved the way for such endeavor. In his opening remarks to the 8th substantive session, Ambassador Burhan Gafoor encouraged Member States to consider the proposal put forward by the stakeholder community to establish a norm limiting the use of cyber-mercenaries[45]. Everything thus remains to be achieved, but the groundwork is set.

[43] Paris Call for Trust and Security in Cyberspace, "Taming the Cyber Mercenary Market: A Multistakeholder Blueprint Towards Increased Transparency and Cyber Stability", November 2023
[44] Louise Marie Hurel, Dr Gareth Mott, Jerome Barbier & al., « The Pall Mall Process on Cyber Intrusion Tools: Putting Words into Practice », Commentary, RUSI, February 2024
[45] Opening remarks by the chair of the open-ended working group on security of and in the use of icts 2021-2025, ambassador burhan gafoor, at the seventh substantive session of the oewg, 4 march 2023

ABOUT THE PARIS PEACE FORUM

Launched in 2018 to mark the commemoration of the end of the First World War, the Paris Peace Forum's (PPF) mission is to address global challenges such as climate change, poverty and inequality, security, and the rise of new technologies, by building new forms of collective action that complement the work of multilateral institutions and contribute to sustainable peace.

To respond to the growing need for global governance in an increasingly conflict-ridden world, in the context of growing distrust towards existing institutions, the PPF carries out political work throughout the year with diverse stakeholders from both North and South - states and international organizations, civil society including NGOs, foundations, academics and the private sector in particular - and organizes an annual event in November.

PARIS PEACE FORUM de PARIS sur la PAIX