STIMSON

EU CYBER DIRECT

DETERRENCE AND ACCOUNTABILITY

Taylor Grossman

STIMSON

EU CYBER DIRECT

DETERRENCE AND ACCOUNTABILITY

Taylor Grossman

July 2024

Suggested citation: Grossman, Taylor (2024) *Deterrence and Accountability* Policy brief, EU Cyber Direct, July 2024.

This publication has been produced in the context of the EU Cyber Direct – EU Cyber Diplomacy Initiative project with the financial assistance of the European Union. The contents of this document are the sole responsibility of the authors and can under no circumstances be regarded as reflecting the position of the European Union or any other institution.

Cover image credits: Verne Ho/Unsplash

Implementing organisations:

EU Institute for Security Studies
Carnegie Endowment for International Peace
Leiden University

Funded by the European Union

Contents

1. Introduction

In 1998, the Russian Federation's delegation to the United Nations made a proposal to the First Committee on Disarmament and International Security to take up the question of information and telecommunications technology (ICT).[1] Questions around ICT had been simmering at the UN for several years; this proposal, however, marked the first time the body formally considered the long-term role these new technologies would have on the global community. The proposal made two key framing choices that remain relevant, even 25 years later. First, the focus of the proposal was on both security and stability: The Russian Federation stated clearly that these "technologies and means may potentially be used for purposes incompatible with the objectives of ensuring international security and stability".[2]

Second, the statement made a clear linkage between ICT and other threat domains, particularly weapons of mass destruction (WMD): Russia was "concerned that new information technologies may be used to improve existing weapons of mass destruction or create new systems of such weapons".[3] Even in its earliest iterations, ICT had a relationship across domains – these new technologies posed not only a novel threat, but a potential amplification of existing threats. Later discussions of WMD threats were watered down and removed from the final proposal, but its existence in this first draft certainly draws attention to the linkages that already existed amongst policymakers about the cross-domain effects of cyber technologies.

Even now, these two elements of the 1998 proposal continue to have significant relevance for discussions of accountability and deterrence. First, the proposal notes the emphasis on creating and maintaining stability, which is quite a different task than deterring malicious or problematic behaviour. Russia was concerned not necessarily about individual cyber activities, but on the impact these operations could have on the global balance of power – on regime stability and regional alliance structures. And second, policymakers have long been aware of the potential relationship across domains; ICT is not just a new threat, but also a potential intervening variable in existing threat landscapes. Activities do not start and end in cyberspace but have important relationships and cross-interactions with other domains.

Broadly speaking, deterrence involves influencing adversary behaviour in a way that prevents them from engaging in malicious behaviour. There are several different forms of deterrence, including deterrence by denial. This paper uses a purposefully broad definition to engage with some current methodologies being employed in this space.

[1] Russian Federation, "Developments in the Field of Information and Telecommunications in the Context of International Security (Draft Resolution)," A/C.1/53/3 § (1998).

[2] Russian Federation.

[3] Russian Federation.

Traditional deterrence theory, such as nuclear deterrence, tends to rely on a few key attributes.[4] First, policymakers assume that there are a limited number of known (state) actors that currently possess (or have the capability to produce) the offensive weaponry in question. Nuclear weapons are incredibly costly to build and maintain at the ready; for smaller states and non-state actors, this cost is often considered prohibitive. Nuclear weapons also leave behind significant physical evidence – centrifuges and enrichment facilities, storage silos, command and control infrastructure – that makes it difficult for states to produce and house them in secrecy. Although states have tried in the past to develop clandestine programs, they have usually failed to keep their designs entirely secret.

Second, deterrence necessitates a degree of openness about capabilities: State actors want their adversaries to know (or at least fear) the retaliatory effects of their weapons programs. The weapons serve as a signal of a nation's power and capacity.[5] So, while states may have short-term incentives to hide their programmes when they are first starting out – to protect a nascent programme from being targeted and eliminated, or to avoid early reputational losses – they eventually need to go (somewhat) public with their new systems for their weaponry to have a deterrent effect on present or future adversaries.

Finally, and relatedly, (state) actors face clear and understood consequences for employing these technologies. In the nuclear area, there is a rich and complex normative environment, including the Nuclear Non-Proliferation Treaty (NPT), that helps hold states accountable. This system is far from perfect, and we still see significant violations. However, the NPT and other mechanisms have had decades to develop and have already weathered quite a few storms.[6]

These conditions cannot be assumed in cyberspace. First, the barriers to entry to create offensive cyber capabilities are not nearly as high as those that exist for other forms of weaponry, including and especially nuclear programmes. Although only a handful of state actors possess the tools, techniques, procedures, human capital, and overall infrastructure to launch large-scale, sophisticated offensive cyber campaigns, many smaller state and non-state actors have developed some form of cyber capabilities and can still cause damage. The nuclear club has remained small and exclusive, with only nine states believed to possess nuclear weapons; the "cyber club", if there is one, is certainly already much more expansive.[7]

---

[4] This is a very robust area of scholarly research. For just a few starting points, consider: Lawrence Freedman, The Evolution of Nuclear Strategy, (London: Palgrave Macmillan, 2003) ed. 3; Robert Jervis, "The Meaning of the Nuclear Revolution," *Political Science Quarterly*, Vol. 21, No.5, 1986: 689-703.

[5] This of course, can be done to greater or lesser success. See: Stephen Pifer, "Russia, nuclear threats, and nuclear signaling," Brookings, October 13, 2023, https://www.brookings.edu/articles/russia-nuclear-threats-and-nuclear-signaling/; Benjamin Hautecouverture, "War in Ukraine: Nuclear Signalling, Coercion, and Deterrence," Canadian Global Affairs Institute, January 2023, https://www.cgai.ca/war_in_ukraine_nuclear_signalling_coercion_and_deterrence.

[6] Alexander Bollfrass and Stephen Herzog, "The War in Ukraine and Global Nuclear Order," *Survival* 64, no. 4 (2022): 7–32.

[7] Max Smeets, No Shortcuts: Why States Struggle to Develop a Military Cyber-Force (London: C Hurst & Co Publishers Ltd, 2022).

The retaliatory consequences for engaging in malicious cyber behaviour are often ambiguous or unknown. We might not know for certain who perpetrated a particular cyberattack, or the state responsible might be able to mask its control over operations by using state proxies. While all-source attribution has certainly improved, political sensitivities and other realities continue to muddy the waters.

The norms surrounding what constitutes malicious, escalatory, or problematic cyber behaviour are also still nascent in many cases. While UN processes, such as the six Groups of Governmental Experts (GGEs) and Open-ended Working Groups (OEWGs), have produced a series of consensus reports and voluntary, non-binding norms, many raise more questions than they answer about defining notions like territory, due diligence, privacy, and critical infrastructure. And a broader question remains: What are we trying to deter in cyberspace?

States have turned to a few different methods to pursue accountability and deterrence in cyberspace. This paper will outline three below. Importantly, these are not mutually exclusive pathways – in fact, in most cases successful deterrence and accountability necessitate some combinations of all three methods. In cyberspace, we can't easily rely on traditional deterrence – we need robust and multifaceted ways of influencing actors to prevent them from engaging in malicious behaviour. There are three primary ways we can do so:

1) Shape the normative environment;
2) Punish bad behaviour; and
3) Deny bad behaviour.

2. Shape the normative environment

Norms can be a powerful method for creating an atmosphere of accountability and deterring the most troubling behaviours. The most prominent international processes have been the UNGGEs and OEWGs. Although begun with the first GGE in 2004, the process really grew in prominence in the 2010s, after the incident in Estonia and the exposed cyber operation of Stuxnet established a very public picture of the possibilities opened by ICTs. States became more interested in what was once a fairly niche topic: GGE membership numbers grew, and countries like the Netherlands took on prominent roles as norm entrepreneurs.

A broad network of regional and multi-stakeholder processes took shape, taking on important and complementary roles, including through ASEAN, the BRICS, the GCSC, the OAS, and others. Many of these processes serve distinct purposes and audiences; some are about articulating agreed upon norms of behaviour, while others focus on implementation. The OSCE Confidence Building Measures (CBMs), for example, serve primarily to build levels of trust between participating states to reduce risks posed by the use of ICTs in an escalatory or ambiguous situation.[8] The African Union, meanwhile, embarked on a robust convention on Cyber Security and Personal Data Protection (also known as the Malabo convention), a set of specific protections that AU member states can sign onto and thereby bind themselves to follow.[9] Thus far, 15 AU members have signed on, although some holdouts like South Africa have implemented many of the Malabo regulations through domestic processes. These two types of normbuilding exercises operate in distinct avenues, but ultimately serve to create more trusted outlets for regional states to operate together in cyberspace.

Regardless, it matters enormously who is defining the behaviours that we are talking about deterring. There are key tensions between activities that might destabilise and activities that are simply deemed malicious or inappropriate. This very tension is evident in the 11 voluntary and non-binding norms of responsible state behaviour in cyberspace that were recommended and eventually endorsed through the UNGGE process – the foundational norms set out.[10] There are eight positive norms and three limiting or restraining norms. Within each, many open questions remain in pinpointing what these norms really mean in practice. Norm three, for example, asks states to "prevent the misuse of ICTs in your territory". What qualifies as misuse? What does territory look like in a cyber incident? What constitutes appropriate due diligence? Norm five asks states to respect human rights and privacy, but these terms mean very different things in

[8] "Cyber/ICT Security, OSCE, accessed March 15, 2024, https://www.osce.org/secretariat/cyber-ict-security.

[9] African Union, "African Union Convention on Cyber Security and Personal Data Protection," Adopted June 27, 2014, https://au.int/sites/default/files/treaties/29560-treaty-0048_-_african_union_convention_on_cyber_security_and_personal_data_protection_e.pdf.

[10] Bart Hogeveen, "The UN norms of responsible state behavior in cyberspace: Guidance on implementation for Member States of ASEAN," Australian Strategic Policy Centre," March 2022, https://ad-aspi.s3.ap-southeast-2.amazonaws.com/2022-03/The%20UN%20norms%20of%20responsible%20state%20behaviour%20in%20cyberspace.pdf?VersionId=pwQNsEIHhDSAx_7gJ4XXSGSupVOpvMJi.

different cultures and environments. Not to mention, of course, that with robust enough surveillance of citizens, a state could reduce misuse of ICT – a tradeoff unthinkable for most democracies, but often appealing to an autocracy. Within these 11 norms are different costs and incentives, as well as key questions of state capacity (and interest) in achieving the ends outlined. These questions about stability trickle down into regional dialogues as well. Some regional groups can agree upon more robust normative frameworks because they have a clearer sense of what stable, malicious, or otherwise problematic activity might look like.

3. Punish Bad Behaviour

Retaliatory consequences for malicious cyber behaviour are often vague or unknown. This method of accountability attempts to solve this problem by articulating clear response mechanisms to potential adversaries. If actors understand the costs of launching specific types of cyberattacks, then perhaps they will think twice before doing so.

The European Union has adopted this approach in part through its cyber diplomacy toolkit, which has the express intention of setting transparent responses for malicious cyber behaviour. The toolkit responses are predicated on the level of confidence in the attribution of the responsible party, as well as the level of coordination necessary across member states to effectively implement the action.[11] In the case of cyber sanctions, for example, the EU announced in 2019 that it would adopt a horizontal regime to be used to prevent, deter, and/or respond to malicious cyberattacks perpetrated by individuals or non-state organisations. These sanctions include travel bans and financial asset freezes, and to date the EU has adopted sanctions against 12 individuals or entities. Other activities require much less coordination across EU institutions and members, as well as lower levels of attribution certainty. However, the idea here is to express with as much clarity as possible the consequences that malicious actors could face in response to a cyberattack that threatens the EU or one of its Member or partner states. The process has not been perfect, and there are still many scholars and practitioners who have vocally criticised its application. Yet, the approach is an interesting advancement in considering multifaceted deterrent strategies.

Other states have engaged in punishment strategies either unilaterally or multilaterally through indictments and sanctions. Over the last 10 years, the United States Government has publicly attributed dozens of cyber operations to foreign, state-affiliated actors. Two arms of the US Government – the Department of the Treasury and the Department of Justice – have launched both punishment and publicity policies. The Treasury Department has levied sanctions against individual actors, freezing financial assets, while the Justice Department has issued indictments that often restrict accused actors' mobility.[12] Sanctions and indictments bring a certain notoriety to individual actors, which

[11] The original EU Cyber Diplomacy Toolbox was launched in 2017. The revised guidance for implementation, launched in 2023, paid special attention to particularly troubling activities like cyber-espionage and foreign information manipulation and interference. "Revised Implementation Guidelines, Cyber Diplomacy Toolbox," Cyber-Diplomacy-Toolbox, June 8, 2023, https://www.cyber-diplomacy-toolbox.com/Revised_Implementing_Guidelines_Cyber_Diplomacy_Toolbox.html.

[12] U.S. Department of the Treasury, "Office of Foreign Assets Control: Cyber-Related Sanctions," accessed March 15, 2024, https://ofac.treasury.gov/sanctions-programs-and-country-information/sanctions-related-to-significant-malicious-cyber-enabled-activities. The Department of Justice operates many different types of programs here. To cite just one, for example, see: The U.S. Department of Justice, "Seven Hackers Associated with Chinese Government Charged with Computer Intrusions Targeting Perceived Critics of China and U.S. Businesses and Politicians," March 25, 2024, https://www.justice.gov/opa/pr/seven-hackers-associated-chinese-government-charged-computer-intrusions-targeting-perceived. They also have a fairly robust "Rewards for Justice" program offering funding in exchange for tips on wanted cyber criminals.

can inhibit future job prospects, make it difficult to travel, and affect their current and future financial security. The results of these approaches are unclear; in some cases, cyber criminals or hackers may crave more publicity for their cause. In the recent LockBit takedown, the group claimed to reform with renewed zeal – although it certainly has lost credibility within the Ransomware-as-a-Service space, and the United Kingdom and the US are confident that it has been disrupted for the foreseeable future.[13]

Like-minded states cannot target every perpetrator of illegal activities, and so many still get away with illicit behaviour Yet the indictments and sanctions do provide a powerful international reminder that punishment mechanisms exist and are being pursued by states on a regular basis – and can help reduce the profitability and anonymity of enterprises.

[13] Andy Greenberg, "Ransomware Groups Are Bouncing Back Faster From Law Enforcement Busts," Wired, February 27, 2024, https://www.wired.com/story/blackcat-ransomware-disruptions-comebacks/; James Coker, "Ransomware Incidents Hit Record High, But Law Enforcement Takedowns Slow Growth," Infosecurity Magazine, January 30, 2024, https://www.infosecurity-magazine.com/news/ransomware-incidients-high-law/#:~:text=However%2C%20law%20 enforcement%20 takedowns%20are,2022%20and%203048%20in%202021; Brian Krebs, "Feds Seize LockBit Ransomware Websites, Offer Decryption Tools, Troll Affiliates," Krebs on Security (blog), February 20, 2024, https://krebsonsecurity.com/2024/02/feds-seize-lockbit-ransomware-websites-offer-decryption-tools-troll-affiliates/?utm_source=pocket_saves.

4. Deny Bad Behaviour

Finally, several states have pursued an active defensive approach focused on denying malicious activity before it has the chance to even arise. The US has been the most famous proponent of this method through its "layered cyber deterrence" strategy that involves significant engagement with nefarious actors. The US has asserted that currently, malicious actors feel "undeterred, if not emboldened" because they do not face enough resistance and threat of punishment from major actors like the US.[14] Now, cybersecurity necessitates anticipating the exploitation of vulnerabilities before they occur and leveraging vulnerabilities among adversaries that already exist. Here, security rests on a logic of "initiative persistence". In other words, it is no longer enough to wait for the enemy to strike – the US has instead advocated for actively pursuing the enemy and denying it any benefits or success before it can launch attacks.

The UK has also endorsed a version of this approach. In April 2023, the UK released a white paper on its National Cyber Force, a joint effort between the intelligence services and the military, which calls for a similar kind of initiative persistence to properly defend the country and its interests. While the UK and the US have been very clear that they do not want to articulate thresholds of activity or transparent response options – lest they give their adversaries a playbook to follow – the UK has opted for more operational transparency, granting the public a clear view of the internal machinations involved in launching cyber operations. The UK is invested in becoming what it calls a "responsible cyber power", which includes embarking on operations that are accountable, precise, and calibrated.[15]

Deterrence by denial has always been a tricky method, as in most cases it requires considerable resources. The US has invested heavily in initiative persistence and layered cyber deterrence; the UK, too, is continuing to step up material contributions to its operational infrastructure. Other countries are likely to follow suit in the coming years as well, albeit on a smaller scale.

[14] "United States of America Cyberspace Solarium Commission," March 2020.

[15] "The National Cyber Force: Responsible Cyber Power in Practice" (National Cyber Force, April 4, 2023), https://www.gov.uk/government/publications/responsible-cyber-power-in-practice.

5. Where Do We Go Now

These are three potential methods for influencing adversary behaviour. The best approaches combine several – indeed, often all three – to achieve some kind of deterrence of the worst behaviours in cyberspace.

What is important to remember, however, is that political relationships and activities do not start nor end in cyberspace. Cyber activity is part of a broader spectrum of behaviour, and actions taken in cyberspace are not divorced from those happening in other domains. Particularly when we home in on deterrence strategies for cyberspace, we focus too myopically on cyber activities without remembering how they fit much more holistically into other realms of state and non-state behaviours. Where are there potential areas for alignment? For conflict and tension?

States are often too quiet about what they are doing, and too concerned that they need to stick to one route – shape, punish, or deny – to get a consistent message to their adversaries. However, states should be using every bit of power in their arsenal to get the attention of the public – both malicious actors and the general populace – to remind them that cyberspace is not the Wild West, but a world of rules and governance, even if some of these rules are still being written. We need more disruption of malicious activity (early and often) and we need to couple it with sharp rebukes that double down on the important norms being advocated for at the UN and other regional and multilateral platforms. We can and should deny bad behaviour, but we can't rely on a single prong of action to get us where we want – not now, and certainly not in the future.

While cyberspace does have its own unique elements, it can complement and intersect what is happening elsewhere. As the earliest UN proposals on ICT remind us, cyber does not, cannot, and should not exist within a vacuum.

About the author

Taylor Grossman Deputy Director for Digital Security at the Institute for Security and Technology (IST), where she works on the Ransomware Task Force and other ongoing projects.

Previously, she was a senior researcher in the Cyberdefence Project at the Center for Security Studies (CSS) at ETH Zurich. There, she served as a policy consultant for the Swiss government. She also conducted independent research on cyber rapid response teams and emergency management, cyber norm development, and other policy issues. Before joining CSS, she worked as a senior research analyst and project manager in the Cyber Policy Initiative at the Carnegie Endowment for International Peace, where her work focused on capacity-building and financial inclusion.

She also serves as a senior editor at Binding Hook, a media organization that publishes articles on technology and security. She holds an MPhil in International Relations from the University of Oxford and a B.A. in Political Science from Stanford University.

About EU Cyber Direct

EU Cyber Direct – EU Cyber Diplomacy Initiative supports the European Union's cyber diplomacy and international digital engagements in order to strengthen rules-based order in cyberspace and build cyber resilient societies. To that aim, we conduct research, support capacity building in partner countries, and promote multistakeholder cooperation. Through research and events, EU Cyber Direct regularly engages in the discussions about the future of international cooperation to fight cybercrime and strengthen criminal justice systems globally.