RESPONSIBLE BEHAVIOUR IN CYBERSPACE

Global narratives and practice

Edited by
François Delerue, Arun Sukumar and Dennis Broeders

print

ISBN 978-92-9462-221-1
CATALOGUE NUMBER QN-04-23-460-EN-C
DOI 10.2815/643871

online

ISBN 978-92-9462-220-4
CATALOGUE NUMBER QN-04-23-460-EN-N
DOI 10.2815/728569

Printed in Belgium by Bietlot.
Luxembourg: Publications Office of the European Union, 2023.
Cover image credit: World map in the style of Hokusai, created by Bing Image Creator.

EU CYBER DIRECT

RESPONSIBLE BEHAVIOUR IN CYBERSPACE

Global narratives and practice

Edited by
François Delerue, Arun Sukumar and Dennis Broeders

Acknowledgements

This volume is based on papers presented and discussed at the conference **Closing the Gap 2022 | Responsibility in Cyberspace: Narratives and Practice**, organized on 8 and 9 June 2022 at the Egmont Palace in Brussels, Belgium, by Leiden University, as part of the EU Cyber Direct project. The conference brought together paper authors and representatives from numerous research institutions and civil society organisations around the world. In addition to the papers presented, there were roundtables on global perspectives on EU cyber diplomacy and scholars and experts' experience on working with policy makers, and opening remarks by Mathieu Michel, Secretary of State for Digitization, in charge of Administrative Simplification, Privacy and Buildings Administration of Belgium.

We would like to thank the **Belgian Federal Public Service Foreign Affairs** for their support in organizing this conference, in particular Pierre Gillon and David Van Lierde of the Directorate for International Governance (M4).

The abstracts and papers presented at this conference and in this edited volume were selected and reviewed with the help of our **Selection Committee**. We would like to thank them for their efforts:

- **Luca Belli**, Fundação Getulio Vargas and CyberBRICS, Brazil
- **Dennis Broeders**, Leiden University & The Hague Program on International Cyber Security, the Netherlands
- **Joe Burton**, University of Notthingham & CYDIPLO, United Kingdom
- **Enrico Calandro**, Research ICT Africa, South Africa
- **Gunjan Chawla**, Centre for Communication Governance at National Law University in Delhi, India
- **Lu Chuanying**, Research Center for the International Governance of Cyberspace, Shanghai Institutes for International Studies, China
- **Frédérick Douzet**, Paris 8 University & GEODE, France
- **Maria Lorena Florez**, Universidad de Los Andes, Colombia
- **Aude Géry**, GEODE, France
- **Joyce Hakmeh**, Chatham House, United Kingdom
- **Ivar Hartmann**, Insper Learning Institution, Brazil
- **Bart Hogeveen**, Australian Strategic Policy Institute (ASPI), Australia
- **Jon Lindsay**, Georgia Tech, United States
- **Alejandro Pisanty**, National University of Mexico-UNAM, Mexico

> **Vera Rusinova**
> **Max Smeets**, ETH Zürich and European Cyber Conflict Research Initiative (ECCRI), Switzerland
> **Motohiro Tsuchiya**, KEIO University, Japan

We would also like to thank our partners within the EU Cyber Direct project for their support and contributions to this conference, in particular Patryk Pawlak, Agnese Olmati and Raluca Csernatoni.

The **European Cyber Diplomacy (EU Cyber Direct)** initiative is funded by the European Union. The implementing partners of the project are: **European Union Institute for Security Studies (EUISS)**, **Institute of Security and Global Affairs at Leiden University** and **Carnegie Europe**.

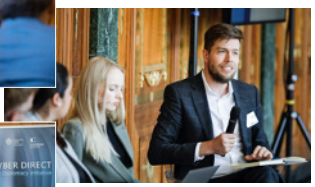
euiss
European Union Institute for Security Studies

Universiteit Leiden
Institute of Security and Global Affairs

CARNEGIE EUROPE

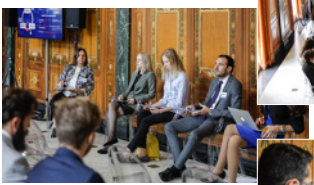


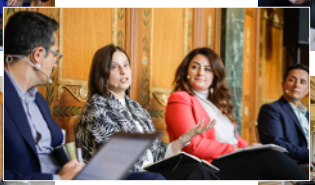
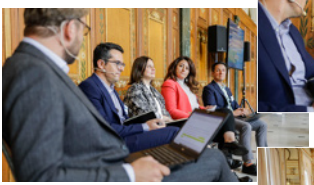


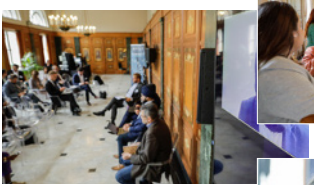


EU CYBER DIRECT



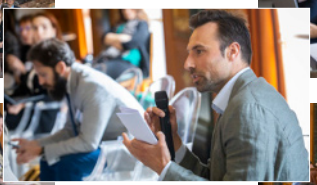


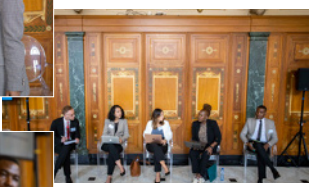
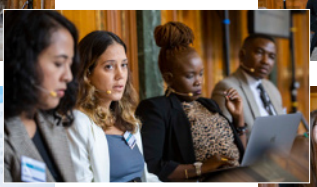





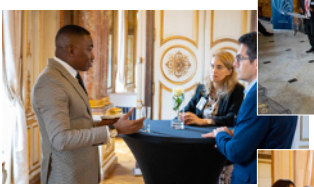






EU CYBER DIRECT






CYBER DIRECT

Contents

Introduction **7**
Responsible behaviour in cyberspace: Global narratives and practice
Arun Sukumar, Dennis Broeders and François Delerue
Responsible state behaviour in cyberspace *7*
Closing the Gap *9*
Overview of the book *10*

REGIONAL AND INTERNATIONAL COOPERATION

1 **The effectiveness of ASEAN regional efforts on cybersecurity** **18**
Monica Nila Sari
Introduction *18*
ASEAN's efforts on cybersecurity *20*
Case study of cyber-attack in Indonesia *23*
ASEAN's regional approach analysis *26*
Conclusion *32*

2 **Online content regulation in the BRICS countries** **34**
A cybersecurity approach to responsible social media platforms
Luca Belli, Yasmin Curzi de Mendonça and Walter B. Gaspar
Introduction *34*
The BRICS and their cybersecurity landscape *36*
Recent developments in the BRICS countries *40*
Conclusion: Choosing between a sledgehammer and a scalpel to regulate content *58*

3 **'We are not quite there yet'** **60**
The Latin-American narrative regarding cyber-norms development
Maria Pilar Llorens
Introduction *60*
Cyberspace norms, narratives and the Global North *63*
The Global South and cyberspace: Latin American experience and narrative *66*
Is there room for a Latin American narrative? *69*
Conclusion *73*

4 **The legal framework for cybercrime accountability in the Western Balkans countries as a turning point for EU integration** **75**
Andreja Mihailović
Overview *75*
Introduction *76*
The importance of a national cybercrime, cybersecurity and cyber-defence framework *78*
State of play in the Western Balkans *80*
The WB's integration roadmap *86*
Conclusion *89*

5 **A looking glass on South–South cooperation to strengthen responsibility in cyberspace** **92**
Moliehi Makumane and Enrico Calandro
Introduction *92*
Methodological approach *94*
Defining the Global South *95*
Defining responsibility in cyberspace and in South–South cooperation *97*
The notion of responsibility in regional grouping statements *99*
Analysis of interviews *101*
Policy considerations and recommendations for developing a South–South cooperation cyber dialogue *108*
Appendix: Interview questions *111*

NATIONAL PERSPECTIVES

6 **Small state, loud voice** **116**
Singapore's regional leadership for norms on responsible state behaviour in cyberspace
Mabda Haerunnisa Fajrilla Sidiq
Introduction *116*
Regional leadership (and the lack thereof) in ASEAN *119*
Norms on responsible state behaviour and Singapore's foreign policy *121*
Gaining legitimacy in ASEAN: fitting the frames into the locale *124*
Conclusion *128*

7 **What does Nigeria's national identity server downtime suggest about accountability and cyber norms in local CERTs? An exploratory study** **129**
Babatunde Okunoye
Overview *129*
The context: Nigeria's national identity database as critical national infrastructure and critical information infrastructure *131*
The cyber incident *135*
Cyber norms and cyberculture as key components of national cyber capacity/maturity *138*
Analysis and conclusion *141*

8 **The role of state–civil society relations in shaping cyber norms in South Korea** **144**
Sofiya Sayankina
Introduction *144*
The establishment of the online democratic sphere in South Korea *146*
The individual and the state in South Korea's cyberspace *150*
The state's policies shaping the norm-building process in South Korea's digital public sphere *154*

CAPACITY BUILDING AND PUBLIC-PRIVATE PARTNERSHIPS

9 **Closing the cyber-capacity gap in digital financial inclusion A critical analysis of prevailing narratives and approaches** **162**
Nanjira Sambuli and Aditi Bawa
Introduction *162*
Cyber capacity and the financial system: insights from the FinCyber Strategy *165*
Towards effective and sustainable cyber capacity-building and digital financial inclusion: tensions and emerging questions *168*
What will count as successful cyber capacity-building? *170*
Conclusion *175*

10 Shaping platform governance in Central Asia **177**
Challenges and opportunities for human rights defenders and journalists
Pavlina Pavlova
Introduction *177*
Internet freedom and the social media landscape in Central Asia *179*
Information control: legislation and practice *181*
The impact of Russia's war in Ukraine *187*
Responses by social media platforms *190*
Recommendations *193*
Conclusion *196*

INTERNATIONAL LAW AND HUMAN RIGHTS PERSPECTIVES

11 Pulling the strings in cyberspace **200**
Legal attribution of cyber operations based on state control
Evgeni Moyakine
Introduction *201*
State responsibility *203*
Control theories in the age of cyber *206*
The Stuxnet incident *210*
Conclusion *217*

12 Is cybersecurity the sole responsibility of states? **219**
The concept of 'active defence' and the role of non-state actors in responsible state behaviour in cyberspace
Jaime Bello
What is 'active defence' and why do some private actors promote its use? *219*
How does the active defence of private organisations affect the responsible behaviour of states in the cyber world? *222*
Balance of interests: prioritising the defence of national interests vs responsible behaviour in cyberspace *229*
A third way? Outsourcing the cyber response *232*
Conclusion *237*

13 Humanitarian organisations under cyber-attack **238**
Emerging threats and humanitarian actors' responsibilities under international human rights law
Francesca Romana Partipilo and Marta Stroppa
Introduction *238*
The 'humanitarian cyberspace' *240*
Cyber-threats against humanitarian organisations and their detrimental impact on vulnerable people's rights *241*
Humanitarian organisations' protection under international law *243*
Humanitarian organisations' responsibilities in the field of data protection *246*
A human-rights-based approach to cybersecurity in humanitarian emergencies *249*
Some recommendations for a 'cyber-secure' humanitarian action *251*
Conclusion *257*

14 A responsibility to improve **258**
How global cybercrime cooperation frameworks must better safeguard human rights and protect the humans of cybersecurity
Raman Jit Singh Chima
Introduction *258*
International cybercrime legal harmonisation: an opportunity for reform or focus on the responsibility to prevent further harm? *264*
Excessively broad cybercrime legal provisions impact cybersecurity research and result in more instability *268*
Ensuring global cyber-coordination helps to further respect for privacy and protected human rights *273*
Conclusion *276*

ANNEX

About the contributors **280**

Introduction

Responsible behaviour in cyberspace: Global narratives and practice

ARUN SUKUMAR, DENNIS BROEDERS AND FRANÇOIS DELERUE

Responsible state behaviour in cyberspace

The global debate on responsible state behaviour in cyberspace has been ongoing since the Russian Federation submitted a proposal leading to the adoption of the first United Nations General Assembly (UNGA) Resolution on the topic in 1998. Since then, developments in the digital sphere have sped up considerably. The significance of the internet for economic activities, communication, government-citizen relations, and interstate relations – including conflict – has increased exponentially. For most countries, the functioning of society is hard to imagine without the underlying infrastructures of the internet and the world wide web. In countries that have come online more recently the face of the internet is usually mobile, but the transformation is no less profound. Large technology companies – including the so-called Big Tech companies – burst onto the global scene not only by providing goods and services, but more significantly by redefining how people connect and form economic, social and political ties. Some of these companies are built on business models that deviated from the classical 'selling goods and services to customers' and operate on the principle of 'selling customers to advertising companies' by

providing goods and services that generate personal data. Governments, and regional organisations like the European Union (EU), have been trying to shape this socio-economic space through competition policy and privacy and data protection policies. Crime also found its way to the digital domain and, especially since the unholy combination between ransomware and cryptocurrency took out the need for money mules, cybercrime has scaled up and became footloose. Fighting cybercrime has proven to be difficult as the phenomenon is profoundly transnational and law enforcement has been struggling to keep up and cooperate across borders.

States have also shaped the internet in a more direct way. Even though most of the political and academic debate about 'cyberwar' has died down, modern day armed conflict has taken on a distinct digital component. Moreover, low-level adversarial activities, below the threshold of the use of force, seem to have become a permanent feature of international relations. Intelligence operations, subversion, and sabotage have all taken a digital shape, leading experts and states to grapple with the exact nature of such operations and how to respond to them. The potential of the internet for both destructive cyber operations as well as subversive information operations also put the phenomenon on the radar of the international community.

Since 1998, digital affairs and adversarial uses of cyberspace have landed on the diplomatic agenda. Between 2004 and 2017, the First Committee of the UNGA (Disarmament and International Security) has convened five so-called Groups of Governmental Experts (GGE) to discuss the risks from the digital revolution to peace and society and to propose elements of a framework for responsible state behaviour in cyberspace. The consensus reports adopted by the GGE in 2010, 2013 and 2015 laid the groundwork for that framework by getting agreement on the threats states face, the applicability of international law to cyberspace as it does to offline activities, and by adopting eleven non-binding norms for responsible state behaviour. In 2018, the plot thickened. The UN General Assembly adopted two concurrent resolutions that started both a new GGE – promoted by the so-called 'like-minded states' – and a new process called an Open Ended Working Group (OEWG), promoted by the Russian Federation, China and other states. With almost fully overlapping mandates, but a very different membership, both processes produced consensus reports in the spring of 2021 that by and large confirmed the '*acquis*' of the previous UN GGE reports, reaffirming the previously adopted framework for responsible state behaviour. The OEWG was open to all UN Member States, conducted its deliberations in the open and for the first time, left a paper trail of state submissions and opinions. The GGE, in contrast, was and is a more traditional closed-door process with a limited membership

(maximum 25 member states) and only spoke through its consensus reports. If there was no consensus, the world was none the wiser as to what had been discussed.

In 2023, the debate about responsible state behaviour in cyberspace at the UN has widened even further. The second Open Ended Working Group (2021-2025) that was voted into existence even before the first OEWG wrapped up its report is currently in its second year of deliberations. The like-minded states preferred the new vehicle of a Programme of Action (PoA), that would focus more on implementation of the *acquis*. On 7 December 2022, the UNGA adopted a Resolution welcoming the proposal for the creation of the PoA as a permanent, inclusive, action-oriented program. In the Third Committee of the UN, meanwhile, the open-ended Cybercrime Ad Hoc Committee has started its work on drafting a new cybercrime convention. As states have very different definitions of what constitutes a (cyber)crime, these are tense negotiations. The fact that all these negotiations are held in times when geopolitical and geo-economic tensions are rising, influences the negotiations and the trust levels needed to get to consensus.

Closing the Gap

Students of international relations are no strangers to power asymmetries, but the study of diplomatic processes suffers from something similar: information asymmetries. We tend to know more about the positions and interests of large and powerful states, as well as those of states that are more similar to ourselves. The self-labelled group of 'like-minded states' is already a good illustration of that. For various reasons, it is often harder to get in-depth knowledge of the interests and positions of states that are neither allied with their own group nor in direct opposition to it. Between the poles of the debate – where many states find themselves – the vision gets blurry. Lack of diplomatic capacity, on both ends, language barriers and sometimes a lack of crystalized policy positions – over a wide variety of cyber-related issues – makes it hard to know, digest, and discuss commonalities and differences.

Especially at this moment in time when there are multiple processes at play at the UN level, wide-ranging discussions at the regional level in organizations like the EU, the African Union, Association of Southeast Asian Nations (ASEAN) and the Organization of American States (OAS), and in non-regional collectives like the BRICS (Brazil, Russia, India, China and South Africa) and the Non-Aligned

Movement, the need for understanding different perspectives, interests and strategies is greater than ever. One of the best ways to create a deeper understanding is by inviting analysts from different parts of the world to convene and discuss and put their thoughts on paper.

In June 2022, a diverse group of scholars and experts gathered in Brussels to discuss various aspects of digitization and cybersecurity around the theme of 'responsible behaviour in cyberspace'. This conference was organized by the EU Cyber Direct program as part of their *Closing the Gap* conference series. The participants spent two days debating various aspects of and perspectives on international cyber security and how to define and organize responsible behaviour of states and other actors. Those discussions are reflected in the chapters that are included in this collection.

Overview of the book

The chapters in this volume are divided into four sections focusing respectively on regional and international cooperation, national perspectives, capacity building and private-public partnerships, and finally on questions of international law and human rights. Each section thematically addresses one or more dimension(s) of 'responsible state behaviour' in cyberspace, and many papers highlight how domestic and international considerations shape the practices of states and non-state actors with respect to the said issue.

The section on regional and international cooperation includes analyses of cybersecurity discussions within the ASEAN, BRICS, Latin America, Western Balkans, and the Global South, broadly defined. **Monica Nila Sari** notes that the ASEAN has attempted to step up to cybersecurity challenges posed by rapid digitalization of its economies following the COVID-19 pandemic through institutional and policy frameworks. Building on the historic efforts of the ASEAN Regional Forum to adopt and implement Confidence-Building Measures (CBMs) in cyberspace, she argues, ASEAN should move towards the creation of a legally binding instrument that establishes 'cybersecurity baselines and compliance mechanisms' for the region. Even as this goal is pursued, ASEAN should enhance existing capacity-building initiatives and in particular, technical cooperation measures. **Luca Belli**, **Yasmin Curzi**, and **Walter Gaspar** survey the changing landscape of cybersecurity policies in BRICS countries. Legislative trends certainly point towards a ratcheting up of intermediary obligations and attempts to increase 'state sovereignty' over territorial digital infrastructure and networks.

Even as these regulations get increasingly sophisticated and cover emerging concerns such as algorithmic accountability, they note, their implementation and review mechanisms remain largely administrative in nature, increasing risks of government overreach as well as non-transparent functioning. **Maria Pilar Llorens**' chapter poses the question whether Latin American countries can find their own voice and 'narrative' regarding the application of international law to state behaviour in cyberspace. Despite exhortations by regional powers of the need to promote a distinct Latin American perspective on the subject, they have not systematically identified their own concerns and strategic motivations to advance international cyber law, notes Llorens. Capacity building initiatives and the articulation of national positions on international law are still largely driven by narratives on cyber operations and insecurities developed by academics and practitioners in the Global North, she concludes. **Andreja Mihailovic** documents recent efforts by Western Balkan states to articulate cybercrime policies and strategies and argues that the 'strengthening of the legal environment for cybercrime' is both an opportunity for regional development but also for its greater integration with the European Union. She prescribes seven measures that Western Balkan states can undertake to harmonize their policies and strengthen cyber defence capabilities, including the setting up of a common cyber risk registry and articulating special protections for regionally important sectors. **Moliehi Makumane** and **Enrico Calandro** take a broad view of South-South cooperation, analysing the multiple lenses from which developing countries view the concept of state 'responsibility' in cyberspace. Drawing *inter alia* on interviews from policy experts from the Global South, they find that developing countries foreground the need for the international community to responsibly preserve gains from ICT for development, and not use cyberspace for overt politico-military goals. At the same time, there is increasing acknowledgement that the Global South should also be seen as responsible participants in cybersecurity regime-building, and build effective domestic infrastructure that secures their own citizens, but also global digital resources from the effects of malicious cyber operations. Given the diversity of topics and entities highlighted in this section, it would be difficult to isolate a specific trend among regional or multilateral coalitions with respect to their articulation of responsible state behaviour. However, it is clear that states across the geopolitical and ideological divide have sought to enhance their 'cyber sovereignty' through tougher domestic cybersecurity, cybercrime, and content regulation laws, and through calls at the global level seeking greater control over digital networks and infrastructure. The adoption of such measures would imply greater responsibility of states for transboundary effects of cyber operations emanating from their territory, including influence operations, but the regional and

international cooperation to realize such measures also raises legitimate concerns about the implementation of human rights online.

The national case studies pertain to Singapore, Nigeria, and South Korea. **Mabda Haerunissa Fajrilla Sidiq** shines light on Singapore's regional 'norm entrepreneurship' at a critical moment in the country's efforts at cyber diplomacy – Singapore is also the chair of the 2021-2025 OEWG, which, at the time of writing, is into its third year of deliberations. Sidiq notes that Singapore is attempting to perform a delicate balancing act: on the one hand, it has tried to work with pre-existing ASEAN institutions to promote CBMs and adoption of GGE norms, and on the other, encouraged ASEAN countries to play a more prominent role in global discussions on cybersecurity, including at the GGE. Singapore's challenge will not only be to elicit but also sustain interest in cyber diplomacy efforts from various ASEAN countries. **Babatunde Okunoye** highlights, through the working of Nigeria's national digital identity program, the need for a cohesive legal, institutional, and 'cybercultural' approach to crisis management that emphasizes redundancy of critical infrastructure service networks, rapid response, and greater synergy on cybersecurity cooperation with the private sector. In February 2022, the national identity platform suffered an extended disruption, resulting in severe hardship for users and economic costs for businesses in multiple sectors that relied on its authentication service. The cyber incident demonstrates the dilemma facing many developing countries: the operation of national digital platforms that offer civic services at scale often requires some element of centralized control and management, but this very feature also opens them up to cybersecurity vulnerabilities and concerns about their resilience in the face of disruptive attacks. **Sofiya Sayankina**'s chapter on the emergence of the 'digital public sphere' in South Korea highlights the evolving dynamic between state and civil society on cybersecurity policymaking. Thanks in part to the historic role of its NGOs and ordinary users in facilitating ubiquitous and affordable internet access, South Korea has developed multistakeholder models of cooperation and governance in the digital domain. However, the growing presence and activism of civil society groups and coalitions presented a challenge to the government in that they sometimes sought to address cybercrime and insecurity by themselves, including in ways that contravened the law. To manage these constituents, South Korea has created channels for the public to offer feedback and opinion on policy directions, while at the same time developing strict standards on identity verification and content monitoring, Sayankina notes.

Since the publication of the landmark 2014-15 UN GGE report, capacity building has been considered one of the four pillars of the 'framework of responsible state behaviour,' alongside voluntary norms, international law, and

confidence-building measures. **Nanjira Sambuli** and **Aditi Bawa** critically analyse the slate of 'supply-driven capacity building' initiatives that populate the field of cybersecurity governance. They argue, using the example of digital financial technologies, that training end users and private companies is just as important as skilling regulators and technical operators, which currently draws the lion's share of capacity-building resources. Sambuli and Bawa point particularly to the lack of sustained initiatives around behavioural aspects of cybersecurity and the relative neglect by development assistance groups of the role that gender plays in shaping roles, identities, and expectations of professionals involved in cybersecurity. **Pavlina Pavlova** points to another aspect of capacity building that is sometimes overlooked in practice: support and training of journalists, civil society organizations, and human rights defenders by social media platforms and technology companies on cybersecurity and data protection, including against state-backed surveillance or harassment. Pavlova highlights the trend of proliferating and increasingly restrictive laws and executive policies in Central Asia on online speech and social media governance. She notes that the war in Ukraine has further exacerbated attempts by states to control narratives critical of Russia's actions in particular. Just as the international community should support and incentivize Central Asian states to comply with human rights commitments, private initiatives at building cybersecurity capacity among various societal actors is also necessary, she concludes.

And finally, questions around the application of international law, especially human rights, to cyberspace are among the most important but also contentious issues facing states today. Examining the challenges involved in attributing cyber operations to states, **Evgeni Moyakine** makes the case for adopting the 'overall control' test under the customary law of state responsibility. The widespread and covert use of hacker collectives and organized private groups by states to execute cyber operations that are potentially internationally wrongful in character poses a serious challenge for legal attribution. Adopting a lower threshold of evidentiary standards required to connect the actions of non-state groups to states will go a long way to ensure the latter do not escape responsibility for such operations – at the same time, it is important to build regional and international attribution mechanisms that can clarify state practice and usher in greater transparency on the execution of such operations and the role of non-state groups, Moyakine notes. **Jaime Bello** reviews the legal dilemma faced by states in authorizing active defence measures by private actors in cyberspace. Given that private actors have the resources and infrastructure to execute sophisticated defensive cybersecurity measures, including measures in anticipation of an imminent cyber operation, states may be tempted to 'outsource' hack backs. But such measures

could also expose the state to international responsibility for wrongful acts, not to mention they run risk of escalation. In this scenario, Bello notes, the insurance industry could play an influential role in laying down standards and 'rules of engagement' for active measures. By auditing responses by private actors and suggesting strategies for risk reduction or management of collateral damage, insurance companies can not only mainstream active cyber defence but also clarify the responsibility of states in relation to such operations, he concludes. **Francesca Romana Partipilo** and **Marta Stroppa** review the evolution and digitalization of humanitarian information systems, and identify their main vulnerabilities to cyber threats. These threats may emerge from both states and non-state actors, and implicate a range of activities from information theft to destruction or manipulation of data held by humanitarian organizations that disrupt their provision of relief services. Partipilo and Stroppa offer six recommendations for international humanitarian organizations, both to defend their networks and to mitigate the aftereffects of serious cyber operations: in particular, they emphasize the need for such organizations to adopt cybersecurity policies and ensure those policies are monitored or audited periodically by independent entities. **Raman Jit Singh Chima** notes that the constitution of a UN Ad Hoc Committee on Cybercrime (UN AHC) is a 'net positive' development in that it is an opportunity to 'foster consensus and international collaboration on countering cybercrime.' From a human rights perspective, Chima notes, the UN Ad Hoc Committee would do well to keep two overarching goals in mind: the need to ensure cybersecurity research is not stunted by legal frameworks targeting criminal activity, and secondly, the importance of embedding safeguards on cross-border law enforcement cooperation with respect to the access and management of data. States should avoid the temptation to use the UN AHC platform to moot overbroad penal provisions for content regulation and use the negotiations as a platform to ensure cybercrime rules deter malicious actors, but also reform and clarify rules on how data is handled by law enforcement agencies, he argues.

These wide-ranging perspectives on the scope and evolving meaning of responsible behaviour in cyberspace may help scholars, policymakers, and diplomats working in the field of digital affairs and cyber security to inform, challenge, and strengthen their own perspectives. Closing a gap starts with identifying the gap, and mapping its various dimensions, which is the objective of this volume.

REGIONAL AND INTERNATIONAL COOPERATION

CHAPTER 1

The effectiveness of ASEAN regional efforts on cybersecurity

MONICA NILA SARI

Introduction

The speed of digitalisation has accelerated since the Covid-19 pandemic, making it one of the most significant growth engines for many developing nations. We are already seeing how digitalisation is reshaping the Southeast Asia region. The Association of Southeast Asian Nations (ASEAN), which consist of 10 member countries (Brunei Darussalam, Cambodia, Lao PDR, Indonesia, Malaysia, Myanmar, Philippines, Thailand, Singapore and Vietnam), is one of the fastest-growing internet markets in the world, with 125,000 new users coming online every day. ASEAN has more than 440 million internet users and, more importantly, 350 million of them (about 80%) are digital customers, i.e. internet users who have bought at least one online service.[1] With a fast-growing

1 'e-Conomy SEA Report 2021', available at: https://seamilano.eu/en/annual-report-2021

base of digital customers and merchants, and acceleration in e-commerce and food delivery, ASEAN's GDP reached more than USD$3.11 trillion in 2020, making it collectively the fifth largest economy in the world.[2] Accelerated digitalisation has helped to grow the region's digital economy, but has also led to new and novel challenges.

An analysis by A.T. Kearney indicated that ASEAN countries are being used as launch pads for cyberattacks—either as vulnerable hotbeds of unsecured infrastructure where numerous computers can be infected easily for large-scale attacks or as hubs for a single point of attack to gain access to the hubs' global connections.[3] Malaysia, Indonesia and Vietnam are global hotspots for major blocked suspicious web activities—up to 3.5 times the standard ratio, indicating that these countries are being used to launch malware attacks.[4] The same analysis argued that ASEAN's policy, governance and cybersecurity capabilities are relatively low. Heinl elaborated that while the ASEAN countries have taken steps to address some of the cyber-related challenges facing the region, these efforts are still at an early stage.[5] Other research on cybersecurity policy in ASEAN countries from Thammsat University[6] suggested that the issue that should be addressed is collaboration, since collaboration and information-sharing are a vital aspect of cybersecurity.

In the past year, cybersecurity has been a priority on the ASEAN agenda. However, ASEAN is characterised by a high degree of heterogeneity in terms of economic development, which resulted in notable gap in terms of cyber maturity and ASEAN countries' commitment and political will to engage with cybersecurity policy. This chapter will analyse the effectiveness of ASEAN's regional approach to cybersecurity issues in the region using Amos N. Guiora's theory of effectiveness, whereby cybersecurity effectiveness relies on policy allocating resources effectively based on a cost–benefit analysis and on accurate risk assessment. The chapter will also present a case study from one ASEAN country—Indonesia, the biggest internet user in the region—to demonstrate whether

2 'ASEAN Key Figures 2021', the ASEAN Secretariat, available at: https://asean.org/wp-content/uploads/2021/12/ASEAN-KEY-FIGURES-Chapter-1-4-Rev-28-Dec-2021.pdf

3 'Cybersecurity in ASEAN: An Urgent Call to Action', Cisco and A.T. Kearney, 5, available at: https://www.southeastasia.kearney.com/documents/1781738/1782318/Cybersecurity+in+ASEAN—An+Urgent+Call+to+Action.pdf/

4 'Internet Security Threat Report Volume 22', Symantec, available at: https://docs.broadcom.com/doc/istr-22-2017-en

5 Caitríona H. Heinl, 'Regional cybersecurity: moving toward a resilient ASEAN cybersecurity regime', *Asia Policy* 18 (18 July 2014).

6 Jirapon Sunkpho, Sarawut Ramjan and Chaiwat Ottamakorn, 'Cybersecurity policy in ASEAN countries', available at: https://www.researchgate.net/publication/324106226_Cybersecurity_Policy_in_ASEAN_Countries

ASEAN's cybersecurity framework has effectively contributed to ASEAN countries. Furthermore, a toolkit from the International Telecommunication Union (ITU) will be used to review the effectiveness of ASEAN's regional approaches.

ASEAN's efforts on cybersecurity

ASEAN leaders shared the vision of a peaceful, secure and resilient regional cyberspace that serves as an enabler of economic progress, enhanced regional connectivity and betterment of living standards, as stated in the ASEAN Leaders' Statement on Cybersecurity Cooperation.[7] Moreover, ASEAN recognises the multifaceted nature of cybersecurity, and the different dimensions of cybersecurity cooperation are discussed under each of the three pillars of ASEAN. On the ASEAN Political–Security Community pillar, it specifically addresses the need to strengthen cooperation on cybersecurity in all aspects, including developing and improving laws and capacity-building for law enforcement. Under the ASEAN Economic Community pillar, cybersecurity is discussed from the angle of cyber infrastructure and information protection, whereas discussion on cybersecurity within the ASEAN Socio-Cultural Community pillar is focused on the promotion of cyber wellness through policy initiatives and activities that relate to developing digital literacy and mitigating the harmful effects of fake news.[8]

Regional framework

Over the past few years, the ASEAN region has shown the way forward on how to build a regional cybersecurity cooperation framework. First, ASEAN updated its cybersecurity cooperation strategy as reflected in the ASEAN Cybersecurity Cooperation Strategy for 2021–2025 in response to recent cyber developments, to strengthen collective efforts to secure cyberspace for the region and to promote growth of the digital economy and community . The updated strategy contains five dimensions of work: (1) advancing cyber-readiness cooperation, (2)

7 'ASEAN Leaders' Statement on Cybersecurity Cooperation', 2018, available at: https://asean.org/wp-content/uploads/2018/04/ASEAN-Leaders-Statement-on-Cybersecurity-Cooperation.pdf

8 Ibid.

strengthening regional cyber-policy coordination, (3) enhancing trust in cyberspace, (4) regional capacity-building and (5) international cooperation.

Second, ASEAN is the first and only regional organisation to have subscribed, in principle, to the UN's 11 voluntary, non-binding norms of responsible state behaviour in cyberspace.[9] This is important to underpin ASEAN's active contribution to maintaining peace and security in cyberspace. In this regard, ASEAN is developing its Regional Plan on the Implementation of UNGGE Norms of Responsible State Behaviour in Cyberspace, which are categorised into several focus areas including international cooperation, development of policy, awareness-raising, strengthening national cybersecurity and cybercrime laws, cybercrime cooperation, incident response cooperation and creation of a trustworthy ecosystem.[10] This initiative has increased ASEAN countries' understanding and awareness of key cybersecurity issues, and will act as a useful guide in ASEAN's work on norms implementation.

Third, ASEAN is establishing the ASEAN Regional Computer Emergency Response Team (CERT) and the ASEAN CERT Information Exchange Mechanism. ASEAN recognised the urgency to secure the growing digital economy in the face of increasingly sophisticated transboundary cyber-attacks, therefore ASEAN CERT will be valuable in terms of facilitating the timely exchange of threat- and attack-related information among ASEAN member states' (AMSs) national CERTs and fostering CERT-related capacity-building and coordination.[11]

ASEAN mechanisms

Relevant ASEAN sectoral bodies and ASEAN-led mechanisms[12] have been working on cybersecurity issues, namely the ASEAN Digital Ministers' Meeting (ADGMIN) and the ASEAN Digital Senior Officials' Meeting (ADGSOM) as its subsidiary body; the ASEAN Regional Forum (ARF); the ASEAN Ministerial Meeting on Transnational Crime (AMMTC); the East Asia Summit (EAS); and the ASEAN Defence Ministers' Meeting-Plus (ADMM)-Plus.

9 'ASEAN Cybersecurity Cooperation Strategy 2021–2025', available at: https://asean.org/wp-content/uploads/2022/02/01-ASEAN-Cybersecurity-Cooperation-Paper-2021-2025_final-23-0122.pdf

10 Ibid.

11 Ibid.

12 ASEAN's constructive engagement with its external partners, through ASEAN-led mechanisms such as ASEAN Plus-One, ASEAN Plus-Three (APT), EAS, ARF and ADMM-Plus, builds mutual trust and confidence as well as reinforcing an open, transparent, inclusive and rules-based regional architecture with ASEAN at the centre.

The AMMTC has the mandate to discuss the the area of cybercrime. Under this mechanism, ASEAN adopted its Declaration to Prevent and Combat Cybercrime in 2017. Recognising the need to address the rapid growth of cybersecurity threats, the ARF established its Inter-Sessional Meeting (ISM) on Security of and in the Use of ICTs in 2017. This serves as a specific platform for ARF participants to promote mutual understanding as well as to discuss and coordinate ARF's efforts on ICTs security, to implement the ARF Work Plan on Security of and in the Use of ICTs and to enhance trust and confidence through capacity-building while ensuring trust and confidence in the conduct of its activities. To guide the work of the ISM on ICTs security, the ARF Work Plan on Security of and in the Use of ICTs was adopted in 2015. It serves to promote a peaceful, secure, open and cooperative ICT environment and to develop transparency and confidence-building measures to prevent conflict in cyberspace between states in the ARF region through capacity-building. ARF adopted 'ARF Terminology in the Field of Security of and in the use of ICTs' in September 2020 to encourage discussion among participants on their domestic views and definitions of key ICTs-related terminology utilised in their respective countries.

Initiatives on cybersecurity under the ASEAN Economic Community pillar are through the ADGMIN (formerly the ASEAN Telecommunication and Information Technology Ministers Meeting (TELMIN), it changed in 2019 to reflect the widening scope of work of the ICT ministries across ASEAN).[13] On cyber-defence, in 2021 the ADMM adopted concept papers on the ASEAN Cyber Defence Network and the ADMM Cybersecurity and Information Centre of Excellence, as important milestones in promoting practical cybersecurity cooperation in ASEAN. These efforts serve as confidence-building measures (CBMs) within the region, and ASEAN would like to encourage other regions to adopt similar measures, with a view to building trust and confidence at the global level. In order to reinforce the leaders' intention to strengthen cooperation in cybersecurity, this issue has been increasingly featured under the ambit of the EAS. This mechanism has provided workshop regional cyber capacity-building as well as leaders' commitment to promote open, secure, stable, accessible and peaceful cyberspace.

ASEAN has various mechanisms dealing with cybersecurity with the aim of facilitating the deliberations on cybersecurity cooperation under its three pillars. As cybersecurity is a cross-cutting issue, in 2020 ASEAN established the ASEAN Cybersecurity Coordinating Committee (ASEAN Cyber-CC) to tackle coordination

13 ASEAN Secretariat, www.asean.org

challenges, and to promote cross-sectoral and cross-pillar cooperation and strengthen cybersecurity in the region. Under this new mechanism, ASEAN is now developing a Regional Action Plan on the Implementation of the Norms of Responsible State Behaviour in Cyberspace to assist with the prioritisation and implementation of the 11 voluntary, non-binding norms of responsible state behaviour in the use of ICTs.

These ASEAN sectoral bodies and ASEAN-led mechanisms are not aiming only to produce chairman's statements or to adopt agreed documents. The regular meetings among regional leaders and officials provide a diplomatic ecosystem where many informal and side-line engagements take place. ASEAN meetings engender a sense of familiarity and a give-and-take approach, which in turn facilitate consensus-building on contentious issues. These mechanisms are also forums for ASEAN countries and partners to discuss relevant issues related to cybersecurity.

Case study of cyber-attack in Indonesia

ASEAN countries' internet penetration is now over 77.6%, which is above the level of internet users worldwide (59.5%).[14] With the ASEAN region seeing exponential growth in the digital technology sector, particularly financial technology and e-commerce, there is an increasing demand for internet and broadband services. However, this increasing reliance on the internet has created many security threats that can cause immense damage. Based on the ASEAN Cyberthreat Assessment 2021 produced by the Interpol ASEAN Cybercrime Operations Desk, ASEAN countries have become a prime target for cyber-attack on account of their position among the fastest-growing digital economies in the world.

Indonesia is the biggest country in ASEAN: it has a population of more than 275 million and, according to Statista, as of July 2021, online penetration in the country stood at around 70%. With over 171 million internet users, Indonesia is one of the biggest online markets worldwide. One of the most serious

14 Internet penetration in Southeast Asia as of June 2021, Statista, available at: https://www.statista.com/statistics/487965/internet-penetration-in-southeast-asian-countries/

cyber-attacks, in 2020, was the data breach incident of Indonesia's e-commerce Tokopedia, in which 91 million users' information was leaked.[15]

Tokopedia is considered the largest e-commerce marketplace in Indonesia, providing business opportunities to various small-scale vendors and small and medium enterprises; its marketplace has become a preferred selling and shopping destination. Its website became the most visited e-commerce site in Indonesia, with monthly traffic reaching 157 million.[16] Indonesia, as one of the fastest growing economies in ASEAN, continues to be a vibrant digital financial services market due to its relatively open regulatory framework, and is showing rapid growth across fintechs and digital platforms. Based on the Google, Temasek and Bein e-Conomy SEA 2021 report, Indonesia's internet economy will likely reach $146 billion by 2025.

After the data breach incident, Tokopedia CEO William Tanuwijaya reported to the Indonesian parliament how the company resolved the cyberattack. At that time, since there was not yet a regulation on data protection in Indonesia, the company claimed to follow international standards by delivering transparency to its customers through providing explanation as to which data had been breached. Furthermore, Tokopedia provided regular updates on how the attack was being handled and improved its system internally to prevent future attacks.[17]

Indonesia does not have a comprehensive personal data protection regulation. What does exist is a multitude of laws and regulations in many sectors governing personal data protection, namely Law No. 11 of 2008 on Electronic Information and Transactions, Government Regulation No. 71 of 2019 on Implementation of Electronic Systems and Transactions, and Minister of Communications and Informatics Regulation No. 20 of 2016 on Personal Data Protection in Electronic Systems.

If we refer to Amos N. Guiora's effectiveness theory of cybersecurity policy,[18] Tokopedia's data breach incident could be analysed by answering these questions: what is the impact and significance of a data breach of 91 million users' information*?* To what extent, from a policy perspective, does this data breach warrant significant attention and resources? And finally, from a policy perspective, what is the impact of Tokopedia's data breach on the overall development

15 'ASEAN Cyber Threat Assessment 2021', Interpol.

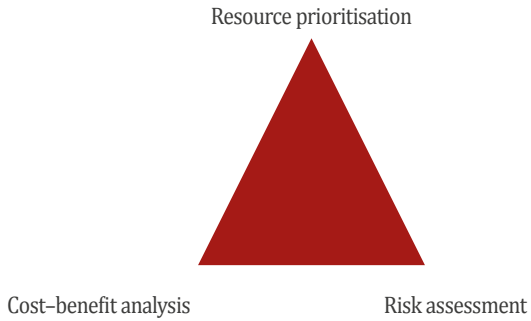
16 'The Map of E-commerce in Indonesia', iPrice, available at: https://iprice.co.id/insights/mapofecommerce/en/

17 Fanny Potkin, 'Indonesia's Tokopedia probes alleged data leak of 91 million users', Reuters (3 May 2020), available at: https://www.reuters.com/article/us-tokopedia-cyber-idUSKBN22E0Q2

18 Amos N. Guiora, *Cybersecurity: Geopolitics, Law, and Policy* (Abingdon: Routledge, 2017).

and assessment of cybersecurity in Indonesia? The question is put forward in the context of *resource prioritisation, cost–benefit analysis* and *risk assessment*.

FIGURE 1 | Resource triangle



In this regard, *what is the impact of significance of a data breach of 91 million users' information?* The significant impact was that data relating to names, emails and telephone numbers of 91 million users had been partly compromised, although Tokopedia explained that financial data were safe. The hackers claimed that email addresses and encrypted passwords from the company's user database were put for sale on the dark web for US$5,000.[19] Since there was not yet a regulation on data protection when the incident occurred, Tokopedia users could not claim and protect their rights. They were only advised to change their password on other digital platforms and not to shar OTP (one-time pin) codes.

To what extent, from a policy perspective, does this data breach warrant significant attention and resources? The Tokopedia incident reminded the Indonesian government of the need to develop and enact a regulation on personal data protection. On the resource dimension, Tokopedia announced that it had appointed an independent global institution specialising in cybersecurity to improve its security system, including the safety and security of its users' accounts and transactions.

From a policy perspective, what is the impact of Tokopedia's data breach on the overall development and assessment of cybersecurity in Indonesia? Indonesia has prioritised drafting a regulation on data protection since businesses need to know they can operate in a secure environment, while customers need to know

19 Potkin (see note 17 above).

that public services supporting their continued safety, health and welfare remain accessible. Thus, the impact of the Tokopedia incident is significant in terms of the development of a cybersecurity framework. Currently, Indonesia's draft regulation on data protection is being discussed intensively at parliament level.

The above case and the report from the Indonesia National Cyber and Crypto Agency that in 2021 there were 1.65 billion cyberattacks in Indonesia[20] demonstrate that Indonesia's cybersecurity framework is in urgent need of further strengthening and developing. In this regard, this chapter argues that ASEAN's cybersecurity framework has not contributed effectively to the biggest ASEAN country, particularly on the legal measures on data protection and technical measures in countering the cyber-attacks.

ASEAN's regional approach analysis

According to IBM Security's 2020 Cost of a Data Breach Report,[21] the average cost of a data breach in ASEAN in 2020 was estimated at US$2.7 million. Interpol's ASEAN Cybercrime Operations Desk reported that data breach was one of the commonest cybercrimes in ASEAN countries in 2021, with the Tokopedia incident on the list.[22] Moreover, ransomware of 1.5 terabytes of sensitive data stolen from a subsidiary of ST Engineering Aerospace in June 2020, ransomware of hospitals and businesses targeted in Thailand in September 2020, and a data breach of 1.1 million accounts of RedMart occurred in October 2020. Besides the cyberattacks and data breaches, there was also an increase in Covid-19-related online fraud, including the sale of medical equipment and personal protective equipment.

20 'The 2021 Security Monitoring Result', National Cyber and Crypto Agency, available at: www.bssn.go.id

21 'Cost of a Data Breach Report 2020', IBM, available at: https://www.ibm.com/downloads/cas/QMXVZX6R

22 Interpol (see note 15 above).

Key challenges

First, as one of the most successful regional organisations in the world, ASEAN has an 'ASEAN way' approach in its decision-making process, which is upholding the consensus principle based on the ASEAN Charter. Some scholars argued that this ASEAN way could limit the group of 10 in finding common ground and mutually acceptable outcomes. Moreover, ASEAN respects the principle of territorial integrity, sovereignty, non-interference and national identities of ASEAN member states.[23] The question arises whether this ASEAN way and principle of ASEAN regionalism are effective in dealing with cybersecurity in the region.

Second, according to ITU's Global Cybersecurity Index (GCI) 2020, ASEAN countries range from 4 to 131 among 194 countries in total (Table 1).

TABLE 1 | Cybersecurity maturity of ASEAN countries.

Country	Rank	Score	Legal measures	Technical measures	Organisational measures	Capacity development	Cooperative measures
Singapore	4	98.52	20.00	19.54	18.98	20.00	20.00
Malaysia	5	98.06	20.00	19.08	18.98	20.00	20.00
Indonesia	24	94.88	18.48	19.08	17.84	19.48	20.00
Vietnam	25	94.55	20.00	16.31	18.98	19.26	20.00
Thailand	44	86.50	19.11	15.57	17.64	16.84	17.34
Philippines	61	77.00	20.00	13.00	11.85	12.74	19.41
Brunei Darussalam	85	56.07	14.06	14.19	10.84	12.85	4.12
Myanmar	99	36.41	9.39	3.64	4.71	8.92	9.75
Lao PDR	131	20.34	11.77	3.27	0.00	1.23	4.07
Cambodia	132	19.12	7.38	2.50	1.69	3.29	4.26

Source: ITU's Global Cybersecurity Index 2020.

Third, ASEAN countries have not yet spent enough on cybersecurity to secure a sustained commitment to cybersecurity and mitigate the investment gap. A.T. Kearney's report argued that to secure sustained commitment to cybersecurity and address the investment gap, ASEAN countries need to spend between 0.35% and 0.61% of their GDP—or US$171 billion collectively—on cybersecurity in the

23 The ASEAN Charter, available at: https://asean.org/wp-content/uploads/images/archive/publications/ASEAN-Charter.pdf

period spanning 2017–2025.[24] Based on a report by Palo Alto Networks, cybersecurity has risen to the top of the leadership agenda for many ASEAN businesses, with the vast majority (92%) believing it to be a priority for their business considering the growing volume of cyber threats in the region.

As surveyed, most ASEAN organisations increased their security investments in 2019. In fact, 46% allocated at least half their total IT budget to cybersecurity. It is also mentioned that 53% of Singapore companies allocated over half their IT budget to cybersecurity and 84% of Indonesian companies increased their cybersecurity budgets between 2019 and 2020, which was the biggest jump in ASEAN.[25] Regarding government-allocated funds, Singapore, as the leading country in ASEAN in terms of cyber maturity, has allocated US$1 billion to build up the government's cyber and data security capabilities for the 2020–2023 budget.[26] Malaysia allocated US$6 million in 2021 to strengthen the nation's cybersecurity capacity[27] and Indonesia allocated US$89 million in 2021 for ICT development.[28] However, other countries in ASEAN have not yet allocated the same proportion of budget for cybersecurity.

Data protection in ASEAN

Since 2020, Malaysia, Singapore, the Philippines and Thailand already have comprehensive general data protection laws in place, while in the other six members the matter is pending passage or covered in various pieces of legislation.[29] The case study in Indonesia has clearly demonstrated that Indonesia is one of the ASEAN countries that are in urgent need of enacting a law on data protection.

ADGSOM has adopted the ASEAN Framework on Digital Data Governance, which aims to align baseline principles and standards for data protection,

24 Cisco and A.T. Kearney (see note 3 above).

25 'The State of Cybersecurity in ASEAN', Palo Alto Networks (2020), available at: https://www.paloaltonetworks.sg/apps/pan/public/downloadResource?pagePath=/content/pan/en_SG/resources/whitepapers/the-state-of-cybersecurity-in-asean-2020

26 Lim Min Zhang, 'Singapore Budget 2020: $1b over next 3 years to shore up cyber and data security capabilities', Straits Times (18 February 2020), available at: https://www.straitstimes.com/singapore/singapore-budget-2020-1b-over-next-3-years-to-shore-up-cyber-and-data-security

27 Angelin Yeoh, 'Budget 2021: RM27mil allocation for CyberSecurity Malaysia hailed by industry players', The Star (6 February 2020), available at: https://www.thestar.com.my/tech/tech-news/2020/11/06/budget-2021-rm27mil-allocation-for-cybersecurity-malaysia-hailed-by-industry-players

28 Indonesia Ministry of Finance, www.kemenkeu.go.id

29 TRPC (2020), 'TRPC Data Protection Index 2020', available at: https:/trpc.biz/old_archive/wp-contents/uploads/TRPC_DPI2020.pdf

advance digital innovation and the use of open and big data, and facilitate data flows.[30] In particular, the ASEAN Data Management Framework and the Model Contractual Clauses for Cross Border Data Flows were approved by the first ADGMIN in January 2021.[31] In addition, the ASEAN Cybersecurity Resilience and Information Sharing Platform (CRISP) has fully operationalised for the participating AMSs that signed the Memorandum of Understanding (MOU) for Sharing of Information during Activities of Digital and Technology Network (DTN) on 1 February 2021, which allows information-sharing to combat cybersecurity threats and to develop collaborative mitigation actions for ASEAN central banks.

However, data protection and cybersecurity are ongoing processes. In order to support the development of regional regulatory environment, ASEAN countries need to make sure their domestic data protection laws are updated regularly to remain relevant to the digital economy, such as enacting coherent and simple rules to both enable and protect cross-border data flows, clear obligations and responsibilities defined for data processors and data controllers, and transparent data breach notification process. In this regard, ASEAN may eventually create a regional framework on data protection in order to mitigate cybercrime in the region.

Existing regional efforts

In order to analyse the effectiveness of ASEAN regional efforts on cybersecurity, this chapter measures cybersecurity commitments across five pillars based on the toolkit from the ITU (Table 2).[32]

30 ASEAN (2018), 'Framework on Digital Data Governance', available at: https://asean.org/storage/2012/05/6B-ASEAN-Framework-on-Digital-DataGovernance_Endorsedv1.pdf

31 ASEAN (2021), 'ASEAN Data Management Framework', available at: https://asean.org/storage/2-ASEAN-Data-Management-Framework_Final.pdf

32 'Global Cybersecurity Index 2020', ITU, available at: https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E

TABLE 2 | Five pillars of cybersecurity commitments.

Pillars	ASEAN commitments
Legal measures	ASEAN has yet to develop a legal framework on cybersecurity. The case of Indonesia demonstrated the urgency to have a legal framework on data protection.
Technical measures	· ASEAN is focused on upgrading the technical capability of its national CERTs. Based on the ASEAN Cybersecurity Cooperation Strategy 2021–2025, each ASEAN country shall assess the technical capability of its national CERT in the areas of cyber-threat monitoring, incident handling, vulnerability handling, evidence handling, alerts and advisory drafting towards achieving a defined level of competency. · ASEAN is establishing ASEAN CERT to facilitate the timely exchange of threat- and attack-related information among ASEAN countries' national CERTs and foster CERT-related capacity-building and coordination.
Organisational measures	· ASEAN has a cybersecurity strategy, as reflected in the document 'ASEAN Cybersecurity Cooperation Strategy 2021–2025', and it is updated from the 2017–2020 document. In this regard, ASEAN updates its cybersecurity regularly. · ASEAN created a new mechanism in 2020 to strengthen cross-sectoral coordination as cybersecurity, which is the ASEAN Cybersecurity-CC. · ASEAN countries have established their national CERTs. · To date, only Singapore, Malaysia, Indonesia, Brunei Darussalam and Myanmar have a national cyber agency, while other ASEAN counties are represented by their relevant ministries. · Development of ASEAN's Critical Information Infrastructure Protection (CIIP) Coordination Framework, built on the 2020 ASEAN CIIP Framework, is to provide strategic recommendations and coordinated approaches to create more resilient cybersecurity across ASEAN's critical information infrastructure.
Capacity development	· ASEAN has three regional initiatives on capacity-building: 1. ASEAN–Japan Cybersecurity Capacity Building Centre (AJCCBC) 2. ASEAN–Singapore Cybersecurity Center of Excellence (ASCCE) 3. ADMM Cybersecurity and Information Centre of Excellence (ACICE) · ASEAN also organises and provides targeted capacity-building through mechanisms such as ARF, ADMM, ADMM-Plus, AMMTC, EAS and ADGMIN.
Cooperation	· ASEAN has established a framework to widen and deepen its relations with external parties through the conferment of the formal status of dialogue partner on Australia, Canada, China, the EU, India, Japan, Republic of Korea, New Zealand, Russia, the US and the UK. · With these dialogue partners, ASEAN established the ASEAN+1 process to discuss and review state cooperation between ASEAN and the dialogue partners as well as strengthening cooperation in priority areas such as cybersecurity. · Through the ASEAN+1 process, ASEAN has managed to enhance cybersecurity cooperation as reflected, for example, in the 2018 ASEAN–US Leaders' Statement on Cybersecurity Cooperation, the 2019 ASEAN–EU Statement on Cybersecurity Cooperation, the inaugural ASEAN–Australia Cyber Policy Dialogue, ASEAN–Japan Cybersecurity Working Groups and Policy Meetings, and an annual workshop on network security with China.

Based on Table 2, in order to counter cybercrime in the region, such as the data breach incident in Indonesia, besides the legal measures, there are two points that most ASEAN countries need to develop. The first is organisation measures. One of the key organisational measures is national cybersecurity strategy, and it has to be updated regularly by assessing current risks, prioritising cybersecurity interventions, and tracking progress and having a clear set of objectives on the protection of critical infrastructure. From the ITU's GCI, we learn that a lack of adequate organisational measures can contribute to a lack of clear responsibilities and accountability in national cybersecurity governance, and can prevent effective intra-government and inter-sector coordination. If all ASEAN countries have established an effective national cybersecurity, this will contribute to the development of ASEAN cybersecurity strategy. Brunei Darussalam, Indonesia, Malaysia, the Philippines, Singapore, Thailand and Vietnam have already developed national strategies related to cybersecurity and can do more to promote regional alignment and assist ASEAN countries that have yet to craft their own cybersecurity roadmaps or implementation strategies.

The second point is the technical measures. Legislation and regulation are important, but the actual implementation of cyber-threat detection systems and the capability to handle cyber risks are more important. In order to improve technical capabilities, ASEAN should enhance its capacity-building programme. Currently there are three ASEAN initiatives on regional capacity-building, namely: (1) the ASEAN–Japan Cybersecurity Capacity Building Centre, established in 2018 in Bangkok, Thailand, (2) the ASEAN–Singapore Cybersecurity Center of Excellence, established in 2019 in Singapore; and (3) the ADMM Cybersecurity and Information Centre of Excellence, established in 2021 in Singapore. The ASEAN–Japan Cybersecurity Capacity Building Centre conducts programmes on technical hands-on computer simulation, digital forensics and malware analysis to improve cybersecurity and trusted digital services among ASEAN countries.[33] The ASEAN–Singapore Cybersecurity Center of Excellence provides training in areas covering cybersecurity norms and policy, CERT-related technical training and virtual cyber-defence training and exercises.[34] The ADMM Cybersecurity and Information Centre of Excellence has three objectives: (a) to function as a node for confidence-building measures, information-sharing and capacity-building among regional militaries; (b) to enhance regional cooperation and information

33 ASEAN–Japan Cybersecurity Capacity Building Center, https://www.ajccbc.org

34 Cyber Security Agency Singapore, https://www.csa.gov.sg/News/Press-Releases/asean-singapore-cybersecurity-centre-of-excellence

sharing, focusing on cybersecurity and disinformation and misinformation threats including through the dissemination of regular and timely reports; and (c) to work with international experts to improve collective resilience against common security threats.[35]

Considering the huge maturity gaps among ASEAN countries, the regional capacity-building should focus on (i) developing the technical ability of ASEAN countries' CERTs, (ii) developing policy, strategy and technical aspects of cybersecurity for ASEAN countries' officials and cybersecurity professionals, and (iii) improving the ability and preparedness of cybersecurity professionals within the ASEAN region for cybersecurity and trusted digital services.

Conclusion

There is no best cybersecurity standard or framework—as new technologies and delivery mechanisms develop, cybersecurity will continue to expand and accommodate change—but there are good examples. Since the ASEAN leaders committed to enhancing cybersecurity cooperation in 2018, ASEAN has made significant progresses. It has strengthened its cybersecurity effort in: (1) the technical dimension by enhancing CERT cooperation; (2) the organisation dimension by updating its strategy and establishing the ASEAN Cybersecurity Coordinating Committee; (3) capacity-building with the three ASEAN initiatives and targeted capacity-building training; and (4) mutually beneficial and effective cooperation within ASEAN countries and with external partners.

Based on the assessment above, ASEAN needs to develop an effective legal framework of cybersecurity baselines and compliance mechanisms that could be implemented in all ASEAN countries, as well as procedures to ensure consistency with international obligations. Indonesia's Tokopedia incident shows the urgency of this. If ASEAN succeeds in producing a legal framework on cybersecurity, all ASEAN countries will have an umbrella to implement the rules in their national laws. ASEAN has been successful on the issue of countering terrorism by agreeing on the ASEAN Convention on Counter Terrorism: a legally binding document that provides a regional cooperation framework to counter, prevent and suppress terrorism. In the future, ASEAN should produce a similar binding

35 ASEAN (see note 9 above).

document on cybersecurity whereby its regional approach could be extended effectively to all ASEAN countries.

ASEAN, as a regional organisation, has limitations in terms of finding mutually acceptable outcomes and implementing the agreed regional framework considering its principle of non-interference and the ASEAN way of decision-making by consensus. In the case of cybersecurity, these limitations are significant since ASEAN countries have a high degree of heterogeneity in terms of economic development, which results in wide disparities in commitment and political will to engage with cybersecurity policy. This is shown in the notable gaps among ASEAN countries in terms of cyber maturity. However, this limitation is also the strength of ASEAN, which has provided forums through various mechanisms to discuss cybersecurity among ASEAN countries and with external partners. This regular interaction between stakeholders serves to increase knowledge and understanding between relevant actors, and to strengthen cybersecurity development. If trust-based relationships can be built, solutions to cybersecurity challenges can be found.

To narrow the gaps among member states, ASEAN can consider focusing on capacity-building in the three centres (AJCCBC, ASCCE, ACICE) to enhance organisational and technical measures. The capacity-building programme can be focused on improving these two dimensions for ASEAN countries with the lowest cyber maturity by improving the capability of national CERTs, training in areas covering cybersecurity norms and policy, and regular assessments of their cybersecurity commitments. At the same time, ASEAN countries with higher cyber maturity could share their best practices in handling cybersecurity challenges regularly. ASEAN has experience with its Initiative for ASEAN Integration (IAI) in providing a framework for regional cooperation whereby the more developed ASEAN countries could provide assistance for those that most need it, with a view of narrowing the development gap and enhancing ASEAN's competitiveness in the region. This IAI has shown its effectiveness through ASEAN's positive GDP trend, which has made it the fifth largest economy in the world. With this experience, ASEAN could undertake a similar regional approach on cybersecurity by narrowing the gaps among member states to improve its cybersecurity framework. Thus, a question remains as to which measures will need to improve next for Southeast Asia to develop its regional cybersecurity resilience.

CHAPTER 2

Online content regulation in the BRICS countries

A cybersecurity approach to responsible social media platforms

LUCA BELLI, YASMIN CURZI DE MENDONÇA AND WALTER B. GASPAR

Introduction

The increasing relevance of digital platforms for everyday societal activities has been generating concerns regarding the concentration of political and economic power in a few private enterprises. The substantial risk of electoral interferences, manipulation and widespread circulation of harmful content has led several countries to draft and enact regulations targeting primarily social media platforms[36] to regain control over such sensitive matters.

36 Platforms can be seen as the technical and governance structures that facilitate relationships and exchange of value between different categories of users. Digital platforms provide a governance structure, via their private ordering, and a technical architecture, via a wide range of standards, protocols and algorithms. See Luca Belli, 'Platform', in Luca Belli, Nicolo Zingales and Yasmin Curzi (eds), Glossary of Platform Law and Policy Terms (Rio de Janeiro: FGV Direito Rio, 2021).

Online content regulation is a core cybersecurity issue as it is instrumental in preserving the security of political infrastructures.[37] In particular, when dealing with the phenomenon of disinformation, there are significant overlaps and even similarities and synergies between the tools and mechanisms through which information disorder is organised and other cyber threats.[38]

This chapter analyses the regulatory state of the art in the BRICS grouping, composed of Brazil, Russia, India, China and South Africa. We consider that, although keeping a low profile as a group, the BRICS countries have acquired an increasing relevance at both regional and global levels, crafting impactful policies and enhancing their cooperation on digital matters. Importantly, their relevance is due not only to their economic weight but also to their mounting influence as policy setters.

Furthermore, it is interesting to highlight that some BRICS countries, notably China and Russia, started defining their content regulation frameworks in the early 2000s and aligned them internationally through the Shanghai Cooperation Organisation (SCO).[39] Indeed, since 2011, the SCO has elaborated on an International Code of Conduct for Information Security[40]—updated in 2015[41]—recognising that information security includes content control within digital media and reaffirming that 'policy authority for Internet-related public policy issues is the sovereign right of States'.

Since 2011, the SCO, which India joined as a full member in 2016, has emphasised that international human rights law (IHRL) allows restrictions to freedom of expression under specific circumstances stated in article 19.3 of the International Covenant on Civil and Political Rights. However, limitations to freedom of expression must be necessary and proportionate to a legitimate aim. As we will discuss, BRICS countries achieve mixed results as regards meeting the tests of necessity and proportionality. SCO states tend to have more pervasive

37 Usually, literature identifies four macro-areas of cybersecurity: data protection, safeguards of financial interests, protection of public and political infrastructures, and control of information and communication flows. See Laura Fichtner, 'What kind of cyber security? Theorising cyber security and mapping approaches', Internet Policy Review 7 (2) (2018), 1–19.

38 Kevin Matthe Caramancion, Yueqi Li, Elisabeth Dubois and Ellie Seoe Jung, 'The missing case of disinformation from the cybersecurity risk continuum: a comparative assessment of disinformation with other cyber threats', Data 7(4) (2022).

39 The SCO is an intergovernmental organisation aimed at political, economic and security cooperation. It covers three-fifths of the Eurasian continent and was established in 1996, in Shanghai, by China, Russia, Kazakhstan, Kyrgyzstan and Tajikistan. See http://eng.sectsco.org

40 See 'Letter dated 12 September 2011 from the Permanent Representatives of China, the Russian Federation, Tajikistan and Uzbekistan to the United Nations addressed to the Secretary-General', available at: https://digitallibrary.un.org/record/710973

41 For the differences between the 2011 and 2015 versions of the Code, see https://openeffect.ca/code-conduct/

information controls, content restrictions, and sanctions—even resulting in criminal punishment. Brazil and South Africa are struggling, so far with limited results, to design frameworks to regulate content effectively.

After providing a brief introduction to the BRICS grouping and the growing importance of digital policies in BRICS fora, stressing the relevance of cybersecurity in the bloc's agenda, we discuss the countries' most recent policy development at the national level. In this sense, this work's research question is to identify the common trends among the BRICS countries regarding cybersecurity and online platforms regulation.

The BRICS and their cybersecurity landscape

The 'BRICS' acronym, coined by Goldman Sachs economist Jim O'Neill, refers to four large emerging economies that experienced a similar and acute phase of economic development: Brazil, Russia, India, and China (South Africa joined the grouping later).[42] After getting acquainted with club governance as key emerging leaders invited to the G7/G8 summits via the so-called 'outreach process',[43] the BRICS countries started to increase their synergies.

Since the creation of the grouping, the number and type of BRICS governmental and multistakeholder gatherings, partnerships and initiatives have grown considerably.[44] In 2014, the bloc established the BRICS-led New Development Bank (NBD)[45] and Contingent Reserve Arrangement—one of its most prominent institutional achievements. Moreover, BRICS heads of state have never missed any of the group summits, thus demonstrating its importance for them.

42 See Jim O'Neill, 'Building better global economic BRICs', Goldman Sachs Global Economic Papers 66 (2001), available at: https://www.goldmansachs.com/insights/archive/archive-pdfs/build-better-brics.pdf

43 The most relevant of such processes was the 'G8 Outreach Five', which added Brazil, China, India, Mexico and South Africa to the 2005 G8 summit (Russia was still part of the G group itself). However, while the outreach model recognised the relevance of emerging economies—notably the future BRICS members—it also led to a shared sense of exclusion, as the countries kept being merely invited as guests, with a marginal role compared to the G members.

44 For detailed overviews of the evolution of BRICS, see Oliver Stuenkel, The BRICS and the Future of Global Order (Lanham, MD: Lexington Books, 2016).

45 See https://www.ndb.int

Regarding cybersecurity, the 2013 revelations of National Security Agency (NSA) contractor Edward Snowden represented a particularly salient event for the BRICS. Most prominently, these illegal activities included wiretapping illegally the Brazilian president's personal phone[46] and the communications of a large number of members of the Brazilian government. This triggered the elaboration and implementation of a wide range of cybersecurity policies in the countries and enhanced their cooperation.[47]

Tellingly, the eThekwini Declaration issued as an outcome of the 2013 Durban Summit of the BRICS included, for the first time, an explicit reference to cybersecurity, stressing the 'paramount importance' of 'security in the use of Information and Communication Technologies (ICTs)'.[48] Furthermore, in 2014, the BRICS technology and communication ministers started a cooperation process establishing the BRICS Working Group on the Security of ICTs,[49] and adopting the BRICS Memorandum of Understanding on Cooperation in Science, Technology and Innovation.[50] Such yearning for cooperation seems to have recently acquired a renewed impetus, with the 2021 BRICS Declaration calling for establishing 'legal frameworks of cooperation among the BRICS States [and] a BRICS intergovernmental agreement on cooperation'.[51]

While the Ukrainian war has indubitably put under strain all diplomatic initiatives involving Russia, it is safe to state that BRICS members' commitment to the grouping remains unchanged. The entire calendar of events was confirmed under the 2022 Chinese rotating presidency. BRICS members continue to consider the group a diplomatic priority, despite the divergence of opinions regarding the

46 See Sônia Bridi and Glenn Greenwald, 'Documentos revelam esquema de agência dos EUA para espionar Dilma', Fantástico, 1 September 2013, available at: http://g1.globo.com/fantastico/noticia/2013/09/documentos-revelam-esquema-de-agencia-dos-eua-para-espionar-dilma-rousseff.html

47 For an analysis of BRICS digital policies and most recent developments, particularly in the field of cybersecurity, see Luca Belli (ed.), CyberBRICS: Cybersecurity Regulations in the BRICS Countries (Cham: Springer, 2021); Luca Belli, 'Cybersecurity policymaking in the BRICS countries: from addressing national priorities to seeking international cooperation', African Journal of Information and Communication 28 (2021); Luca Belli and Danilo Doneda, 'Data protection in the BRICS countries: legal interoperability through innovative practices and convergence', International Data Privacy Law (ipac019) (2022).

48 See BRICS (Fifth BRICS Summit), 'eThekwini Declaration' (Durban, 2013) para 34, available at: http://mea.gov.in/bilateral-documents.htm?dtl/21482

49 For an analysis of such documents and their impact, see Vladimir Kiselev and Elena Nechaeva, 'Priorities and possible risks of the BRICS countries' cooperation in science, technology and innovation', BRICS Law Journal 5 (4) (2018), 33–60.

50 See BRICS (Second BRICS Science, Technology and Innovation Ministerial Meeting), 'BRICS Memorandum of Understanding on Cooperation in Science, Technology and Innovation' (Brasília, 18 March 2015), available at: https://www.gov.br/mre/pt-br/canais_atendimento/imprensa/notas-a-imprensa/ii-reuniao-de-ministros-de-ciencia-tecnologia-e-inovacao-do-brics-documentos-aprovados-brasilia-18-de-marco-de-2015

51 BRICS (XIII BRICS Summit), 'New Delhi Declaration' (9 September 2021), available at: https://brics2021.gov.in/brics/public/uploads/docpdf/getdocu-51.pdf

Ukrainian war. A meeting of the BRICS Ministers of Foreign Affairs in May 2022 was remarkably cooperative, culminating with the release of a Joint Statement on 'Strengthen[ing] BRICS Solidarity and Cooperation, Respond[ing] to New Features and Challenges in International Situation'.[52]

After the abovementioned BRICS meeting, the Brazilian Ministry for Foreign Affairs 'reiterated its support for intra-BRICS cooperation'[53] and highlighted that the grouping has 'shown concrete results',[54] emphasising that BRICS is a forum focused on international cooperation and sustainable development and on building a more robust multipolar order and inclusive global governance for the benefit of developing countries.

Recent developments in the BRICS provide substantial evidence that these countries' roles and interactions are starting to acquire global relevance for digital policymaking, besides their national and regional impact. Notably, the 13th BRICS Summit, hosted by India in September 2021, gave particular prominence to cybersecurity.[55]

While the five countries' national approaches diverge in many aspects, it is possible to identify several points of overlap and even tendencies towards convergence. Remarkably, their approaches to cybersecurity have started to converge and intensify ever since the creation of the 'Working Group of Experts of the BRICS States on security in the use of ICTs' in 2014, with a mandate to, inter alia, 'develop practical cooperation with each other in order to address common security challenges in the use of ICTs'.[56]

While agreeing on shared principles and high-level objectives through the annual declarations, the countries have crafted a unique blend of normative and developmental approaches to shape how (cybersecurity) cooperation and

52 See BRICS Joint Statement on 'Strengthen BRICS Solidarity and Cooperation, Respond to New Features and Challenges in International Situation', press release no. 76 (19 May 2022), available at: https://www.gov.br/mre/en/contact-us/press-area/press-releases/brics-joint-statement-on-201cstrengthen-brics-solidarity-and-cooperation-respond-to-new-features-and-challenges-in-international-situation201d

53 See the official Twitter account of the Brazilian Foreign Affairs Ministry: https://mobile.twitter.com/Itamaraty_EN/status/1527398486454460417

54 Ibid.

55 See BRICS (2021), BRICS India 2021—XIII BRICS Summit—New Delhi Declaration, available at: https://brics2021.gov.in/brics/public/uploads/docpdf/getdocu-51.pdf

56 BRICS (2015). VII BRICS Summit—Ufa Declaration, available at: https://www.brics2021.gov.in/BRICSDocuments/2015/Ufa-Declaration-2015.pdf

regulation should unfold.[57] However, such an approach is not immediately understandable for an observer used to consider only the normative side of regulation, i.e. regulation by prohibiting undesired behaviours and oversight by a specific authority. Indeed, cooperation and regulation, on cybersecurity or any other matters, can be achieved, arguably more effectively, through other means than mere norm-making, such as investments and standardisation.

Lastly, it is essential to emphasise that, despite the ambitions and intentions expressed in the BRICS annual declarations and official documents, the ease with which intra-BRICS cooperation on cybersecurity issues can occur remains unclear. On the one hand, most content regulation issues are highly sensitive, and national policymakers' decisions regarding content restrictions represent the quintessence of domestic cultural, political and legal peculiarities, thus making them less than ideal candidates for international consensus.[58] Nevertheless, the likeliest rapprochement is in the form of information and good practice (or bad practice, depending on the observer's standpoint), for which a dedicated intra-BRICS body already exists.

Hence, it is crucial to evaluate the domestic approach of the various BRICS members to cybersecurity to understand in which areas and to what extent coordination, convergence or divergence is most likely to occur. In addition, content regulation and online platform responsibility have become prominent in national debates, mainly due to disinformation. The following sections provide an overview of the latest national developments to shed light on what BRICS approaches converge, or even reproduce each other, and on what elements the countries are taking different paths.

57 See Luca Belli, 'Data protection in the BRICS countries: enhanced cooperation and convergence towards legal interoperability', CyberBRICS (2020), available at: https://cyberbrics.info/data-protection-in-the-brics-countries-enhanced-cooperation-and-convergence-towards-legal-interoperability/; Luca Belli, 'CyberBRICS: a multidimensional approach to cybersecurity for the BRICS', in Luca Belli (ed.), CyberBRICS: Cybersecurity Regulations in the BRICS Countries (Cham: Springer, 2021).

58 See Belli, 'Cybersecurity policymaking in the BRICS countries' (note 12 above).

Recent developments in the BRICS countries

Brazil

Brazilian social media regulation relies on the Brazilian Civil Rights Framework for the Internet, Law n. 12,965/2014, aka 'Marco Civil da Internet' (MCI), which is in the process of being supplemented by Draft Bill n. 2,630/2020, aka the 'Fake News Bill.'

Brazilian Civil Rights Framework for the Internet

The MCI is Brazil's primary law regarding internet regulation and the first and only general law for internet governance adopted in Latin America. It establishes rules and principles for a democratic, plural and neutral internet and defines general provisions for application providers. Article 19 establishes a general regime[59] of a judicial notice-and-takedown[60] system whereby application providers can only be liable for user-generated content (UGC) if failing to comply with court orders for the removal of specified content within 24 hours, granted that they have the technical capacity to do so. The rationale was that by the imposition of such legal procedure, abusive requests would not follow through and only valid demands would come to the judiciary,[61] ensuring legal certainty for the companies.

59 The exceptions are articles 19.2 and 21, which respectively refer to copyright infringement and intimate imagery and provide a notice-and-takedown regime.

60 Before MCI, the Brazilian Superior Court of Justice (STJ) was in the process of 'unifying' its jurisprudence to establish the notice-and-takedown regime as the general regime in the country, influenced by the North American Digital Millennium Copyright Act (DMCA). In its session 512, DMCA enacts a 'safe harbor' for service providers, which are exempt from liability if they have set notice-and-takedown procedures enabling users (copyrights holders) to request a quick removal of infringing content. STJ justice Nancy Andrighi even mentioned such legislation in a case against Google to condemn the search engine for not complying with a takedown request by an offended user. However, this majority opinion neglected the massive number of requests for content removal—not all of which are valid and lawful.

61 It is also relevant to mention that Brazil has a relatively functional public judicial system. 'Access to justice' is, in fact, a constitutional right (article 5°, XXXV), and in article 19.3, MCI assures that users can refer their cases to Special Courts, where they can count on free legal assistance and an expedited judicial procedure.

The Brazilian Supreme Court will soon assess article 19's constitutionality in Extraordinary Appeals ('RE') n. 1,037,396/SP and n. 1,057,258/MG[62]—both questioning intermediaries' role in amplifying users' rights violations. Furthermore, following other countries' initiatives to curb disinformation and other harms, the Brazilian legislators started drafting bills towards this goal, establishing platforms' responsibilities and transparency. The main result—now under discussion in the Federal Congress—is the Draft Bill on Freedom, Responsibility and Transparency on the Internet, PL n. 2,630, presented in 2020, aka 'PL das *Fake News*'.

Draft Bill on 'Freedom, Responsibility and Transparency' of application providers

In 2020, Senator Alessandro Coronel presented to the Brazilian Federal Senate Draft Bill n. 2,630/2020, submitting it to the National Chamber on 3 July for appreciation. Experts and civil society organisations criticised the draft's first version due to problematic provisions such as traceability of communications for tackling disinformation, criminalisation of disinformation spread, and the absence of more sophisticated transparency and users' rights provisions and a proper governance model.

The Draft Bill is currently under debate at the National Chamber, having as its rapporteur Deputy Orlando Silva. The Chamber held multiple public hearings in 2021, counting on the participation of civil society organisations and experts to improve the draft, which culminated in entirely new versions presented by its rapporteur on 4 November 2021 and 31 March 2022.

Some improvements of the current version merit highlighting: the provision on the criminalisation of disinformation dissemination now targets only

62 In the first case, a Facebook user had a fake account created in her name and issued a lawsuit for Facebook to delete it, requesting compensation. The regional appeals court not only ordered Facebook to delete the fake profiles and to pay for damages but also declared the 'incidental unconstitutionality' of article 19, considering that Facebook did not act expeditiously before the lawsuit. The regional judges argue that article 19 is incompatible with the Brazilian Federal Constitution regarding consumer protection (art. 5°, XXXII) and general civil rights provisions, such as intimacy, privacy, honour and reputation (art. 5°, X). Facebook appealed to the Supreme Court, remarking that article 19 determines that intermediary liability should only stem from failing to comply with a judicial request when it is proven that the company could do so. In the second case, students from a school in Minas Gerais created a forum on the social network Orkut (controlled by Google) to criticise a teacher. She demanded that Orkut remove the page. This case followed Facebook's in much the same way, with Google losing and appealing to the Supreme Court—which joined the two appeals, due for judgment in June 2022. The whole lawsuit can be accessed at the Brazilian Federal Supreme Court website: https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5160549&numeroProcesso=1037396&classeProcesso=RE&numeroTema=987

coordinated actions by enterprises/companies, not by individuals. In addition, it altered the traceability provision in compliance with due process in criminal law—it must be based on (1) previous intelligence work, (2) presumption of innocence and user privacy, and (3) security of communications.

Nevertheless, the Draft Bill left broad room for platforms' self-regulation. According to the current version, it is up to them to create their own codes of conduct to assure transparency and accountability—which the CGI.br (the Brazilian Internet Steering Committee), a multisectoral entity, must certify. However, the Steering Committee does not have enforcement tools or power under the law to enforce regulations. Therefore, there is a high risk that the codes of conduct will deviate entirely from what the law intended.

Regarding transparency reports duties, the draft only requires information on the total number of actions taken to moderate content (e.g. the amount of social media posts that have been removed) that fail to provide meaningful transparency and do not allow the identification of biases and failures in moderation or recommending systems, according to several experts.[63] In addition, the draft does not present a methodology or model for presenting reports, making it challenging to monitor failures and biases.

Final considerations regarding Brazil's social media regulation

Arbitrary removal, shadow bans[64] and lack of transparency are often pointed out as issues impacting free speech and democracy. With the growth of platforms' powers, governments must move toward platform observability[65] to assure non-discrimination and democratic legitimacy of their actions before civil society.

In this sense, within the constitutionality of MCI's article 19 debate, the Supreme Court[66] can propose the differentiation of duties between very large platforms and other actors, fostering fundamental rights and innovation. In

63 Nicolas P. Suzor, Lawless: The Secret Rules That Govern Our Digital Lives (Cambridge: Cambridge University Press, 2019).

64 'Shadow ban' refers to a relatively common moderation practice of lowering a user's visibility, content or ability to interact without them knowing, so that they can continue to use the platform normally. See Courtney Radsch, 'Shadowban/Shadow banning', in Belli et al. (note 1 above).

65 Bernhard Rieder and Jeanette Hofmann, 'Towards platform observability', Internet Policy Review 9 (4) (2020), 1–28.

66 The lawsuit at the Brazilian Federal Supreme Court is available at: https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5160549&numeroProcesso=1037396&classeProcesso=RE&numeroTema=987

addition, the legislator could enact duties, such as the Digital Services Act,[67] for those with power and technical capacities to implement efficient monitoring of inappropriate content and risk assessment obligations, especially considering that content moderation technologies have improved with AI advances.[68]

It could also define a new civil liability scheme, which should ensure both an innovative ecosystem and legal certainty for enterprises, as well as duties and increased responsibilities for very large platforms, in harmony with new regulations that attempt to tackle issues derived from the unprecedented economic and political power of such actors. Despite this, as it is possible to conclude from the analysis of the 'Fake News' Draft Bill's most recent versions, Brazil did not make much progress in creating a governance model that would affect platforms' activities. As a result, the current version of the Draft Bill is not a moderate but a conservative piece of legislation that enables platforms to regulate themselves at will.

Russia

In recent years, Russia has adopted multiple restrictive normative provisions crafting a vision of 'Russian internet sovereignty',[69] consisting of provisions on personal data localisation, content regulation and a new type of 'infrastructure-embedded control',[70] and inspiring governments and legislators globally.[71]

67 Cf. European Commission, 'Questions and answers: Digital Services Act', available at: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348

68 Cf. Robert Gorwa, Reuben Binns and Christian Katzenbach, 'Algorithmic content moderation: technical and political challenges in the automation of platform governance', Big Data & Society 7 (1) (2020).

69 See A. Shcherbovich, 'Data protection and cybersecurity legislation of the Russian Federation in the context of the "sovereignisation" of the internet in Russia', in Belli, CyberBRICS (see note 12 above), pp. 67–131; F. Daucé and F. Musiani (eds), 'Infrastructure-embedded control, circumvention and sovereignty in the Russian Internet', First Monday 26 (5) (2021), available at: https://firstmonday.org/ojs/index.php/fm/issue/view/693

70 See Daucé and Musiani (note 34 above).

71 See e.g. Nigel Cory and Luke Dascoli, 'How barriers to cross-border data flows are spreading globally, what they cost, and how to address them', Information Technology & Innovation Foundation (2021), available at: https://itif.org/publications/2021/07/19/how-barriers-cross-border-data-flows-are-spreading-globally-what-they-cost/

From a liberal to a sovereignty-led approach

The main goal of recent digital policies adopted at the Russian level has been the establishment of an autonomous Russian segment of the internet, dubbed the 'Runet', allowing increased control on national digital infrastructures, and largely reproducing the strategies deployed by China since the early 2000s with the so-called 'Great Firewall of China'.[72]

Unlike China, however, the internet in Russia remained relatively free from regulation for more than a decade, with the introduction of light regulation in the mid-2000s. Only in recent times has Russia tightened its control on online media. While a certain degree of censorship has always existed, until the early 2010s Russia maintained a somewhat liberal[73] approach.

In 2006, Russia adopted Federal Law n. 149-FZ 'On Information, Information Technologies and Protection of Information', based mainly on the EU approach to intermediary liability, exempting intermediaries from civil liability related to UGC. Since the early 2010s, however, the initial liberal approach has been replaced by an increasingly heavy-handed approach.

Regulating terrorist content, fake news and insults to public officials

Since March 2019, Russia has moved towards a new content regulation regime aimed at regulating disinformation and restricting opinions on public authorities. Federal Law n. 31-FZ introduced the first set of provisions on Amending Article 15.3 of the Federal Law 'On Information, Information Technologies, and Protection of Information', 18 March 2019, aka the 'Fake News Law.' It prohibits publication of 'socially important information' and defines disinformation[74].

As pointed out by Shcherbovich.[75] the Explanatory Note to the Draft Bill states that the optimal way to implement it is by vesting the Prosecutor General of the

72 The Chinese approach led to the creation of an internet with Chinese characteristics that observers compare to an extensive national intranet connected to the global internet through limited channels.

73 Especially considering the media regulation during the Soviet era.

74 Disinformation is defined under the 'Fake News Law' as 'information of public interest, which is known to be unreliable, is disguised as accurate information and poses risks of harm to the life and/or health of citizens or property, mass disruption of public order and/or public safety, or impeding or halting the functioning of critical, transport or social infrastructures, lending institutions, or power generation, industrial or communications facilities'.

75 See A.A. Shcherbovich, 'Exploring the new Russian measures against "fake news" and online insults' (5 April 2019), available at: https://cyberbrics.info/exploring-the-new-russian-measures-against-fake-news-and-online-insults/

Russian Federation or his deputies with the power to request Roskomnadzor[76]—the Russian Media, Telecommunications and Information Regulator—to restrict access to information resources that disseminate disinformation. Hence, to implement the provisions, the Prosecutor General or his deputies request that Roskomnadzor orders providers to remove information within a specific deadline. Failing to comply with this request allows the authority to add the corresponding IP address to one of the state registers, obliging providers to block the IP address and prevent users from accessing the content.

Hovyadinov highlights the hybrid nature of Russian social media governance, as internet businesses with close ties to the government play a key role in conducting 'censorship and surveillance activities'.[77] These partnerships, enabled through state bodies' purchase of tech companies' shares, allow the federal government to count on the cooperation of intermediaries to control information flows and user activities. For example, a leading state bank, Sberbank, is a majority shareholder controlling Russian search engine and e-commerce giant Yandex, while email portal Mail.ru and the social media platform VKontakte are controlled by entrepreneurs closely affiliated with the Kremlin.[78]

Since 2019, Russia has limited the right to express 'disrespectful' opinions on public officials, society and symbols of the Russian Federation,[79] by passing Federal Law n. 30-FZ.[80] Under it, certain types of online content can be deemed illegal and taken down or blocked. Some cases do not even require a court order, thus allowing the government to directly instruct Roskomnadzor to request ISPs to block access to webpages or websites. After the Roskomnadzor notification to the ISPs, it must inform as to the content removal. Finally, Roskomnadzor verifies that the illegal content is inaccessible and tells the access providers to restore the access resource.

76 Roskomnadzor is the Russian Federal Service for Supervision of Communications, Information Technology and Mass Media. This executive agency is responsible for controlling and regulating all Russian mass media, including online media and internet networks, supervising compliance with data protection legislation, implementation of content regulation and telecoms law, and the operation of the Russian Autonomous Internet Subnetwork, better known as 'RuNet', in compliance with the Russian Sovereign Internet Law.

77 See Sergei Hovyadinov, 'Intermediary liability in Russia and the role of private business in the enforcement of state controls over the internet', in Giancarlo Frosio (ed.), Oxford Handbook of Online Intermediary Liability (Oxford: Oxford University Press, 2020).

78 Ibid.

79 The amendment prohibits 'the spreading of information which shows blatant disrespect for society, the government, official state symbols of the Russian Federation, the Constitution of the Russian Federation or authorities exercising governmental authority in the Russian Federation'.

80 'On Amendments to the Federal Law on Information, Information Technologies and Protection of Information.'

In June 2020, the European Court of Human Rights (ECtHR)—to which Russia was subject, as a Member of the Council of Europe, until September 2022—delivered a series of judgments assessing the implementation of Russia's Law on Information, Information Technologies, and Protection of Information. The ECtHR held that blocking entire websites was an extreme measure, which can be only justified in exceptional circumstances, as it is equivalent to banning a newspaper or a television station, having collateral effects on lawful content.[81] After the court rulings, the Duma introduced new amendments[82] to regulate platforms. These entered into force in February 2021, requiring social media platforms to monitor content and 'immediatelyly restrict access' to users that post information about state secrets, justification of terrorism or calls to terrorism; pornography; violence and cruelty; obscene language; drugs manufacturing; and information on methods to commit suicide, as well as calls for mass riots.

The consequences of the Ukraine war

Recent developments related to the Ukrainian war have had repercussions regarding online content regulation. First, the State Duma has adopted amendments to the Criminal Code of the Russian Federation, increasing responsibility for spreading 'fake news' about Russian Armed Forces' actions or calling for sanctions against Russia on social media.[83] They establish punishments with fines of 700,000 to 1.5 million roubles or imprisonment for up to three years. Moreover, if the illegal behaviour derived from 'abusing one's official position, based on political, ideological, racial, national or religious hatred or enmity, or based on hatred or enmity against any social group', then the term of imprisonment can be up to 10 years.

In addition, administrative sanctions and criminal liability might apply in case of 'public actions aimed at discrediting the exercise by state bodies of the Russian

81 See Gurshabad Grover and Anna Liz Thomas, 'Notes from a foreign field: the European Court of Human Rights on Russia's website blocking' (22 February 2021), available at: https://cyberbrics.info/notes-from-a-foreign-field-the-european-court-of-human-rights-on-russias-website-blocking/

82 Law 149-FZ, 'On Information, IT and Protection of Information'.

83 See State Duma of the Federal Assembly of the Russian Federation, 'Responsibility for the dissemination of fakes about the actions of the Armed Forces of the Russian Federation is introduced', available at: http://duma.gov.ru/news/53620/

Federation of their powers outside the territory of the Russian Federation'.[84] Special additional sanctions apply in cases of threat to 'public order',[85] where the Code of Administrative Offences foresees administrative fines of 50,000 to 100,000 roubles for individuals, from 200,000 to 300,000 roubles for officials, and from 500,000 to 1 million roubles for legal entities.

Since the early 2010s, the liberal approach has been replaced by an increasingly heavy-handed approach. Russia has amended its national framework on content regulation, introducing 'normative packages' to combat terrorism and preserve national sovereignty and, more recently, to regulate 'fake news', online insults to public authorities and war-related disinformation. The most recent amendments have confirmed a trend towards a stringent regime.[86]

India

A long line of rules and judicial decisions affect platform regulation in India, starting with the Information Technology Act of 2000 (IT Act[87]) and its subsequent rules. In 2021, a new set of rules concerning media intermediaries was enacted, the IT (Intermediary Guidelines and Digital Media Ethics Code) Rules. This scenario may soon change with the Digital India Act, currently being drafted by the Minister of State for IT and expected to be publicly debated in 2023.[88]

In terms of security and data protection concerns, the IT Act originally contained civil sanctions for 'cyber contraventions' (Section 43(a)–(h)) and criminal sanctions for 'cyber offences' (Sections 63–74). The Act was amended in 2008 to include Sections 43A ('Compensation for failure to protect data'[89]), 66A ('Punishment for sending offensive messages through communication

84 The fine will range from 100,000 to 300,000 roubles or imprisonment for up to three years. If these actions generate concrete consequences beyond the circulation of the disinformation, the maximum term of imprisonment is up to five years. See http://duma.gov.ru/news/53773/

85 'Calls for holding unauthorised public events, as well as posing a threat of harm to the life and [or] health of citizens, property, a threat of mass disruption of public order and [or] public safety, or a threat to interfere with the functioning or termination of the functioning of life support facilities, transport or social infrastructure, credit institutions, energy, industry or communications facilities.'

86 See State Duma of the Federal Assembly of the Russian Federation, 'Amendments on responsibility for fakes about the work of state bodies of the Russian Federation abroad were adopted', available at: http://duma.gov.ru/news/53773/

87 Amended in 2008.

88 PTI, 'Significant work done, draft Digital India Act framework by early 2023: MoS IT', The Hindu (6 November 2022), available at: https://www.thehindu.com/business/Economy/significant-work-done-draft-digital-india-act-framework-by-early-2023-mos-it/article66103357.ece

89 India, IT Amendment Act 2008 (5 February 2009).

service'[90]) and 72A ('Punishment for disclosure of information in breach of lawful contract'[91]).

Article 79 of the IT Act provides immunity to network service providers (meaning intermediaries) for UGC. This immunity is conditional on their due diligence (according to applicable rules) and participation solely as an intermediary. It is lost if the intermediary 'fails to expeditiously remove or disable access to that material' after having actual knowledge[92] or receiving a notification from a government agency. This provision was criticised at the time for casting too wide a net, potentially bringing liability to intermediaries conducting simple content moderation operations.[93]

The relevant provision on content-blocking is Section 69A,[94] which led to a judicial controversy between Twitter and the Ministry of Electronics & Information Technology (MeitY), whereby the company questioned the government's block notices on thousands of accounts.[95] It considers these orders procedurally and substantially flawed for not providing prior judicial review and hearings to content creators, besides failing to demonstrate the public interest necessity on a case-by-case basis.[96,97] Moreover, as noted by Bhandari (2022),[98] the interplay between Section 69A and the IT (Procedure and Safeguards for Blocking for Access of Information by Public) Rules of 2009, as interpreted by the government,

90 Ibid.

91 Ibid.

92 The Indian Supreme Court clarified the meaning of 'actual knowledge' in Shreya Singhal v. Union of India, which addressed 'the issue of intermediaries complying with takedown requests from non-government entities and has made government notifications and court orders to be consistent with reasonable restrictions in Article 19(2)'. Jyoti Panday, 'The Supreme Court judgment in Shreya Singhal and what it does for intermediary liability in india?', Centre for Internet and Society (11 April 2015), available at: https://cis-india.org/internet-governance/blog/sc-judgment-in-shreya-singhal-what-it-means-for-intermediary-liability

93 This immunity can be guaranteed according to their due diligence (following applicable rules) and their participation solely as an intermediary (i.e. without 'select[ing] or modify[ing]' the information). It is lost if the intermediary 'fails to expeditiously remove or disable access to that material' after having actual knowledge or receiving a notification from a government agency. This provision was criticised at the time for casting too wide a net, potentially bringing liability to intermediaries conducting simple content moderation operations. Pranesh Prakash, 'Short note on IT Amendment Act, 2008', Centre for Internet and Society (February 2009), available at: https://cis-india.org/internet-governance/publications/it-act/short-note-on-amendment-act-2008

94 India, Section 69A in the Information Technology Act, 2000 (2000), available at: https://indiankanoon.org/doc/10190353/

95 ETech, 'Twitter-ministry hearing in Karnataka HC adjourned till August 25', Economic Times (26 July 2022), available at: https://economictimes.indiatimes.com/tech/technology/twitter-ministry-hearing-in-karnataka-hc-adjourned-till-august-25/articleshow/93129940.cms

96 Vrinda Bhandari, 'Twitter case underlines web moderation issues', Hindustan Times (8 July 2022), available at: https://www.hindustantimes.com/opinion/twitter-case-underlines-web-moderation-issues-101657209298117.html

97 Saptaparno Ghosh, 'Twitter's petition on Section 69A of the IT Act', The Hindu (12 July 2022), available at: https://www.thehindu.com/sci-tech/technology/twitters-petition-on-section-69a-of-the-it-act/article65623202.ece

98 Bhandari (see note 61 above).

creates an opaque system whereby content creators face an 'arduous legal process to first try and secure a copy of the blocking order and then challenge it'.

From the free speech perspective, one highlight is the decision in *Shreya Singhal v. Union* of India,[99] whereby the court declared Section 66A unconstitutional under article 19(1)(a) of the Indian Constitution.[100] The court found that the section's vagueness in terms such as 'annoyance' and 'inconvenience' could create a chilling effect over a 'large amount of protected and innocent speech' (para. 83).

More recently, the 2021 intermediary Rules[101] have raised attention in platform regulations. The Rules create due diligence duties for social media intermediaries and 'significant'[102] social media intermediaries, thus specifying the conditions of liability immunity for these actors.

Among the due diligence obligations in the 2021 Rules, intermediaries are required to publish monthly grievance reports and to appoint a Chief Compliance Officer, a Grievance Officer and a Nodal Contact Person, all residing in India.[103] Furthermore, the 2021 Rules demanded that intermediaries implement 'content takedown within tight deadlines [Rule 3(1)(d)], automated content filtering [Rule 4(4)] and voluntary identification of users on social media intermediaries [Rule 4(7)]'[104] They also enacted a traceability obligation,[105] which was criticised for its potential to break end-to-end encryption in messaging applications.

Although the Indian government has proposed two models that allegedly allow this traceability obligation without disclosing the content of messages,

99 Shreya Singhal v. Union of India (March 2015), Columbia Global Freedom of Expression, available at: https://globalfreedomofexpression.columbia.edu/cases/shreya-singhal-v-union-of-india/

100 Aditi Subramaniam and Sanuj Das, 'In a nutshell: data protection, privacy and cybersecurity in India', Lexology (22 October 2020), available at: https://www.lexology.com/library/detail.aspx?g=04c38a97-f6cb-4d23-ae95-00df33df8a68

101 India, Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules (2021), available at: https://egazette.nic.in/WriteReadData/2021/225464.pdf

102 Distinguished by the number of users in India, according to a threshold determined by the central government (currently it is set at 5 million or more registered users), Notification S. O. 942(E), Pub. L. No. S. O. 942(E), Gazette of India (2021), available at: https://egazette.nic.in/WriteReadData/2021/225497.pdf

103 Twitter, for example, had problems when the rules were enacted. ET Bureau. 'Twitter now in compliance with IT rules, govt tells court', Economic Times (10 August 2021), available at: http://www.ecoti.in/KAdCwb47; Alnoor Peermohamed, 'Delhi High Court gives Twitter "last opportunity" to show compliance with IT rules', Economic Times (28 July 2021), available at: http://www.ecoti.in/c0hCoZ; Alnoor Peermohamed, 'Twitter lost immunity under IT Act: Centre to HC', Economic Times (6 July 2021), available at: http://www.ecoti.in/3OGSqY

104 Neeti Biyani and Amrita Choudhury, 'Internet Impact Brief: 2021 Indian Intermediary Guidelines and the internet experience in India', Internet Society (8 November 2021), available at: https://www.internetsociety.org/resources/2021/internet-impact-brief-2021-indian-intermediary-guidelines-and-the-internet-experience-in-india/

105 To identify the 'first originator' (in India) of certain information shared through an intermediary's messaging application (such as WhatsApp or Signal) [Rule 4(2)].

these might require breaking of end-to-end encryption, nonetheless abandoning forward secrecy or simply being based on faulty assessments of how encrypted messaging applications work.[106] In addition, the rule has been criticised[107] for raising other operationalising costs—particularly data storage to trace every message on a messaging thread—thus increasing barriers for smaller competitors. Another provision pointed out as problematic for similar reasons is Rule 4(4), which requires client-side scanning for matches against certain types of material (e.g. rape or child sexual abuse material)—an intrusive manner of content control and not necessarily practical.[108]

Finally, Rule 3(1)(d) of the 2021 Rules requires content removal upon court order or governmental notice[109] in up to 36 hours. This provision aims to expedite content removal related to various subjects listed in the rule. However, in doing so it creates a wide net of hypotheses for content removal based on open-ended juridical terms such as 'public order' and 'incitement to an offence' and subjective terms such as 'decency' and 'morality'.

In summary, the 2021 Rules have been criticised for conflicting with the IT Act whence they come[110] and, through vague wording, creating space for arbitrariness. They also came under scrutiny for establishing obligations to implement technical procedures that have been widely regarded as incompatible with end-to-end encryption and data privacy, potentially creating a harmful chilling effect over legitimate forms of speech and exposing minority and sensitive political groups to risks online. Criticism over the traceability rule went beyond simple discourse: WhatsApp and the Foundation for Independent Journalism have filed suits questioning the IT Rules 2021's constitutionality and legality,

106 Biyani and Choudhury (see note 69 above); Namrata Maheshwari and Greg Nojeim, 'Part 2: New intermediary rules in India imperil free expression, privacy and security', Center for Democracy & Technology (4 June 2021), available at: https://cdt.org/insights/part-2-new-intermediary-rules-in-india-imperil-free-expression-privacy-and-security/; Riana Pfefferkorn, R., 'New intermediary rules jeopardize the security of Indian internet users', Brooking's TechStream (2021), available at: https://www.brookings.edu/techstream/new-intermediary-rules-jeopardize-the-security-of-indian-internet-users/

107 Biyani and Choudhury (see note 69 above).

108 This case brings to mind Apple's plan, revealed in 2021, to implement a similar mechanism on its iOS devices, which was promptly dropped after it was heavily criticised for infringing users' privacy and putting sensitive information at risk. See India McKinney and Erica Portnoy, 'Apple's plan to "think different" about encryption opens a backdoor to your private life', Electronic Frontier Foundation (5 August 2021), available at: https://www.eff.org/deeplinks/2021/08/apples-plan-think-different-about-encryption-opens-backdoor-your-private-life

109 '[U]pon receiving actual knowledge in the form of an order by a court of competent jurisdiction or on being notified by the Appropriate Government or its agency.'

110 N. Behera, 'Legal protection of right to privacy in cyberspace' (2020); Biyani and Choudhury (see note 69 above).

respectively.[111] All this comes on top of an already contentious system of content-blocking and liability for content posted, with open concepts that give rise to curtailments on speech based on deficient procedural checks and balances—all of which have been or are currently under litigation.

China

Over the past two years, China has considerably updated its cyberspace regulations. For example, it adopted the Provisions on the Governance of the Online Information Content Ecosystem in 2020, the Data Security Law (DSL, effective in September 2021) and the new Personal Information Protection Law (PIPL, effective in November 2021). In terms of cybersecurity, these build on the foundations established by the 2017 Cybersecurity Law (CSL). Taken together, they create a comprehensive cybersecurity framework,[112, 113]

In January 2022, a regulation on algorithmic recommendation systems, published for comments in 2021,[114] was adopted.[115] Press announced the algorithmic recommendation regulation as 'pioneering' and 'groundbreaking',[116] and it seems so: the closest existing norm at the time of its discussion would be the

111 Surabhi Agarwal, S., 'WhatsApp sues Government of India over new IT rules', Economic Times (29 May 2021), available at: https://economictimes.indiatimes.com/tech/technology/whatsapp-sues-india-govt-says-new-it-rules-mean-end-to-privacy/articleshow/82963637.cms; Joseph Menn, 'WhatsApp sues Indian government over new privacy rules', Reuters (26 May 2021), available at: https://www.reuters.com/world/india/exclusive-whatsapp-sues-india-govt-says-new-media-rules-mean-end-privacy-sources-2021-05-26/; The Wire Staff, 'Why The Wire wants the new IT rules struck down', The Wire (9 March 2021), available at: https://thewire.in/media/why-the-wire-wants-the-new-it-rules-struck-down

112 Dehao Zhang, 'China: The interplay between the PIPL, DSL, and CSL', DataGuidance (April 2022), available at: https://www.dataguidance.com/opinion/china-interplay-between-pipl-dsl-and-csl

113 Belli, 'Cybersecurity policymaking in the BRICS countries' (see note 12 above).

114 China, 'Notice of the state Internet Information Office on the provisions on the administration of internet algorithmic recommendation (draft for solicitation of comments)', Cyberspace Administration of China (27 August 2021), available at: http://www.cac.gov.cn/2021-08/27/c_1631652502874117.htm

115 Rogier Creemers, Graham Webster and Helen Toner, 'Translation: Internet Information Service Algorithmic Recommendation Management Provisions', Digichina (1 March 2022), available at: https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/; Harry Chambers and Julian Sun, 'China: The Internet Information Service Algorithm Recommendation Management Regulations', DataGuidance (March 2022), available at: https://www.dataguidance.com/opinion/china-internet-information-service-algorithm

116 Shen Lu, 'Chinese tech companies now have to tell users about their algorithms', Protocol (1 March 2022), available at: https://www.protocol.com/bulletins/china-algorithm-rules-effective; Yan Luo, Vicky Liu and Irina Danescu, 'China takes the lead on regulating novel technologies: new regulations on algorithmic recommendations and deep synthesis technologies', Covington Inside Privacy (8 February 2022), available at: https://www.insideprivacy.com/artificial-intelligence/china-takes-the-lead-on-regulating-novel-technologies-new-regulations-on-algorithmic-recommendations-and-deep-synthesis-technologies/; Helen Toner, Paul Triolo and Rogier Creemers, 'Experts examine China's pioneering draft algorithm regulations', Digichina (27 August 2021), available at: https://digichina.stanford.edu/work/experts-examine-chinas-pioneering-draft-algorithm-regulations/

United Kingdom's algorithmic transparency standard.[117] This rule provides a useful example of the Chinese strategy towards platform regulation—containing strong bureaucratic, content and technical controls, in a fashion similar to other specific regulation and to the more general 'Internet Information Service Management Rules' and 'Provisions on the Governance of the Online Information Content Ecosystem'.[118] Due to its novelty and specificity, as well as the growing importance of algorithmic recommendation systems underlying the operations of digital platforms of various kinds, it merits a detailed description.

The regulation defines algorithmic recommendation systems as 'the use of generative or synthetic-type, personalised recommendation-type, ranking and selection-type, search filter-type, dispatching and decision-making-type, and other such algorithmic technologies to provide information to users' (art. 2). As such, it covers a wide array of standard practices in digital platforms' activities—content recommendation, ranking, selection, search filters and others.

Some highlights, divided by the authors into broader thematic categories below, are as follows.

1. **Platforms' duties:**
   - 1.1. Duty to mark algorithmically generated or synthetic information before dissemination (art. 9);
   - 1.2. Duty to remove unlawful or harmful information, preserve records and alert cybersecurity and other competent authorities (art. 9);
   - 1.3. Control of algorithmic processes to the level of the tagging of user profiles/models, which shall avoid unlawful or harmful keywords (art. 10);
   - 1.4. Duty to establish systems of manual intervention by users in algorithmic recommendation processes directed at them and to promote 'autonomous user choice' (art. 11);
   - 1.5. Duty of transparency and understandability concerning algorithmic recommendation processes (art. 12);
   - 1.6. Duty to provide users with complaint and reporting mechanisms (art. 22);

117 CDDO, 'Algorithmic Transparency Reports' (29 November 2021), available at: https://www.gov.uk/government/collections/algorithmic-transparency-standard

118 China, 'Internet Information Service Management Rules' (25 September 2000), available at: https://chinacopyrightandmedia.wordpress.com/2000/09/25/internet-information-service-management-rules/; Cyberspace Administration of China, 'Provisions on the Governance of the Online Information Content Ecosystem' (15 December 2019), available at: https://wilmap.stanford.edu/entries/provisions-governance-online-information-content-ecosystem

1.7. A general prohibition of various behaviours enabled by algorithms, such as account and likes/comments/shares manipulation (seemingly aimed at bot activity) and manipulative administration of listings and topics to influence public opinion (art 14);
1.8. A general prohibition on anti-competitive behaviours enabled by algorithms (art. 15);

2. **User rights:**
2.1. Notification and information about algorithmic recommendation systems in use (art. 16, with special protection of minors and the elderly in arts 18 and 19 respectively);
2.2. Granular control over algorithmic recommendation services, including the capacity to choose and delete user tags (art. 17);
2.3. Special protection to workers in labour relations intermediated by algorithmic services, upholding interests 'such as obtaining labour remuneration, rest and vacation, and others' (art. 20);
2.4. Special protection to consumers in consumer relations, hinting at predatory marketing practices ('they may not use algorithms to commit acts of extending unreasonably differentiated treatment in trading conditions such as trading prices, and others', art. 21);
2.5. Duty to provide users with complaint and reporting mechanisms (art. 22);

3. **Content control:**
3.1. A general duty to prevent harmful content (various articles), including through active technical measures such as 'content de-weighting, scattering interventions, and others' (art. 12);
3.2. Requirement of a permit for news information services, accompanied by a fake news prohibition: 'They may not generate or synthesise fake news information, and may not disseminate news information not published by work units in the State-determined scope' (art. 13);

4. **Security measures:**
4.1. Duty to establish security plans, incident response processes and regular revisions of algorithms (art. 8);
4.2. Graded and categorised algorithm security management system (art. 23);

4.3. Exceptional cybersecurity and reporting/registering duties for 'algorithmic recommendation services with public opinion properties or social mobilisation capabilities' (arts 24–27);
4.4. Cybersecurity assessments by authorities and a duty to preserve network records 'according to the law' (art. 28).

All these provisions are notable for either their qualities or their defects. Some provisions are too general (e.g. arts 6, 10, 14, 17, 21), becoming possibly over-inclusive and thus potentially harmful to innovative efforts, day-to-day operations of the regulated firms or users' rights, such as free speech. Others are strongly pro-user and go into minutiae of the realisation of user rights (arts 16–22), revealing a clear view of how these algorithmic systems work and how their adverse effects might be halted or mitigated. Finally, other provisions seem aspirational or closer to public policy aims, such as observing 'science and reason, and sincerity and trustworthiness' (art. 4) and advancing the use of algorithms 'in the direction of good' (art. 6).

Overall, the regulation touches upon many subjects involved in using algorithmic systems. Its preoccupation with the generation of addiction and excessive consumption (arts 8 and 18) resonates with studies of the addictive effects of social media recommendation systems. Its inclusion of particular mention of the rights of workers mediated by algorithms (art. 20) seems to recognise potential vulnerabilities in the algorithmic labour organisation. Its inclusion of fake news (art. 12) and manipulative practices, including false likes, comments and shares (art. 14), echoes some pervasive practices that threaten political systems worldwide. Finally, the inclusion of granular user control of algorithmic systems, including the possibility to outright deactivate those systems (art. 17), marks a strong position in empowering users in the face of data-intensive digital platforms.

On the other hand, the regulation contains several references to state control of news media (art. 13) and overly broad provisions and insufficiently defined terms (e.g. art. 6, 'mainstream value orientations', 'positive energy'), and does not go into detail on the administrative structure that will be needed to operate the level of control the regulation aims to implement. Moreover, the use of broad language in the definition of controlled content—including encouraged internet

content—follows the previous tendency set by the Provisions on the Governance of the Online Information Content Ecosystem[119] enacted in 2020.

Overall, the Chinese framework provides interesting case studies for Western legislators[120] in terms of dos and don'ts. The incisiveness in dealing with the technical details of algorithmic recommendation systems and their societal and economic consequences demonstrates possible strategies for dealing with the harms of surveillance capitalism and the attention economy. However, the use of overly broad legal terms, especially concerning content control, and the lack of an independent regulator implementing the provisions may lead to frameworks considerably unaligned with the West's paradigm on due process, necessity and proportionality in the case of restrictions to speech.

South Africa

Regarding the legal and regulatory environment for intermediary liability in South Africa, Zingales[121] points out that the Republic of South Africa's Constitution enacts equality, dignity, freedom and advancement of human rights as its central values. In its democratisation process, the country prioritised promoting equality, stating it in several legal provisions such as the Promotion of Equality and Prevention of Unfair & Discrimination Act (PEPUDA), from 2000. This Act also defines hate speech, binding application providers to combat explicitly hateful content and content that might 'be reasonably construed to have a clear intention to be hurtful'.[122] In this context, the DoC called, in 1999, for laws regarding intermediaries' liability, pointing out concerns about their roles in disseminating or allowing unlawful content. According to Zingales,[123] this led to a public consultation resulting in the Electronic Communications and Transactions Act (ECTA), passed in 2001.

119 Bolin Zhang and Joan Barata, 'Provisions on the Governance of the Online Information Content Ecosystem', Wilmap (1 March 2020), available at: https://wilmap.stanford.edu/entries/provisions-governance-online-information-content-ecosystem

120 Tom Wheeler, 'China's new regulation of platforms: a message for American policymakers', Brookings (14 September 2021), available at: https://www.brookings.edu/blog/techtank/2021/09/14/chinas-new-regulation-of-platforms-a-message-for-american-policymakers/

121 See Nicolo Zingales, 'Internet intermediary liability: identifying best practices for Africa', Association for Progressive Communications (26 November 2013), available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2359696

122 Nicolo Zingales, 'Intermediary liability in Africa: looking back, moving forward?', in Giancarlo Frosio (ed.), Oxford Handbook of Online Intermediary Liability (Oxford: Oxford University Press, 2020), 213–235: 216.

123 Ibid., p. 217.

ECTA is considered 'to date, the most articulate framework for dealing with intermediary[124] liability in Africa'.[125] Its development aimed explicitly to deal with the growth of e-commerce in the country, promoting legal certainty for enterprises with safe harbours similar to those present in the US's DMCA and the EU's E-Commerce Directive. It establishes that the law cannot require a service to actively monitor data, facts or circumstances indicating unlawful activity. Nevertheless, it limits liability to two additional requirements: '(1) the intermediary's membership of an industry representative body (IRB); and (2) adoption and implementation of the corresponding code of conduct'.[126] Furthermore, the Minister of Communications issued a document titled 'Guidelines for recognition of industry representative bodies of Information System Service Providers' in 2006, which integrates the code of conduct requirement. It states that 'the only monitoring or control done by the State … is to ensure that the IRB and its ISPs meet certain minimum requirements'.[127]

Despite such advanced provisions, intermediaries have been in relative juridical uncertainty, frequently subject to injunctions and lawsuits under criminal law for their users' behaviours. Moreover, in addition to the failures in the safe harbours application, the hopes for building a more democratic social media governance are now on hold with the approval of several laws that constitute the so-called Internet Censorship Bill, as explored below.

Social media legislation and the 'Internet Censorship Bill'

The primary South African legislation for regulating online content is the Film and Publications Act (FPA), 1996. The enactment of such a law aimed to repeal prior legislation that aimed at censoring cultural productions in the apartheid context. It also established the Film and Publications Board (FPB) to receive complaints or applications to evaluate the classification of cultural production regarding its suitability for an audience.

By the end of 2019, South African President Cyril Ramaphosa signed an amendment to the FPA (aka the FPAA), dubbed the 'Internet Censorship Bill' by opponents. The new version of the bill shifts the intermediary liability completely

124 Which is defined as 'any person providing information system services' by its Chapter IX.

125 Ibid.

126 Ibid., p. 8.

127 Zingales (see note 86 above), p. 11.

by imposing new duties and obligations on ISPs, which become obliged to monitor illicit, abusive and harmful content, such as child exploitation and abuse imagery, war propaganda, incitement to violence and hate speech. If the ISP fails to remove such content promptly, it could suffer sanctions such as fines of up to ZAR 50,000 (approximately 3,200 USD) and even imprisonment for six months. The bill also establishes criminal provisions for individuals that distribute prohibited content.

In addition, FPAA changes the role of the FPB, transforming it from a classification authority into a full regulator, with powers to renew or not the certificates of its applicants and request them to submit their content for evaluation. The FPB is also allowed, under the FPAA—which started to take effect in March 2022—to issue takedown notices for ISPs regarding potentially prohibited content. But experts at the ISPA[128] have been pointing out the possible censorship nature of FPB. Furthermore, they highlight the possible impacts of such measures given that the body does not have the same capacities as the courts for weighing rights.

Other relevant legislation

a) Cybercrimes Act

The Cybercrimes Act, passed in May 2021, aims at tackling harmful speech in the online environment, including incitement of violence and other harms. It designates several specific offences as cybercrimes and criminalises 'malicious communication', such as sending data messages with violence, threats, harm or non-consensual intimate imagery.

(b) South African Disaster Management Regulations

The Covid-19 pandemic led the South African government to declare a 'state of disaster' in March 2020, enacting the 'Disaster Management Act', which criminalises disinformation. However, according to a report by Mawarire to USAID,[129]

128 Freedom House. Freedom on the Net 2022: South Africa. (2022) available at: https://freedomhouse.org/country/south-africa/freedom-net/2022

129 Teldah Mawarire, '"Things will never be the same again": Covid-19 effects on freedom of expression in Southern Africa', 2020 Research Report' (2020), available at:https://internews.org/wp-content/uploads/2021/02/Internews_Effects_COVID-19_Freedom_of_Expression_Southern_Africa_2020-12.pdf

'the Act had been amended at least three times within a month, making it difficult for ordinary citizens to interpret it'. In addition, article 19 has pointed out some concerns with such measures, highlighting that it could be 'a dangerous trend of countries using the Covid-19 pandemic to enforce disinformation laws in the region'.[130]

Conclusion: Choosing between a sledgehammer and a scalpel to regulate content

With the increase in digital platforms' impact on political and economic systems, the BRICS countries are establishing regulations to tackle malicious activities and unlawful content, establishing intermediary obligations for transparency and accountability. The enactment of laws geared explicitly towards digital platforms aims to reassert state sovereignty in the online environment, preserving the stability and security of the national political infrastructures. Nevertheless, historical institutional complexities and disputes affect how regulations and approaches are shaped and chosen for such sensitive matters.

Not surprisingly, we can remark on an inevitable overlap between the cultural specificities of the country at stake and its approach to content regulation. In Brazil, Draft Bill 2,630/2020 is extremely moderate because the Brazilian democratic model is historically sceptical towards media regulation and relatively permeable to lobbying from private companies. While the Brazilian legislature might want to avoid using a sledgehammer to tackle disinformation with strict legislation, the proposed framework so far has failed to propose an effective scalpel to fight disinformation in a surgical fashion.

Interestingly, the Russian online content-blocking regime is remarkably similar to the Indian regime defined in Section 69A of the Information Technology (IT) Act. While no official document explicitly acknowledges the Russian influence, it is safe to assume knowledge of the Russian system on the part of India (and other BRICS countries), given the existence of a specific intra-BRICS body

130 'South Africa: prohibitions of false COVID-19 Information must be amended', Article 19 (23 April 2021), available at: https://www.article19.org/resources/prohibitions-of-false-covid-information-must-be-amended/

for information exchange on cybersecurity for almost eight years. However, both regulatory frameworks have been criticised for their tendency towards a sledgehammer approach to platform regulation, which may easily be abused.

The Chinese approach seems to be the most coherent and structured, as well as the most innovative. While it adopts a rigorous approach to content regulation, it offers valuable food for thought regarding what practical measures can be considered and the difficulty that legislators might have in regulating disinformation effectively without engaging in draconian norms. The South African approach is an example of how even countries that are internationally renowned for their commitment to democracy and human rights, and strive to elaborate a well-articulated framework to regulate content properly, will inevitably end up being criticised for censorship.

Despite the divergences in the BRICS online content regulations, some common trends can be highlighted. First, almost all countries are drafting or have passed legislation outlawing specific types of online content and frequently defining transparency obligations, from moderate laws such as the Brazilian to stricter ones such as the Chinese. Duties of care are present in most legal frameworks. The oversight mechanism allowing the implementation of the content regulation provisions is usually an administrative procedure. As such, this governance model may lead to concerns regarding the independence of the process and the proportionality and full respect of rule-of-law criteria, especially when the administrative body competent for the oversight is not an independent body.

To conclude, we provide the reader with a visual representation[131] of the primary norms regulating online content, the type of content deemed illegal and the bodies competent to implement the regulatory framework in each BRICS country. The regulatory choices of the BRICS members will naturally influence the countries' regional neighbours, but these frameworks should also be carefully analysed by non-BRICS nations struggling with similar issues. While the BRICS have long been transplanting Western policy elements in their national frameworks, some of the BRICS countries are among the most 'experienced' regarding content regulation. Their experiences offer valuable insights into what could, should or should not be reproduced by others.

131 A detailed visual representation of the online content normative frameworks of the BRICS countries can be found at https://cyberbrics.info/map-online-content-normative-frameworks-in-the-brics/

CHAPTER 3

‘We are not quite there yet’

The Latin-American narrative regarding cyber-norms development

MARIA PILAR LLORENS

Introduction

Since the late 1990s the United Nations (UN) has been addressing how cyberspace should be regulated. It was not until recently that a consensus regarding international law’s application to cyberspace was reached.[132] However, the question of ‘the specific interpretation and application of the

132 For example, in 2021 the consensus was reinforced by the UN Group of Governmental Experts (GGE) and the Open Ended Working Group (OEWG). Their final reports restated that the existing international legal framework applies to cyberspace. UN General Assembly, ‘Open-Ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security. Final Substantive Report’, UN Doc. A/AC.290/2021/CRP.2, 10 March 2021, available at: https://front.un-arm.org/wp-content/uploads/2021/03/Final-report-A-AC.290-2021-CRP.2.pdf; UN General Assembly, ‘Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security’, UN Doc. A/76/135, 14 July 2021, available at: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/075/86/ PDF/N2107586.pdf

norms of International Law to cyberspace and cyber operations'[133] remains unanswered.[134]

This consensus, and hence the development of cyber norms, was the result of a long (and sometimes difficult) process. The debates were shaped by states' different and usually incompatible stances on cyberspace governance and international law's application in this domain. As with other fields of international law, the debates have been led by states, and also by scholars, from the Global North (GN). As a result, their voices have expressed the prevailing, if not the only, views. Hence the narrative from the GN is the one that matters. This could prove problematic as it could be perceived (once again) as 'global law made by the West'.[135]

While it could be argued that during the past five years more states from the Global South (GS) have been taking part in the debate,[136] broader participation does not necessary mean that more voices are heard. Furthermore, participation does not ensure that new or different narratives are being proposed by other states. Latin American states are an example in this regard. Their participation has increased (at both the UN and regional levels), and they have engaged in the ongoing debates. However, they do not seem to be bringing anything new to the table. They have released hardly any public document highlighting their position, and have privately acknowledged that they face difficulties that prevent them from engaging meaningfully.[137]

I argue that what is partly behind these difficulties that Latin American states face is an appropriation of narrative from the GN. This in turn silences peripheral voices, mostly the GS. I maintain that the international legal framework in

133 François Delerue, Cyber Operations and International Law (Cambridge: Cambridge University Press, 2020), 2.

134 Among others: Dennis Broeders, Els de Busser, Fabio Cristiano and Tatiana Tropina, 'Revisiting past cyber operations in light of new cyber norms and interpretations of international law: inching towards lines in the sand?', Journal of Cyber Policy 7 (1) (2022), 97–100; Dapo Akande, Antonio Coco and Talita de Souza Dias, 'Drawing the cyber baseline: the applicability of existing international law to the governance of information and communication technologies', International Law Studies 99 (2022), 6; Dapo Akande, Antonio Coco, Talita de Souza Dias, Duncan Hollis, James O'Brien and Tsvetelina van Benthem, 'The Oxford Statement on International Law Protections in Cyberspace: The Regulation of Information Operations and Activities' (EJIL: Talk!, 2 June 2021), available at: https://www.ejiltalk.org/the-oxford-statement-on-international-law-protections-in-cyberspace-the-regulation-of- information-operations-and-activities/

135 Brian-Vincent Ikejiaku, 'International law is Western made global law: the perception of third-world category', African Journal of Legal Studies 6 (2–3) (2013), 341.

136 For example, see the discussion of African Union (AU) and ASEAN engagement with the ongoing debate in Irene Poetranto, Justin Lau and Josh Gold, 'Look South: challenges and opportunities for the "rules of the road" for cyberspace in ASEAN and the AU', Journal of Cyber Policy 6 (3) (2021), 318–339.

137 Inter-American Juridical Committee, 'Improving Transparency: International Law and State Cyber Operations. Fifth Report', CJI/doc. 615/20 rev. 1, 7–8 paras 17–21 (Organization of American States, 7 August 2020), available at: https://www.oas.org/en/sla/iajc/docs/themes_recently_ concluded_International_law_State_cyber_operations_FINAL_REPORT.pdf

this field is, once again, being shaped by the GN's interest while the GS is following behind.[138]

In this work I use the framework provided by Third World Approaches to International Law (TWAIL). This framework helps to understand how the GN's interests prevail in the construction of the narrative regarding cyber-norms development. TWAIL critique highlights how international law is used to 'protect, project, promote (3Ps) or to safeguard the interest'[139] of the GN. Moreover, TWAIL helps to understand how actors from the GS can provide their own narrative and make themselves heard on this topic. In this contribution I draw from Latin American states' narrative regarding cyberspace regulation, built on official documents released by Latin American states and also on their statements at the Organization of American States (OAS)—particularly the reports of the Inter-American Juridical Committee (IAJC) on the topic[140]—and at the UN. In line with the proposed framework, the contribution aims to understand whether Latin American states have room to develop their own narrative and to contribute to the international cyber-norms production process.

In this work, narrative is understood as 'stor[ies] that order experience, render experience meaningful or tentatively explain acts and events, for a particular audience'.[141] The documents presented by states and their statements at meetings at global and regional levels can be read as narratives. One limitation that I have encountered is the unavailability of official documentation. As a consequence, this work heavily relies on the reports of the IAJC on the question of international law's application to/in cyberspace. Both reports highlight that Latin American states do not have official positions on the topic. However, where available these official documents are referred to.

To address this issue, this contribution is organized in five sections. The following section aims to explain how the (global) narrative regarding cyberspace regulation has been (and still is) constructed. I argue that the GN's interest prevails in the development of cyberspace norms. Usually, this normative

138 On the contrary, some authors maintain that AU and ASEAN member states have been key stakeholders in the process; see Poetranto et al. (note 5 above).

139 Ikejiaku (see note 4 above).

140 CJI/doc.615/20/rev. 1 and also Inter-American Juridical Committee, 'International law applicable to cyberspace', CJI/doc.671/22 rev. 2 (OAS, 24 August 2022), available at: http://www.oas.org/es/sla/cji/docs/CJI-doc_671-22_rev2_ESP.pdf (Spanish version).

141 Anette Bringedal Houge, 'Narrative expressivism: a criminological approach to the expressive function of international criminal justice', Criminology & Criminal Justice 19 (3) (2019), 279.

development follows certain events that affect GN states (e.g. the Estonian cyber-attack in 2007 or the interference with US presidential elections in 2016).

I then examine Latin American experience concerning normative development in cyberspace. In this section I look into how Latin American states have addressed the topic, particularly at the regional level. These states have an emerging interest in developing a distinct narrative regarding international law's application in/to cyberspace.

The following section looks into the question of Latin American states' likelihood of providing their own narrative on the topic. I maintain that whereas Latin American states show interest in developing their own narrative, several factors prevent them from achieving this goal. As a result, Latin American states are trying to play catch-up, but while doing so they are being absorbed by the GN's narrative.

This chapter concludes that despite their early engagement with the process of development of international regulation for cyberspace, Latin American states are still struggling to provide their own narrative, mainly because they have yet to develop their own narrative, one that detaches from the hegemonic narrative of the GN.

Cyberspace norms, narratives and the Global North

The debates on cyberspace regulation have been ongoing for almost two decades. How these debates have evolved is relevant because they have the potential to promote states' own narratives regarding cyber-norms development to protect their interests.[142] I argue that during this ongoing debate the GN's attitude, and hence its narrative, has shifted from expressing almost non-interest in the question to becoming the champion of the cyber-norms development process. The GN narrative can be distinguished in two periods. The first (1998–2007) is characterised by GN's (almost) lack of interest in cyberspace norms development. The second (2007–2022), on the contrary, is characterised by the emergence of

142 Eneken Tikk-Ringas, 'International cyber norms dialogue as an exercise of normative power', Georgetown Journal of International Affairs 17 (3) (2016), 47–59: 48.

GN interest in cyberspace regulation, where GN states have developed and reinforced its narrative regarding international law's application to cyberspace.

During the first phase, the GN lacked the interest to engage in the public dialogue initiated by the UN General Assembly (UNGA) in 1998.[143] Many GN states even disputed UNGA's competence to discuss normative development in cyberspace.[144] In this period GN states favoured a laissez-faire view of the internet, as they were interested in having as little regulation as possible.[145] As a result, the predominant narrative was that the development of international norms for responsible state behaviour in cyberspace was not necessary.

The second phase was triggered by the Estonian cyber-attack of 2007.[146] As a consequence of this, GN states changed their attitude to cyberspace regulation: the GN abandoned its earlier reluctance to address the topic and started to actively promote the normative development process. GN states began to advance their own understanding of how cyberspace regulation should be. For instance, until today, the development of certain cyber norms is closely related to major cyber incidents affecting GN states.

An example can be seen in the treatment of the use of force in cyberspace. This subject gained momentum after the Estonian cyber-attack and the cyber-attacks

143 In 1998 the Russian Federation submitted to the First Committee of UNGA a draft resolution that called on member states to consider at multilateral level the 'existing and potential threats in the field of information security'. It cited concerns on the emergence of 'information weapons and the threat of information wars'. Russian Federation, letter dated 23 September 1998 from the Permanent Representative of the Russian Federation to the United Nations addressed to the Secretary-General, UN Doc. A/C.1/53/3, 2 (30 September 1998) available at: https://digitallibrary.un.org/ record/261158?ln=en

144 For example, submission of the United States in UN Secretary General, Developments in the Field of Information and Telecommunications in the Context of International Security, UN Doc. A/54/213, 11 (10 August 1999); submission of Sweden on behalf of the European Union in UN Secretary General, Developments in the Field of Information and Telecommunications in the Context of International Security, UN Doc. A/RES/56/164, 4 (3 October 2001). It was also noted that the GN response was a reaction to Russian interest in establishing some limits to (informatics) weapons development: Eneken Tikk and Mika Kerttunen, The Alleged Demise of the UN GGE: An Autopsy and Eulogy (New York: Cyber Policy Institute, 2017), 9, available at: https://cpi.ee/wp-content/uploads/2017/12/2017-Tikk-Kerttunen-Demise-of-the-UN-GGE-2017-12-17-ET.pdf. Also Eneken Tikk-Ringas, Developments in the Field of Information and Telecommunication in the Context of International Security: Work of the UN First Committee. 1998–2012, Brief, Cyber Policy Process (Geneva: ICT4Peace, 2012), 5–6, available at: https://ict4peace.org/wp-content/uploads/2012/08/Eneken-GGE-2012-Brief.pdf

145 Tikk and Kertunnen (see note 13 above), 8–9. It was noted that GN states did not want to negotiate a new international treaty as international law provided for sufficient regulation: Dennis Broeders and Fabio Cristiano, Cyber Norms and the United Nations: Between Strategic Ambiguity and Rules of the Road, ISPI Dossier (Milan: Italian Institute for International Political Studies, 2020), 8, available at: https://www.ispionline.it/sites/default/files/pubblicazioni/dossier_cyber_april_2020_0.pdf. Also Christian Henderson, 'The United Nations and the regulation of cyber-security', in Nicholas Tsagourias (ed.), Research Handbook on International Law and Cyberspace (Cheltenham: Elgar, 2021), 585.

146 For three weeks a denial of service (DoS) and a distributed denial of services (DDoS) attack targeted Estonia, resulting in economic and communication disruption as government, media, banks and other websites were offline. Marco Roscini, Cyber Operations and the Use of Force in International Law (Oxford: Oxford University Press, 2014), 4–5. For an analysis of the cyber-attack see e.g. Delerue, note 2 above, 146–149; Eneken Tikk, Kadri Kaska and Liis Vihul, International Cyber Incidents: Legal Considerations. (Tallinn: Cooperative Cyber Defence Centre of Excellence (CCDCOE), 2010), 14–35.

that occurred during the Georgia–Russia conflict in 2008,[147] allegedly carried out by Russia,[148] when NATO (and like-minded states) were prompted to address the issue.[149] As a result, the Tallinn Manual process[150] was launched. This effort aimed to 'bring some degree of clarity to the complex legal issues surrounding cyber operations'.[151] Consequently it stated some rules that 'reflect consensus among the Experts as to the applicable *lex lata*, that is, the law currently governing cyber conflict'.[152] However, the final result considered only GN views, as the group of experts comprised only nationals of GN states and like-minded states.[153]

Furthermore, since its release in 2013, and despite the absence of authors from the GS in the Tallinn process, the Tallinn Manual has become a 'compulsory' benchmark regarding use of force in cyberspace (as will be shown in the following section). Almost every scholarly work and a majority of statements of international law's application in cyberspace (SILACs) refer to the Tallinn Manuals when examining this topic. As a consequence, the GN's narrative is reinforced, as the understanding on how a violation of the prohibition of the use of force occurs in cyberspace reflects the GN's position on the issue. This is problematic because there is little room for (even beginning to think about) other narratives. As many authors have noted, academics from the GN are seen as the 'authoritative voices' (the only voices) on the topic.[154] As a result, they have the power to influence the development of international norms.

Throughout the 2010s, GN states also started to release SILACs. These documents share their views on cyberspace regulation and aimed to achieve more

147 For an analysis of this conflict see Eneken Tikk, Cyber Attacks Against Georgia: Legal Lessons Identified (Tallinn: Cooperative Cyber Defence Centre of Excellence, 2008). Also Tikk et al. (note 15 above), 66–90.

148 See for example Delerue (note 2 above), 42, 68.

149 Michael N. Schmitt, Tallinn Manual on the International Law Applicable to Cyber Warfare (Cambridge: Cambridge University Press, 2013), 1–2.

150 The Tallinn Manual process refers to a research initiative from the CCDCOE that aims to address the challenges posed by cyber operations. To date, two Tallinn Manuals have been published (in 2013 and 2017) and a third is currently under development. See https://ccdcoe.org/research/tallinn-manual/

151 Schmitt (see note 18 above), 3–4.

152 Ibid., 5.

153 The experts were from a few Western countries (e.g. US, UK). This lack of diversity in the countries of origin of the experts has been criticised. See for example Dieter Fleck, 'Searching for international rules applicable to cyber warfare—a critical first assessment of the new Tallinn Manual', Journal of Conflict & Security Law 18 (2) (2013), 331–351: 335; Papawadee Tanodomdej, 'The Tallinn Manuals and the making of the international law on cyber operations', Msarryk University Journal of Law and Technology 13 (1) (2019), 67–86: 75.

154 B.S. Chimni, 'Third World approaches to international law: a manifesto', International Community Law Review 8 (2006), 3–27: 15; David Kennedy, 'My talk at the ASIL: what is new thinking in international law?', Proceedings of the ASIL Annual Meeting 94 (2000), 104–125: 121.

clarity on the question of how international law applies to/in cyberspace.[155] With this practice the GN could have the ability to influence the international dialogue concerning normative development in cyberspace. The documents signal the aspects the GN considers relevant to be addressed by the international community. For example, most SILACs released after the 2016 US presidential election interference started to discuss non-intervention and sovereignty, showing that the GN's problems are the relevant problems.

During this second phase, the GN's narrative has become the dominant narrative on the matter. As a result, there is little room for other perspectives.

The Global South and cyberspace: Latin American experience and narrative

Since the early 2000s Latin American states have shown interest in the regulation of activities taking place in cyberspace. They have been taking part in debates concerning cyberspace at both the regional and global levels. It can be said, then, that Latin American states are not new players concerning cyberspace regulation. Yet Latin America is still struggling to develop its own narrative and make its voice heard.

This struggle could be explained by the way that Latin American states have approached the topic. At the regional level, for almost two decades, they

155 Some notable exemptions are the statements of Iran, Israel and, more recently, Brazil and Kenya. See General Staff of the Armed Forces Islamic Republic of Iran, Declaration of General Staff of the Armed Forces of the Islamic Republic of Iran Regarding International Law Applicable to the Cyberspace, 2020, available at: https://www.aldiplomasy.com/en/?p=20901; Roy Schönford, 'Israel's perspective on key legal and practical issues concerning the application of international law to cyber operations', International Law Studies 97 (2021), 395–406. For the Brazil and Kenya statements see the UN official compendium of national contributions. UN General Assembly, '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', UN Doc. A/76/136 (13 July 2021), available at: https://digitallibrary.un.org/record/3933543/files/A_76_136-EN.pdf

neglected the question of applicability of international law to cyberspace,[156] as their focus was exclusively on cyber security and cyber-criminality issues.[157]

Latin American states only began to address international law's application in cyberspace in 2018. The IAJC decided to examine the topic under the title 'International law and State cyber operations: improving transparency'.[158] It aimed to survey (Latin) American states'[159] positions regarding how international law applies in cyber operations rather than to codify or progressively develop (international law on) the topic.[160]

This attempt resulted in Duncan B. Hollis' fifth report. The report highlighted that (Latin) American states had a mixed experience concerning the topic. As a result, Latin American states have not developed a common approach or their own narrative on responsible behaviour in cyberspace. In this sense, the report found that (Latin) American states 'have said relatively little to date about how international law applies to State [behaviour] in cyberspace'[161] and their domestic efforts have centred on cyber security and cyber-criminality.

At the same time, the report showed that (perhaps) Latin American states are trying to find their own voice on the matter. Despite expressing doubts on 'how universally the extant law might apply',[162] they have a certain level of agreement on the 'overall application of international law to cyber operations'.[163] I argue that this shows an emerging Latin American agreement on international law as the necessary framework for cyberspace regulation. For instance, Latin American

156 IACJ's 5th report on international law's application to cyberspace reinforces this idea. CJI/doc. 615/20 rev. 1, 16 para. 4.

157 At the regional level they have engaged with the development of the new hemispheric security agenda, which was updated in 2003 to encompass non-traditional threats such as cyber-attacks and cyber-criminality. The close relationship between cyberspace regulation and the hemispheric security agenda had unintended consequences for Latin American states' approach to cyberspace. See e.g. Concepción Anguita Olmedo and Mariano Bartolomé, 'El reto de la gobernanza global en la ciberseguridad. La gestión de la Unión Europea (UE) y la Organización de Estados Americanos (OEA)', in Comunicación Política en el Mundo Digital: Tendencias Actuales en Propaganda, Ideología y Sociedad (Madrid: Dykinson S.L., 2021), 623–648; Louise Marie Hurel, 'Beyond the Great Powers: challenges for understanding cyber operations in Latin America', Global Security Review 2 (1) (2022), 21–31.

158 CJI/doc. 576/18, 154–155.

159 It is safe to assume that the proposal aimed to understand the GS position regarding international law's application in cyberspace, as 33 of the 35 members of the OAS are GS states and 24 of these are Latin American states. Also, the GN position was already known as the USA had a public stance on the topic while Canada was actively engaging with the debates at the UN level. For example, for the US point of view see Brian J. Egan, Remarks on International Law and Stability in Cyberspace, 10 November 2016, available at: https://2009-2017.state.gov/s/l/releases/remarks/264303.htm; Harold H Koh, International Law and Cyberspace: Remarks, 18 September 2012, available at: https://2009-2017.state.gov/s/l/releases/remarks/ 197924.htm

160 CJI/doc. 615/20 rev. 1, 1 para. 1.

161 CJI/doc. 615/20 rev. 1, 16 para. 4.

162 Ibid., 17 para. 8.

163 Ibid., 17 para. 9.

states seem to accept the application of international law on subjects such as use of force, international responsibility of states, sovereignty and international humanitarian law (IHL) while, at the same time, not agreeing on the particulars of each regime.

Overall the report's findings show that there is increasing interest among Latin American states in addressing this matter. This conclusion is reinforced by their engagement with UN process at the global level. During the former OEWG several Latin American states had increasingly become involved with the debates in this forum. They have not only made public statements on international law's application but also released public documents on the topic.[164]

Despite this emerging interest, Latin America is still struggling to develop its own narrative on responsible state behaviour in cyberspace and to make its voice heard. Particularly troublesome for the development of this Latin American perspective is the lack of transparency with which Latin American states approach the matter. While showing interest on the topic, they are not willing to publicly debate their stance. A variety of external and internal factors, such as capability deficiencies and geopolitical confrontation, explain this behaviour, but at the same time prevent the development of Latin America's own position on the subject. As one state representative put it, '[we] are not quite there yet'.[165]

To help with the process, the IACJ has retained the subject on its agenda. As it has recently highlighted, a 'clear attitude regarding the scope of the application of international law in the context of cyberspace'[166] is necessary to be able to address the wide range of implications that cyber operations can have for states. The importance of cyberspace for everyday life requires that Latin American states overcome their shortcomings and address the topic. The regional level continues to offer a good forum for this aim.

164 For example, Brazil is the only Latin American state that has released a comprehensive SILAC. Colombia, Costa Rica and Cuba have released (at the UN level) some documents highlighting their position regarding international law's application in cyberspace. However, they are very limited as they do not articulate a comprehensive stance on the issue. They can be found on the OEWG 2021-2025 website: https://meetings.unoda.org/section/oewg-ict-2021_general-statements_14537_general-statements _16368/

Several attempts have been made to get access to official documentation developed by Latin American states regarding the OAS process, such as contacting the Department of International Law of the OAS and the foreign affairs ministries of the states that have submitted an answer to the IACJ questionnaire. However, it was almost impossible to get the documents as in most cases the answer was that this documentation was confidential. Only Peru and Guatemala have supplied the required documents.

165 CJI/doc. 615/20 rev. 1, 7 para. 21.

166 Inter-American Juridical Committee, 'Annual Report of the Inter-American Juridical Committee to the General Assembly: 2021', CJI/doc.657/21, 50 (OAS, 11 August 2021), available at: http://www.oas.org/en/sla/iajc/docs/INFOANUAL.CJI.2021.ENG.pdf

Is there room for a Latin American narrative?

Even though Latin American states have engaged with discussions about cyberspace regulation in recent decades, it is difficult to identify a Latin American narrative concerning international law's application in cyberspace. As shown in the previous section, some tentative attempts suggest that Latin American states want to develop their own vision on how international cyberspace norms should be.

During the OAS process, Latin American states have voiced their concerns on 'whether differences in legal capacity might impact the law's actual application or evolution'.[167] They fear that GN states may have the 'capacity to disproportionately influence the content and boundaries of rules for cyberspace over states lacking such a capacity'.[168] At the same time they have called for 'developing a distinctly Latin American perspective on the international governance and legal framework of cyberspace'.[169] But that is easier said than done.

How could Latin American states develop their own narrative? There is no simple answer to this question. Their approach has been somewhat contradictory. Their documents and statements on the topic suggest that they are not necessarily satisfied with the narrative provided by the GN. One of their key concerns is how the development gap in cyber capabilities impacts (and will continue to influence) the development of cyber norms. Hence they consider that these

167 CJI/doc. 615/20 rev. 1, 17 para. 10.

168 Ibid., 17–18 para. 10.

169 Ibid., 34 para. 54.

aspects should be taken into account when discussing these issues.[170] As a result, it seems that they are not willing to accept the imposition of GN narratives on the matter.

Yet when dealing with the subject their (often perceived) lack of capability[171] evokes an excessive reliance on the same narrative that they call into question. For instance, some Latin American states (e.g. Ecuador, Brazil and Guatemala)[172] refer to the GN's academic production, such as Tallinn Manuals, to assess their positions regarding international law's application in/to cyberspace.[173] This may not be seen as a concern per se; it is problematic, though, when the GN's academic production is presented as one of the only (authoritative) sources in the matter because the GN's narrative tends to be seen as the only valid experience.

An excessive reliance on the GN's narrative renders the GS's experience invisible. As Gathii has noted, there is a 'limited geography of places and ideas that

170 For example, Venezuela stated that 'The future development of international standards and regulations applicable to Information Technology and Telecommunications in the field under consideration by the UN Working Group should be the result of the agreement of all States, in accordance with the legitimate interests and concerns of all parties involved (as universally as possible), and not only of those countries with a higher level of economic, technological and industrial development, however legitimate they may be.' Ecuador emphasised 'the need for a wider recognition of asymmetries in the capacity to implement norms, rules and principles of responsible behaviour of States; as well as the differentiated effects that an ICT incident, for example, would have on a specific critical infrastructure in a developing country.' Venezuela, 'Preliminary Considerations of Venezuela to the Initial Pre-Draft Report of the OEWG on Developments in the Field of Information and Telecommunications in the Context of International Security', 1 para. 3 (OEWG, 2020), available at: https://front.un-arm.org/wp-content/uploads/2020/04/nv-00069-annex.pdf; Ecuador, 'Ecuador preliminary comments to the Chair's 'Initial pre-draft' of the Report of the United Nations Open Ended Working Group on developments in the field of information and telecommunications in the context of international security (OEWG)', 2 (OEWG, April 2020), available at: https://front.un-arm.org/wp-content/uploads/2020/04/ecuador-comments-on-initial-pre-draft-oewg.pdf.

Similar concerns have been voiced by, among others, Colombia, Chile and Nicaragua. See Chile, 'Comentarios de Chile al pre-informe del Chair, 2 para. c (OEWG, 2020), available at: https://front.un-arm.org/wp-content/uploads/2020/04/comentarios-de-chile-al-pre-informe-del-chair-oewg-2020-v2.pdf; Chile, 'Statement 2Rev. Draft', 2 (OEWG, 28 July 2022), available at: https://documents.unoda.org/wp-content/uploads/2022/08/4.-Remarks-Chile-tercera-reunion-OEWG-2021-2025_28JUL2022.pdf; Colombia, 'Colombia's comments on the initial "Pre-draft" of the report of the OEWG on developments in the field of information and telecommunications in the context of international security', 1–2 (OEWG, 16 April 2020), available at: https://front.un-arm.org/wp-content/uploads/2020/04/colombia-general-comments-pre-draft-oewg-16-04-2020.pdf; Nicaragua, 'Nicaragua's considerations to the initial document of the Open-Ended Working Group on progress in the field of information and telecommunications in the context of international security', 4 para. 7 (OEWG, 3 April 2020), available at: https://front.un-arm.org/wp-content/uploads/ 2020/04/minic-mis-143-04-2020-permanent-mission-of-switzerland.pdf.

171 CJI/doc. 615/20 rev. 1, 7 paras 17–21.

172 For example, when examining the treatment of use of force in cyberspace Ecuador relies on the Tallinn Manual 2.0. CJI/doc. 615/20 rev. 1, 18 note 50. Brazil and Guatemala highlight that they have relied on Tallinn Manuals to articulate their positions. Brazil SILAC, UN Doc. A/76/136, 18; Guatemala official response to OAS questionnaire, answer 10.

173 Due to the scarcity of documents, in this chapter I do not discuss to what extent Latin American states support or do not support the scope of international law rules as they are lay out by Tallinn Manuals. Latin American statements regarding international law's application in/to cyberspace tend to be very general. As a result it is difficult to assess whether they agree with each of the provisions in the manuals, and thus with the narrative proposed by the GN.

dominate'[174] international law. The international law produced in this limited geography 'influences and reinforces our understandings'[175] of the international practice. Hence the GN's narrative is the one that matters and becomes the benchmark for how international cyberspace norms should be.[176]

Moreover, this contradictory approach makes it harder to identify the emergence of a Latin American narrative regarding international cyberspace norms. Are Latin American states supporting the Tallinn Manual's approach to cyberspace regulation because they are convinced that this is the way to go, or because it is the benchmark that they are supposed to conform to? When the GN's narrative receives much (almost all) the attention, there is little room to develop other narratives.

No single reason explains this contradiction; rather there are multiple intertwined factors. I will address the three main ones: (a) different stages of cyberspace capabilities development; (b) geopolitical debates; and (c) the approach of the IAJC to the topic.

With respect to cyberspace capabilities, Latin American states present different stages of development. Hollis' fourth report highlighted that there is an uneven distribution of cyber capabilities (both technical and legal). While some state responses showed a deep knowledge and understanding of the complexities involved when discussing cyber operations and related issues, others' responses were much more limited in their understanding.[177]

As suggested by IACJ, one way to tackle this problem is by promoting cyber capability-building efforts. However, these initiatives come with caveats, as they do not correspond to 'the reproduction of donor–recipient/north–south logic'.[178] For instance, almost every cyber capability-building initiative undertaken by OAS

174 James Thuo Gathii, 'The promise of international law: a Third World view,' American University International Law Review 36 (2021), 377.

175 Gathii, 378.

176 An example in this regard is the Colombian statement on the international law's application in/to cyberspace submitted to the First Substantive Session of the OEWG 2021–2025. In this statement, Colombia stressed the need for developed states to share their practice regarding international law's application to/in cyberspace in order to foster a better understanding of the norms. Colombia, 'Intervención de Colombia. Primer período de sesiones sustantivas', 2 (OEWG, 15 December 2021), available at: https://documents.unoda.org/wp-content/uploads/2021/12/INTERVENCION-DE-COLOMBIA-DERECHO-INTERNACIONAL-PUNTO-5c-OEWG-CIBERSEGURIDAD-PRIMERA-SESION-DIC.-15-2021.pdf

177 Inter-American Juridical Committee, 'Improving Transparency. International Law and State Cyberoperations: Fourth Report', CJI/doc. 603/20 rev. 1 corr.1, 5 para. 11 (OAS, 5 March 2020), available at: https://www.oas.org/en/sla/iajc/docs/CJI_doc_603-20_rev1_corr1_ eng.pdf

178 Hurel (note 26 above), 21.

throughout 2019 was related to GN partners.[179] This means that in order to improve cyber capabilities, Latin American states need to turn to those that have them: the GN. As a result, the GN narrative could be reinforced by these efforts.

Regarding geopolitical confrontation, the 'great powers rivalry' (US, UK, EU on the one hand, China and Russia on the other) has permeated cyberspace regulation debates[180] since the early beginnings. Fuelled by this confrontation, debates on cyberspace regulation have unfolded around the democracy/freedoms vs autocracy/control dichotomy.[181] When dealing with responsible state behaviour in cyberspace, Latin American states could not escape that logic: they are caught up in 'great powers' confrontation, and assuming one or the other position could be seen as taking a stance in this rivalry. Therefore, Latin American states show 'reluctance to make similar signals lest they embroil that State in the competition and conflict among these actors'.[182] As a result, to date it has been difficult to develop a distinctly Latin American position on the matter, as many states have preferred to remain silent.[183]

Lastly, the IACJ's approach to the topic might have played a significant role in how Latin American states have addressed/examined cyberspace regulation. As the advisory body on juridical matters of the OAS, the IACJ influences (Latin) American states regarding the progressive development and codification of international law.[184] Hence, when suggesting how a topic should be addressed, the IACJ is setting the tone for future debates on the matter. The IACJ's influence is not problematic per se, but its approach to international law's application in cyberspace is troublesome because rapporteurs relied on the GN's narrative to address the subject. For instance, almost every report focuses on the GN's

179 This was the last report publicly available. Inter-American Committee Against Terrorism (CICTE), '2019 Annual Report of the Inter-American Committee Against Terrorism (CICTE) to the Fiftieth Regular Period of Sessions of the General Assembly', CICTE/doc.5/20 rev. 1, 3–6 (Washington, DC: OAS, 25 September 2020), available at: https://www.oas.org/es/sms/ cicte/sesiones/ordinarias/2020/

180 See the analysis of Hurel (note 26 above).

181 See e.g. Henderson (note 14 above), 593–595.

182 CJI/doc. 615/20 rev. 1, 7 para. 20.

183 For example, only seven Latin American states answered the questionnaire prepared by the IACJ in 2019 to survey international's law application in cyberspace in the region: the others remain silent. Moreover, Salazar's report highlighted that (Latin) American states preferred to engage in cyber capability-building efforts rather than answer another questionnaire or release a public statement on the topic. CJI/doc. 671/22 rev. 2, 12.

184 See e.g. Dante Mauricio Negro Alvarado, 'La Labor Del Comité Jurídico Interamericano', Agenda Internacional XXII (33) (2015), 211–230; Dante Negro Alvarado, 'El Comité Jurídico Interamericano Como Órgano Consultivo de La Organización de Estados Americanos', Agenda Internacional 11 (21) (2004), 269–282; José Luis Siqueiros, 'La OEA y El Derecho Internacional', Revista Mexicana de Política Exterior 54 (1998), 37–67. Compare Lunardelli Caldeira, who maintains that IACJ has lost its relevance and has been replaced by other bodies within OAS: Alberto Lunardelli Caldeira, 'El Comité Jurídico Interamericano de La OEA y Sus Perspectivas: La Necesidad de una Reforma Profunda y de una Corrección de Curso', Revista Iberoamericana de Derecho Internacional y de La Integración 11 (November 2019), 170–208.

experience as the one that matters as it calls attention to cyber-attacks targeting GN states and their responses.[185] They also highlight GN academic production, i.e. the Tallinn Manuals or the Oxford Process are regarded as the (sometimes sole) reference works regarding the study of international's law application in cyberspace.[186] Hollis' reports also underline that IACJ work on the topic aligns with the EU's call to submit national views, as if (Latin) American s should follow the EU's lead.[187] Moreover, in 2022, to help (Latin) American states develop and release their own SILACs, the IACJ convened a forum to debate international law's application to cyberspace. However, the panel almost entirely comprised GN experts.[188]

Thus, the IACJ's heavy reliance on the GN's perspective on cyberspace regulation could have harmful effects for a Latin American narrative. The GN's narrative could be perceived by Latin American states as the only legitimate or valid one for norms development. Therefore Latin American states could be tempted to follow the imposed narrative and not be able to develop a distinctive view on the subject.

Is there room for a Latin American narrative? Overall it seems that while trying to play catch-up, Latin American states are being absorbed by the GN's narrative. Therefore they cannot even begin to think of their own narrative. However, there is still hope and room: they only need to step up and not to conform blindly to the GN's narrative.[189]

Conclusion

Since 2007, the GN's narrative has become the dominant narrative concerning international cyberspace norms development. During the past 15 years, cyber-attacks targeting GN states usually have triggered the normative development in

185 See e.g. Inter-American Juridical Committee, 'International Law and State Cyber Operations: Improving Transparency', CJI/doc. 570/18, 1–2 paras 2–3 (OAS, 9 August 2018), available at: https://www.oas.org/en/sla/iajc/docs/CJI_doc_570-18.pdf; CJI/doc. 657/21, 50; CJI/doc. 671/22 rev. 2, 4–5.

186 See e.g. CJI/doc.657/21, 52-53; CJI/doc. 671/22 rev. 2, 6, 10.

187 See e.g. CJI/doc. 615/20 rev. 1, 5 para. 10.

188 CJI/doc. 671/22 rev. 2, 12.

189 David P. Fidler, 'Revolt Against or From Within the West?: TWAIL, the Developing World, and the Future Direction of International Law', Articles by Maurer Faculty (Bloomington: Maurer School of Law, Indiana University, 2003), available at: https://www.repository.law.indiana.edu/facpub/2126/

this domain. These experiences have influenced the dialogue, as GN states have started to flag the norms likely to be affected by those cyber-attacks as relevant for the international community as a whole.

Despite their early engagement with the process of development of international regulation for cyberspace, Latin American states have struggled to articulate a clear position in the debates. This is mainly because they have yet to develop their own narrative, one that detaches from the hegemonic narrative of the GN.

Several internal and external factors prevent Latin American states from achieving the goal of developing a distinctly Latin American narrative. These states need to enhance their cyber capabilities to be able to meaningfully engage with the ongoing debate regarding international cyberspace norms development. The OAS continues to offer a good forum for this aim. However, OAS efforts should take account of Latin American uniqueness and not rely blindly on GN narratives.

CHAPTER 4

The legal framework for cybercrime accountability in the Western Balkans countries as a turning point for EU integration

ANDREJA MIHAILOVIĆ

Overview

This chapter analyses the Western Balkans' (WB) major challenges in building a cybersecurity legislative framework and combating cybercrime in the light of EU integration. WB is a term referring to countries in the Balkans that are not currently members of the EU (excluding Turkey). A number of WB countries (such as Montenegro and North Macedonia) are actively pursuing EU membership, as well as Bosnia and Herzegovina (BiH) and Kosovo, which are currently part of the EU's enlargement process. The tumultuous past of the WB region, as well as the need to re-establish connections through programmes that foster regional cooperation and post-conflict rehabilitation, has shaped the political landscape of the region. Most of the countries in this region

have adopted the unified legal framework established by the former Yugoslavia, which is notable for its high homogeneity and clearly articulated legal identity. There is no reason to believe that the WB region is trailing behind when it comes to ICTs' growth. As a result of the large numbers of internet users in these countries, their citizens are at risk both of becoming victims of cybercrime and of being educated to commit crimes utilising or against ICTs.

All jurisdictions in the WB region (excluding BiH) have national cybersecurity strategies, although the majority of cybercrime operations take place through communication tactics, with most of them involving the exploitation of private information (in many cases belonging to EU citizens). Efforts are being made to address these issues through market orientation and investment policies, as well as uniform data monitoring. Even so, 'unregistered labor', tax evasion and other organised crime and corrupt practices are the main issues in the WB area. To address those challenges, the region needs updated current legislation and improved broadband infrastructure and access, digital literacy and information security awareness, especially regarding the importance of responsibility in cyberspace and comprehensive information exchange in order to effectively process cybercrimes.

It is undeniable that the WB's inclusion in the pan-European digital market will result in enhanced regional industry structure, innovation, reduced administrative barriers (as a prerequisite for strengthening public sector efficiency), enhanced knowledge transfer and protection of privacy. The chapter highlights the fact that for the WB region, strengthening the legal environment for cybercrime responsibility is a development opportunity, an engine of innovation and a cornerstone of a comprehensive 'reform agenda' to enhance productivity in this region, which has tremendous economic potential and strategic importance.

Introduction

In recent decades, users all around the world have seen adverse effects of the rapid expansion of digitalisation, as well as societal and cultural advantages. The fundamental roadways that rely on information and communications technology (ICT) and digital innovations may be a tremendous facilitator of inclusive growth, but only in a protected, sustainable and resilient cyberspace. The proliferation of cybercrime is correlated with the development of information technology, broadband access and the possible applications of sensitive information. This resulted in cybercrime becoming one of the most widespread and expensive

forms of crime worldwide. Whereas traditional forms of crime might happen on a social and economic basis, cybercrime is a business that seeks rapid gain, which includes financial benefits as well as systematically chosen targets.190 It is anticipated that global accumulated data will surpass 175 zettabytes by 2025, covering everything from streaming content to healthcare system records. At the same time, the latest reports forecast that the cost of cybercrime will reach $10.5 trillion by 2025.191

As a result of emerging digital threats, the EU constructed a regulatory architecture in order to: maximise resilience and establish best practices in cyberspace; enable its capacity to detect, mitigate and react appropriately to cyber intrusions; and expand its alliances in favour of a globalised and inclusive virtual world.192 The General Data Protection Regulation (GDPR) completely replaced the EU's 1995 data protection regulation, Directive 95/46/EC of the European Parliament, in 2016. Since the EU considers privacy to be a basic human right, the GDPR provides data protection norms with which Member States must comply, although there is some flexibility for states to endorse 'specific exemptions' from certain aspects of the GDPR.193 Shaping Europe's Digital Future, the Commission's Recovery Plan for Europe194 and the Security Union Strategy 2020–2025195 all make extensive use of the novel functional cybersecurity strategy for the digital decade that was utilised by the EU with the goal of enhancing Europe's technology and digital sovereign control. They outline legislative measures that would better integrate cybersecurity into the EU's legal provisions on privacy, technologies, markets and digital services,196 and include specific ideas for the implementation of three fundamental factors of a secure and resilient cyberspace: (1) resilience, technical sovereignty and leadership; (2)

190 A. Alexandrou, *Cybercrime and Information Technology: Theory and Practice – The Computer Network Infrastructure and Computer Security, Cybersecurity Laws, Internet of Things (IoT), and Mobile Devices* (Boca Raton: Taylor & Francis, 2022).

191 Cisco Cybersecurity Ventures, *2022 Cybersecurity Almanac*, available at: https://cybersecurityventures.com/cybersecurity-almanac-2022/; Accenture, *How Aligning Security and the Business Creates Cyber Resilience*, available at: https://www.accenture.com/_acnmedia/PDF-165/Accenture-State-Of-Cybersecurity-2021.pdf

192 EEAS, *Cybersecurity: EU External Action, 2022*, available at: https://www.eeas.europa.eu/eeas/cybersecurity-eu-external-action_en

193 J. Kosseff, *Cybersecurity Law* (Chichester: Wiley, 2020); EC, 'Data protection', available at: https://ec.europa.eu/info/law/law-topic/data-protection_en

194 'Europe's Moment: Repair and Prepare for the Next Generation', COM 98 final, 2020.

195 The EU Security Union Strategy 2020–2025.

196 A. Bendiek and M.C. Kettemann, 'Revisiting the EU Cybersecurity Strategy: a call for EU cyber diplomacy', *Stiftung Wissenschaft und Politik German Institute for International and Security Affairs*, no. 16 (2021).

the operational ability to prevent, deter and respond; and (3) collaboration to promote a global and open cyberspace.197

The importance of a national cybercrime, cybersecurity and cyber-defence framework

The creation of an environment that encourages continuous economic growth and digital transformation cannot happen unless data protection issues, the growth of ICT-enabled architecture and online services are wisely aligned with priority areas and schemes for building capacity. Therefore, governments should integrate their economic development agendas with overall cybersecurity objectives in order to actualise the capabilities of advanced technologies. Establishment of the national strategy, goals and deliverables encourages policymakers to engage in cyberspace holistically throughout the entire value chain, rather than focusing on a single segment, target or hazard identification, which empowers countries to provide a systematic and comprehensive approach.198 One thing is certain: the challenge of a uniform approach in the cybersecurity realm can't be addressed overnight.

Contemporary analysis of national cybersecurity strategies reveals that open, guarded and resilient cyberspace must be explained in a constructive triptych of cybersecurity concepts: three distinct but closely connected frameworks striving to achieve the common objective of a protected digital world. Even though the term 'active defence' is often used in armies to denote offensive activities, it remains unclear once transferred to the cyber realm, suffering from confusion in relevant legislation and government systems. Nonetheless, efforts to define the notion have been made. Dewar offers a definition of active cyber defence (ACD) based on proactive system security flaw tracking, evaluation and prevention as 'a method of achieving cyber security based on the deployment of measures

197 EC, 'New EU Cybersecurity Strategy and new rules to make physical and digital critical entities more resilient', 2020, available at: https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2391

198 International Telecommunication Union (ITU), 'Strategic Engagement in Cybersecurity: Guide to Developing a National Cybersecurity Strategy' (Geneva, 2021).

to detect, analyze, identify, and mitigate threats to and from cyberspace in real-time, combined with the capability and resources to take proactive or aggressive action against threat agents in those agents' home networks'.199 ACD is essential for the understanding of the methodological approach to cybersecurity since it represents proactive external actions that distinguish it from other methods referred to as 'passive cyber defence' (PCD). PCD is defined by Farwell and Rohozinski as 'firewalls, cyber "hygiene" that trains an educated workforce to guard against errors or transgressions that can lead to cyber intrusion, detection technology, "honey pots" or decoys that serve as diversions, and managing cyberspace risk through collective defense, smart partnerships, information training, greater situation awareness, and establishing a secure, resilient network environment'.200

On the other hand, there is currently no consistent and widely accepted definition of cybercrime. Given the fact that the Convention on Cybercrime offers effective guidelines for standardising the efforts against cybercrime, legislators utilise a range of terminology.201 Subsequently, the most accepted definition of cybercrime is as follows: '(1) a crime threatening ICT – information and network safety (computer integrity crime or cybercrime in the narrow sense), (2) a crime using ICT to commit conventional crime (computer-related crime), and (3) a content-related crime, such as child pornography, hate speech, and infringement of intellectual property rights'.202 The Convention identifies the groups of cybercrime as: '(1) offences against the confidentiality, integrity, and availability of computer data and systems (illegal access, illegal interception, data interference, system interference, misuse of devices), (2) computer-related offences (computer-related forgery and computer-related fraud), (3) content-related offences (offences related to child pornography); and (4) offences related to infringements of copyright and related rights.' Despite the diversity in interpretations, existing findings indicate that cybercrime may be understood as a technological matter,

199 R. Dewar, *The 'Triptych of Cyber Security': A Classification of Active Cyber Defence*, International Conference on Cyber Conflict, CYCON, 2014.

200 J. Farwell and R. Rohozinski, 'The new reality of cyber war', *Survival* 54 (4) (2012).

201 S. Boes and E.R. Leukfeldt, 'Fighting cybercrime: a joint effort', in R.M. Clark and S. Hakim (eds), *Cyber-Physical Security: Protecting Critical Infrastructure at the State and Local Level* (New York: Springer, 2017); B. Diamond and M. Bachmann, 'Out of the beta phase: Obstacles, challenges, and promising paths in the study of cyber criminology', *International Journal of Cyber Criminology* 9 (1) (2015); M. Goodman, 'International dimensions of cybercrime', in S. Ghosh and E. Turrini, *Cybercrimes: A Multidisciplinary Analysis* (New York: Springer, 2010).

202 A. Završnik, 'Cybercrime – definitional challenges and criminological particularities', *Masaryk University Journal of Law and Technology* 2 (2), 2008.

typically white-collar criminal activity, a geopolitical concern, or the result of a socioeconomic perspective.

State of play in the Western Balkans

The European and international regulatory environment on cybercrime and data security offers WB countries critical and comprehensive mentorship for continuously improving institutional and regulatory frameworks. The WB countries are nominally aligned with the main global pathways, primarily the Budapest Convention.203 Although substantial progress has been made in all countries to adhere to the legislative structure for combating cybercrime, the efficient enforcement of these processes remains a major obstacle associated with lack of coordination, competence and funding for the streamlined investigation of crime and organisational sources. To illustrate, sufficient harmonisation and implementation of the Directive on security of network and information systems (NIS Directive)204 provisions into national practices is a fairly complex model, mirroring a new vision throughout sector management and stronger ties with the industry; correspondingly, economic support and information security professional training, policy reforms and functional enforcement of policies in this area are imperative.205

WB countries were already driven by digital transformational change while highly aware of the potential risks to their essential systems, infrastructural facilities, industries and corporate entities, along with regional stability, that might result from the widespread abuse of communications technology and insufficient resilience. In the WB area, a lot of progress has been made in setting up the conceptual framework for cybersecurity, including optimising the existing institutional architecture to prevent or reduce cyberspace vulnerabilities.

There are numerous worrisome occurrences in the WB region as an important hub with a gateway location that has many advantages in terms of the distribution of goods, services, consumers, money and innovations. The most critical

203 Council of Europe Portal, The Budapest Convention on Cybercrime, available at: https://www.coe.int/en/web/cybercrime/the-budapest-convention

204 NIS Directive 2016/1148 of the European Parliament and of the Council of 6 July 2016, available at: https://eur-lex.europa.eu/eli/dir/2016/1148/oj

205 DCAF – Geneva Centre for Security Sector Governance, *National Cybersecurity Strategies in Western Balkan Economies* (Geneva, 2021).

issues identified are organised crime, computer hacking, terrorism, violent extremism, radicalism and espionage. The dynamics of viral dangerous phenomena, as well as current global and regional security-political trends, require the much more advanced structure of the cybersecurity sector.206

Although the bulk of cybercrime operations take place via communication methods, including the exploitation of private information, all jurisdictions in the WB area (except BiH) have national cybersecurity strategies (including many that apply to EU citizens). The fundamental objectives of all the strategies implemented among all WB economies are to optimise institutional data protection; ensure normal function and endurance of security mechanisms of essential services; strengthen the safeguarding of critical infrastructures; improve capacities for combating high-tech crime; and increase awareness and level of data protection. All regional national cybersecurity strategies (NCSs) envision expanding the legal foundation by setting in place multi-sectoral national cybersecurity coordination bodies or councils, entrusting relevant institutions to enforce computer security incident response teams (CIRTs) or defining professionals whose primary responsibilities will be related to cybersecurity activities. Nevertheless, the continuously evolving dynamics of the online world, significantly larger reliance on ICTs, and the spread of digital hazards do require adjustments to coherent national policies. Policymakers should define the emergence of facilities within the coordinated development of cybersecurity regulations in order to exploit the advantages and overcome the threats of the digital revolution.

Montenegro

The Montenegrin cybercrime landscape encompasses several legal documents, including the Criminal Code, Criminal Procedure Code, law on information security, law on the national security agency, law on protection of critical infrastructure and law on personal data protection, as well as laws on electronic communications, electronic commerce and electronic signature. Since the adoption of the law on information security and the regulation on information security measures in 2010, Montenegro has shown strong dedication in the development of a cybersecurity architecture and introduced a national cybersecurity strategy in 2013. Following the achievements of the previous cybersecurity strategies, it

206 M. Trbojević and B. Svirčević, 'Strategic directions of activities of the intelligence and security agencies of the Western Balkans', *Kultura Polisa* 19 (1) (2022).

adopted its third national cybersecurity strategy[207] for the period 2022–2026, establishing five major goals: (1) cyber-defence and crisis management capabilities; (2) critical infrastructure security; (3) cybercrime and personal data protection; (4) education, research and development; and (5) public–private partnerships (PPPs) and international collaboration.

Remarkable progress has been made with the formation of a cybersecurity organisational unit within the Ministry of Defence and the Montenegrin army, which expanded organisational and administrative mechanisms for cyber-defense development and cyber operations. In light of this, the Security Operations Center–SOC MO was launched, which used cutting-edge technical solutions and developed a framework for preventing cyber intrusions and dealing with cyber-attacks. Regarding international cooperation, it is noteworthy that Montenegro joined the European Centre of Excellence for Combating Hybrid Threats in 2019, which facilitated the exchange of experiences and best practices with NATO and EU member states and aided Montenegro's efforts to empower national capacities to combat hybrid threats. Furthermore, Montenegro joined the NATO Center of Excellence for Cooperative Cyber Defense in Tallinn in 2020 to facilitate interoperability in this area through multidisciplinary research activities and professional assistance.

Serbia

Serbia has extensive potential to foster cybercrime competence, via both professional practice and institutional pathways, although it faces challenges with retaining a highly skilled workforce and meeting diverse market needs for cybersecurity experts.[208] Considering the responsible approach to overcoming the issues of cybercrime and a stable information security institutional and legal base, Serbia has empowered the country to grow security protocols, such as the National CERT (computer emergency response team), that could safeguard the resilience of key infrastructure throughout the state. The Serbian army has a growing cyber-defence system and collaborates with civilian organisations and partner countries. Under the Telecommunications and Information Technology

207 Cybersecurity Strategy of Montenegro 2022–2026, available at: https://www.gov.me/dokumenta/8a2de214-c58e-4524-9196-c08886f5829b

208 World Bank, 'Serbia Has Undertaken Critical Steps in Cybersecurity', available at: https://www.worldbank.org/en/news/press-release/2020/12/21/serbia-has-undertaken-critical-steps-in-cybersecurity-says-first-cybersecurity-capacity-maturity-model-assessment

Directorate (J-6), as an inter-service organisational unit of the Serbian Armed Forces General Staff, there is a Department for Information Security and Cyber Defence intended for coordination and control of information security.209

The importance of information security is acknowledged by the Information Society and Information Security Development Strategy for the period 2021–2026, which is aligned with the NIS Directive and encompasses all priority areas related to the growth of the information society, such as electronic communications; e-government, e-health and e-justice; ICT in education, science, and culture; e-commerce; the ICT business sector; and data protection.210 In terms of cybercrime regulation, the Serbian Criminal Act prescribes the unauthorised collection of personal data as a felony. The law on information security defines actions for protecting information systems from privacy issues and the responsibility of legal entities, and determines relevant agencies responsible for implementation of prevention measures, alignment and tracking of the judicious handling of the stipulated provisions and operating system progress.211

Bosnia and Herzegovina

BiH seems to be in the initial phases of institutional transformation. The strategic approach for public service reform has indeed been endorsed by all areas of government, but without a comprehensive mechanism for addressing cybersecurity incidents. BiH has been mandated to ensure the commitments, fundamentals and requirements resulting from participation in international institutions—the UN and the Organization for Security and Cooperation in Europe (OSCE)—as well as the regional initiatives. Even though several initiatives address cybersecurity in core components, BiH remains the only country in southern Europe without a national cyber strategy and CIRT.212 Suboptimal alignment, an inadequately

209 Serbian Armed Forces, Telecommunications and Information Technology Directorate (J-6), available at: https://www.vs.rs/en/units/serbian-armed-forces/general-staff/telecommunications-and-information-technology-directoratej-6-

210 Information Society and Information Security Development Strategy of the Republic of Serbia 2021–2026, available at: https://mtt.gov.rs/extfile/sr/35315/Information%20Society%20and%20InfoSec%20Strategy%202021-2026111.pdf

211 CMS, 'Data Protection and Cybersecurity Laws in Serbia', available at: https://cms.law/en/int/expert-guides/cms-expert-guide-to-data-protection-and-cyber-security-laws/serbia

212 D. Maravić, 'Cybersecurity Policy Development and Capacity Building – Increasing Regional Cooperation in the Western Balkans', *DCAF High Level Regional Conference on Cyber Resilience and Cybersecurity Capacity Building in the Western Balkans*, 2021, available at: https://www.dcaf.ch/sites/default/files/imce/Events/CybersecurityConference_DiscussionPaperPanel2_PublicCapacityBuildingRegionalCooperation.pdf

integrated policy and limited capabilities continue to be difficulties. In consideration of its pledge to EU accession, BiH needs to establish a robust cybersecurity policy and harmonise existing cyber-security regulations. The dictating instruments are the EU GDPR and the NIS Directive. BiH is indeed obligated by OSCE cyber confidence-building activities.213

Therefore, the informal initiative by the multidisciplinary working group of experts from different administrative levels in BiH to set a recommendation for a strategic cybersecurity framework, under the supervision of the OSCE mission, was a promising effort. The major accomplishment of that initiative is setting preconditions for the highly inclusive policy creation process and cross-sectoral cooperation in this field, which provides a basis for continuous improvement.

Kosovo

According to a report which utilises the Cybersecurity Capacity Maturity Model for Nations (CMM) methodology developed by the University of Oxford's Global Cyber Security Capacity Centre (GCSCC),214 Kosovo has managed to improve many aspects of cybersecurity potential in recent years and gained a broad insight into gaps and potentials for inclusive growth. Kosovo was the first country in the world to pilot the CMM examination in 2015, demonstrating that it has achieved the basic components in creating cybersecurity infrastructure, most noticeably by adopting its first NCS. As shown in the study, the NCS has provided encouragement to an extensive reform effort, including the rewriting of cybercrime laws and the provision of a legal foundation for identifying key essential infrastructure. This monitoring is carried out with funding provided by Korea's Ministry of Economy and Finance through the Korea–World Bank Group Partnership Facility (KWPF), managed by the World Bank, which continuously assists citizens, as well as governmental and educational organisations, in improving access to online information sources, networks, and economic activities, via the Digital Economy Project for Kosovo.215 In addition, the UNDP Kosovo

213 OSCE Confidence-Building Measures to Reduce the Risks of Conflict Stemming from the Use of Information and Communication Technologies, PC.DEC/1202 of 10 March 2016, available at: https://www.osce.org/pc/227281?download=true

214 https://www.worldbank.org/en/news/press-release/2020/06/29/kosovo-has-undertaken-critical-steps-in-cybersecurity; https://mzhe-ks.net/index.html.

215 World Bank, 'Kosovo Has Undertaken Critical Steps in Cybersecurity, Says New Cybersecurity Capacity Maturity Model Assessment', available at: https://www.worldbank.org/en/news/press-release/2020/06/29/kosovo-has-undertaken-critical-steps-in-cybersecurity

has commissioned the final evaluation of the Combating Cyber Crime in Kosovo (C3K) proposal in addition to assessing the C3K's timeliness, in order to expand on the experience gained and provide direction for further improvements and procedures in mitigating and countering cybersecurity threats.216

North Macedonia

North Macedonia's regulatory regime is divided into multiple legal documents that address digital security concerns, such as the law on personal data, the law on electronic commerce, the law on electronic communications, the law on interception of communications, the law on free access to public information and the law on data in an electronic form and electronic signature. Similarly, the 2013 revisions to the law on criminal procedure address cybercrime and criminal acts via the use of technology, as well as government agencies' acquisition of digital evidence. A national MKD-CERT operating under the authority of the Agency for Electronic Communication was established in 2015 as a result of the completion of an EU-funded cybersecurity pilot programme as part of the EU ENCYSEC project.217 The GCSCC performed its CMM analysis of North Macedonia in early 2018 in collaboration with the World Bank through the Korea–World Bank Group Partnership Facility. The review's main findings revealed that North Macedonia's cyber ecosystem was still in its initial phases, mostly because online users were unaware of the potential dangers, but showed the country's determination to engage on crucial objectives towards its improvement.218 As a result, the National Cyber Security Strategy for the period 2018–2022, as the main policy statement, focuses on: (1) establishing a cyber-resilient ICT infrastructure, as well as identifying and implementing appropriate solutions to safeguard national interests; (2) encouraging a cybersecurity culture to raise awareness and comprehension of cyber threats, as well as to establish and advance the necessary protective capabilities; (3) improving national capacity; (4) strengthening national defence

216 K. Loshi, UNDP, *Final Evaluation Combating Cyber Crime in Kosovo (C3K)* (Kosovo, 2021).

217 DIPLO, 'Cybersecurity Capacity Building and Research Programme for South-Eastern Europe', Research report, Federal Department of Foreign Affairs of Switzerland (Geneva, 2016).

218 Oxford Martin School, *North Macedonia*, available at: https://gcscc.ox.ac.uk/north-macedonia

capabilities and mitigating current and future cyberspace threats; and (5) domestic and international cooperation and information exchange.219

Albania

Along with the rapid development of digital services, Albania has experienced an increase in numerous types of cybercrime, particularly phishing and spam, leading digital payment crimes. The first National Security Strategy, established in 2014, presented the national framework and pillars for increasing security in the country, and the current strategy for the period 2020–2025220 is in accordance with the government's objective to harmonise its legislative and policy framework with the EU. The plan is noteworthy because, for the first time in the country's history, strategy includes a specific chapter on children's online safety, elevating it to a higher-priority position and reaffirming the state's willingness to keep children safe in all contexts. The strategy also announced the launch of the first national cybercrime strategy, as a significant step forward in the creation of laws and actions against cybercrime.

The principal rules covering cybercrime in Albania are the Criminal Code, law on cybersecurity, law on electronic communications and law on personal data protection. In addition, the National Electronic Certification and Cybersecurity Authority (NECCA) has been established as the authority in charge of supervising the law's implementation within the cybersecurity regulatory regime.

The WB's integration roadmap

The EU and WB countries are intricately intertwined. With 81% of exports and 58% of imports, the EU is the WB's most important trade partner.221 So, it is clear that the WB's participation in the pan-European digital market would lead to a better regional industrial structure, more innovation, fewer administrative

219 Republic of Macedonia National Cyber Security Strategy 2018–2022, available at: https://www.mioa.gov.mk/sites/default/files/pbl_files/documents/strategies/cyber_security_strategy_macedonia_2018-2022_-_eng.pdf

220 Albanian National Cybersecurity Strategy and Its Action Plan 2020–2025, available at: https://www.unicef.org/albania/media/3526/file/Albanian_National_Cybersecurity_Strategy.pdf

221 EUROSTAT, available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=International_trade_in_goods_-_a_statistical_picture

hurdles (needed to make the public sector more efficient), a better flow of information and the protection of individual rights.

The European Commission has scheduled a targeted timeframe of 2025 for WB admission, while underlining that the region's economic underdevelopment is strongly linked to political turmoil. The WB region's turbulent history, as well as the imperative to re-establish links via activities that promote regional cooperation and post-conflict rehabilitation, has altered the geopolitical landscape. Most of the countries in this area have taken on the uniform legal structure of the former Yugoslavia, which is known for having a clear legal identity.

The strengthening of the legal environment for cybercrime responsibility is a development opportunity, a driver of innovation and a cornerstone of a comprehensive 'reform agenda' to boost productivity in this region of tremendous economic potential and strategic importance. There is no evidence that the WB area is underperforming in terms of ICT development, due to the high number of internet users in these nations. Their inhabitants are in danger of being victims of cybercrime as well as being taught to perpetrate crimes using or directed towards ICT.

The assimilation of roughly equivalent frameworks and comparable architectures as stipulated in the NIS Directive, which strives to standardise terminology, methodologies, policies and processes with established European and worldwide standards, is the cornerstone of these national cybersecurity strategies. All WB economies have made commendable initiatives to advance critical policy documents and safety regulations based on assessments in accordance with the desired practice in view of the foregoing EU regulations. Still, for the WB area to be able to deal with cybercrimes effectively, more work needs to be put into harmonising regulatory platforms; raising digital literacy, trust and information security awareness in all sectors; and strengthening cyber-defence capabilities.

The core results of the Cybersecurity Ecosystem Report,222 which reflects security breaches in the WB, critical concerns, incidents and ransomware patterns, commissioned by the UK government, demonstrate that regional cybersecurity actors do not perceive the region as a foremost target of cybercriminals. Therefore, malware and incidents are acknowledged as unexpected consequences of operations on some other priority targets. In particular, decision-makers haven't yet revealed any region-specific security breaches. Meanwhile, considering the fast rate of digital innovation and geopolitical tensions, this shouldn't

222 ISAK, PwC, 'Cybersecurity Ecosystem Report: Western Balkans: Emerging Cyber Threats' (2022).

exclude the likelihood of a stronger regional dimension to cyber-attacks occurring in the fairly near future.

In contrast, flexibility and access to digital activity, especially in recent periods, have triggered a rise in the proportion of security events captured by governmental bodies. A large percentage of the government's digital infrastructure in Montenegro has been rendered inoperable by multiple cyber intrusions starting in August 2022,223 following a mid-July cyber-attack on government sites in neighbouring Albania and a failed attempt on the cyber network of public bodies as revealed by Kosovo.

Considering that all six countries of the WB are at a very similar level of human and technical cybersecurity capabilities, the further development of cybersecurity infrastructure could have substantial consequences for the initiative to fortify the digital integration of the region, and could therefore serve as a catalyst on their path towards European integration. First among the essential guidelines that can be outlined are:

- setting up a regional cyber-risk registry in the form of an interactive database and platform to facilitate the efficient exchange of information between nations on reported events;
- in accordance with S3 strategies,224 promoting cross-sectoral collaboration and developing the network of PPPs, as well as collaborations with academic institutions;
- development of specialised frameworks for identifying needs in key sectors in terms of both technological and human resources, with an emphasis on common vulnerabilities in critical infrastructure;
- consistently monitoring improvements in information security by maintaining a comprehensive periodic assessment;
- introducing new educational programmes that more effectively suit the requirements of the cybersecurity industry as a regional socioeconomic priority and provide measures for altering the labour market;

223 *Reuters*, 'Montenegro blames criminal gang for cyber attacks on government', available at: https://www.reuters.com/world/europe/montenegro-blames-criminal-gang-cyber-attacks-government-2022-08-31/

224 Smart Specialisation is a place-based approach characterised by the identification of strategic areas for intervention based both on the analysis of the strengths and potential of the economy and on an Entrepreneurial Discovery Process (EDP) with wide stakeholder involvement. It is outward-looking and embraces a broad view of innovation including but certainly not limited to technology-driven approaches, supported by effective monitoring mechanisms. See also: https://s3platform.jrc.ec.europa.eu/.

- implementation of a set of cybersecurity awareness programmes aimed at encouraging a cybersecurity culture and implementing cyber-hygiene best practices.
- capacity-building of productive and bureaucracy-free operational processes for combating cybercrime, particularly the institutions responsible for digital investigations and prosecutions of cyber-criminal offences.

Conclusion

The economies of the WB have aggressively sought to fortify the durability of existing national cybersecurity infrastructures. In addition, they have formed bilateral and multilateral collaborations, with a general inclination favouring EU and NATO policies, although with certain national variations.

Considering this remains a nascent industry, it is anticipated that e-commerce customers will encounter an increasing number of hazards due to the current pace of digitisation and the tendencies seen in the past year. The anticipated persistence of a blended work style is expected to result in irreversible changes to countermeasures. To encourage broader participation in countering cybercrime, regional governments must devote more effort to facilitating sectoral capacity development and strengthening PPPs on a regional level.

Cybersecurity is not just a flourishing field; it is also becoming an essential factor for every business and organisation. Especially considering that cyberspace has progressed into a major source of criminal activity, this research highlights the need for further regional empowerment of the platform of specialised bodies for preventing cybercrime and supporting information security and data protection policies, while creating a stronger regional and European coordination platform. WB economies must remain focused on establishing an effective governance platform among the sectors responsible for developing a regional cyber-safe environment capable of systematically reviewing, assessing and adjusting national cyber policies. Even though each country's main projects and level of technical development are slightly different, the WB region needs better regional cooperation, resource integration, interconnectivity and institutional strengthening in the most important parts of the cybersecurity ecosystem.

Raising awareness of cybersecurity at the national level is extremely important and is a precondition for creating a stimulating environment for continuous economic development and digital transformation. Setting up a national

cybersecurity strategic and normative structure can help attain the goals of a safe, trustworthy and long-lasting digital landscape that is supported by high efficiency and built on the integrity of the regional cybersecurity partnership.

At present, the human and logistical resources for criminal prosecutions in the WB region do not complement the acquisition of information security on a broad scale. Aside from the aforementioned issues, the region is constantly confronted with evolving and versatile challenges related to vulnerable institutions and poor trustworthiness of authorities—from policymakers to the media—in addition to other tangled domestic matters that can amplify the impacts of criminal groups, further polarising the general public. In order to overcome these drawbacks, governments must improve their understanding of these security problems, foster trust and undertake long-term activities to mitigate susceptibility to outside impact and keep raising the general level of digital literacy, which is a basic requirement for sustained national and regional cyberesilience. To reduce resilience shortfalls, each country should recognise specific risks and national administrative flaws, and utilise a tailored plan for a safe digital environment. On this basis, it is necessary to optimise the ability to manage security breaches, which involves a modification in organisations, methods, and operational and administrative capabilities. A coherent strategic framework for cybercrime cannot be implemented solely via the utilisation of technology solutions; it must be supplemented by an appropriate and up-to-date legislative model that focuses on the increasing complexity of the ICT environment and the rapidly growing scope of digital threats. Innovative approaches to associating with the workforce, end-users, civil society groups and local communities can be good measures for elevating the process of digitalisation, achieving greater trustworthiness, and potentially building capacities for effective prevention, early detection and coordinated defence against complex cyber-threat vectors. In order to establish long-term and future-oriented regional cyber-resilience, policymakers should devote greater effort to fragmentation, diffusion of technical and financial resources, and lack of institutional coordination through streamlined operations.

Therefore, to encourage sustainable cyber-resilience, WB governments should continue to engage in: greater harmonisation of legal and strategic frameworks with EU legislation; stronger coordination and information exchange between regional partners in the acquisition, analysis and categorisation of data on identified cyber-incidents; regular publication of country reports on security risks and continuous updating of action plans in the direction of novel cyber-security challenges; reinforcing the inclusive platform of public, private, civil and academic sectors in optimally addressing the ICT industry labour market needs; capacity-building of relevant bodies and institutions in charge of cybercrime

investigation and prosecution; and implementation of holistic programmes intended to optimise digital literacy, cyber-risk awareness and general cyber-hygiene. This requires additional efforts to advance cross-sectoral (particularly among four key stakeholders: public, private, business, and academia), interregional and multilateral collaboration, which includes promoting public engagement and critical reasoning; strengthening the PPP network and civil society; and comprehensive digital skills education.

CHAPTER 5

A looking glass on South–South cooperation to strengthen responsibility in cyberspace

MOLIEHI MAKUMANE AND ENRICO CALANDRO

Introduction

The UN discussions on developments in the use of ICT and advancing responsible state behaviour in cyberspace have cast a spotlight on issues from the global South.[225] This part of the world represents approximately 130 of the UN member states and largely comprises countries with overlapping membership from the regions of Latin America, Asia, Africa and Oceania.

The context and environment within which we understand South–South cooperation is not static, but is evolving. The first South–South cooperation[226]

225 *World Population Review* (2022), Global South Countries, https://worldpopulationreview.com/

226 South–South cooperation is defined as the 'technical cooperation among developing countries in the Global South'. It is a tool based on the 1978 Buenos Aires Plan of Action, https://www.unsouthsouth.org/about/about-sstc/

period (SSC 1.0) ran from 1950 to early 2000. Common themes from countries of the Global South are growth; greater visibility of countries and actors of the Global South, including media, political actors and partners; and increased numbers of diplomatic and commercial engagements; but with this, institutions and their capacities have been strained. Mawdsley describes this as typical of the second phase of SSC 2.0, from early 2000 to around 2016.[227] During this second period, 'North' and 'South' identities and agendas are eroding given the increased number of dialogues, and relationships of cooperation and collaboration.[228] As a result, countries of the Global South are articulating and projecting narratives and identities in a complex international landscape in which shared experiences among developing countries are more open to contest.[229]

In the context of multilateral cyber governance, countries of the Global South have the numbers. Particularly in the Open-Ended Working Group (OEWG) on developments in ICT, countries of the Global South have demonstrated considerable substantive negotiating power by leveraging regional groups and alliances in UN negotiations, including on advancing and supporting the resolution on countering the use of information and communications technologies for criminal purposes, which passed with 88 'yes' votes (to 58 'no' votes; 34 abstentions), primarily from countries in the Africa and South Asia regions.[230] The UNGA 73 Developments in the field of information and telecommunications in the context of international security passed with 119 'yes' votes.

Translating numbers to UN processes, and processes to substantive reports that reflect the values, priorities and aspirations of the Global South, is a high ambition. A Global South–South cyber dialogue as a pathway to this will be explored thoroughly in this chapter.

227 Emma Mawdsley, 'South–south cooperation 3.0? Managing the consequences of success in the decade ahead', *Oxford Development Studies* 47 (3) (2019), 259–274.

228 Ibid.

229 Ibid.

230 Third Committee 50th plenary – Item 107, Draft resolution A/C.3/74/L.11/Rev.1, Countering the use of information and communications technologies for criminal purposes, available at: https://www.un.org/en/ga/third/74/docs/voting_sheets/A.C3.74.L.11.Rev.1.pdf

Methodological approach

By adopting a Global South perspective as a magnifying glass on the negotiations processes at the UN OEWG on ICT state security, this chapter aims to unravel what 'responsibility is and is not', as understood by countries of the Global South.

First, we present some considerations on the definition of concepts such as the Global South and South–South cooperation as foundational concepts for this chapter.

Second, the study examines how the concept of responsibility has been introduced and prioritised in this region of the world: (a) by analysing publicly available statements of regional groupings at the OEWG and reviewing them across the thematic areas of international law, cyber norms and cyber capacity; and (b) qualitatively, by conducting semi-structured interviews with purposefully selected senior officers and cyber diplomats from the Global South who are involved in regional groupings at the UN.

We conducted semi-structured interviews between 3 May and 22 May 2022. We invited 11 experts to respond to our questions, and five individuals responded. Due to the small sample of senior public officers interviewed, in this exploratory research, opinions cannot be generalised to all countries of the Global South. We expect future research to validate findings and recommendations.

The interviews were 45 to 60 minutes long and conducted online. The protocol adhered to and upheld ethical conduct. Participants were informed to the greatest possible extent concerning the nature of the research,[231] and they offered their consent to be interviewed. We discussed with the interviewees possible outcomes of the interviews and we did not envisage any potential harm emerging, as we did not engage with vulnerable subjects. Since all our interviewees are public officers who were not officially representing their countries during the interviews, we offered confidentiality and privacy. Therefore, all responses are anonymised.[232] The authors recorded[233] the interviews and took notes during them.

231 Michael Jefford and Rosemary Moore, 'Improvement of informed consent and the quality of consent documents', *The Lancet Oncology* 9 (5) (2008), 485–493.

232 Angelica Orb, Laurel Eisenhauer and Dianne Wynaden, 'Ethics in qualitative research', *Journal of Nursing Scholarship* 33 (1) (2001), 93–96.

233 Recordings are hosted in an encrypted and password-protected online repository.

The authors developed thematic codes[234] to select and narrow down the information received from the interviewees into a consistent and manageable clusters of data. The codes were generated from the topics addressed in the semi-structured questionnaire, and were useful in clustering similar ideas. Subsequently, recurring themes were identified and interviews were analysed on three main issues of the questionnaire:

1. relevance of South–South cooperation for UN discussions on cybersecurity;
2. priorities related to cyber norms, confidence-building measures (CBMs) and cyber capacity-building for the OEWG and in a framework of South–South cooperation;
3. the nexus between cybersecurity and development, and the Global South's approach to cybersecurity capacity through finding local solutions to its development problems.[235]

Defining the Global South

Generally whenever we speak of Global South , we speak of countries in Africa, Latin America and the Caribbean, and areas of Asia and Oceania.[236] According to Haug, three interpretations of the 'Global South' have emerged in studies of world politics.[237] First, the label 'Global South' is commonly used to refer to low-income and economically marginalised areas of the world. Second, the 'Global South' has been portrayed as a site of resistance to neoliberal capitalism. Beyond country-based viewpoints, the 'Global South' has been reframed as a signifier for anti-hegemonic movements that can occur anywhere.

234 Thematic coding involves 'interpreting the information' and categorising textual extracts with reference to 'themes in the context of a theory or conceptual framework'. Richard E. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development* (Thousand Oaks, CA: Sage, 1998), p. 11.

235 The full questionnaire is in the appendix of the chapter.

236 Sebastian Haug, Jacqueline Braveboy-Wagner and Günther Maihold, 'The "Global South" in the study of world politics: examining a meta category', *Third World Quarterly* 42 (9) (2021), 1923–1944.

237 Sebastian Haug, 'What or where is the "Global South"? A social science perspective', *LSE* (28 September 2021), available at: https://blogs.lse.ac.uk/impactofsocialsciences/2021/09/28/what-or-where-is-the-global-south-a-social-science-perspective/

For the purpose of conceptualising a South–South Cooperation Cyber Dialogue, in this study, we define the Global South as cross-regional and multilateral alliances among the developing countries of Latin America, Africa and Asia.[238]

Countries of the Global South are represented in a few formal and regional groupings of the United Nations that include both developed and developing countries. For instance, the Pacific Islands Forum (PIF) comprises 18 members, both small Pacific islands and developed economies such as Australia and New Zealand; similarly, the Organization of American States gathers 35 independent countries of North and South America; the Association of Southeast Asian Nations includes 10 member states with many differences in terms of macroeconomic indicators; the Caribbean Community (CARICOM) is a grouping of 20 countries; and the Group of African States, which includes 53 states, is the largest regional group and speaks in the General Assembly (GA) and not the African Union.

Almost all countries of the Global South are primarily represented in two multilateral alliances: first, the Non-Aligned Movement (NAM), which remains the largest political grouping of countries apart from the UN itself;[239] and second, the Group of 77 (G77), the largest intergovernmental organisation of developing countries in the UN.[240]

The two alliances have different mandates: the G77 provides the means for the countries of the South to articulate and promote their collective economic interests, enhance their joint negotiating capacity and promote South–South co-operation for development; the NAM has a highly visible political role in representing the interests of developing countries, particularly in the eradication of colonialism and supporting struggles for liberation and self-determination.

In 1998, to revitalise the NAM and its structures, the XII Non-Aligned Movement Summit adopted a decision to cooperate with the G77.[241] The decision called for close cooperation to enhance the solidarity of developing countries in the UN system and South–South Co-operation in general. Since then, a Joint Co-ordinating Committee of the NAM and G77 (JCC), co-chaired by the chairs

238 The definition is in line with Haug (ibid.).

239 The NAM is a group of states that are not explicitly aligned with or opposed to any major power bloc. The movement included 120 members and 17 observer states as of 2011. In general (as of 2011), NAM members are all Group of 77 members (along with Belarus and Uzbekistan) that are not observers in the NAM and are not Oceanian states (with the exception of Papua New Guinea and Vanuatu). For the full list of NAM countries see http://cns.miis.edu/nam/about.html#:~:text=The%20Non%2DAligned%20Movement%20(NAM,members%20and%2017%20observer%20countries

240 Group of 77 at the UN, 'About the Group of 77', available at: https://www.g77.org/doc/

241 Department of International Relations and Cooperation (DIRCO), 'Non-Aligned Movement: History and present status', available at: http://www.dirco.gov.za/foreign/Multilateral/inter/nam.htm

of NAM and the G77, meets to coordinate inputs, ensuring that the concerns of the developing countries are represented. The joint views were reflected in the Millennium Summit planning and draft outcome document.[242] The 18th Summit outcome agreed to 'Strengthen coordination and cooperation, as well as the formulation of common strategies with the Group of 77 and China, through the Joint Coordinating Committee (JCC), on issues relative to transnational organized crime'.[243]

The Buenos Aires Plan of Action[244] and the Nairobi Outcome Document of the High-Level United Nations Conference on South–South Cooperation (2010)[245] defined South–South cooperation as 'technical cooperation among developing countries in the Global South' and a 'partnership among equals, based on solidarity and guided by principles of respect for national sovereignty and ownership, free of any conditionality'.[246]

Defining responsibility in cyberspace and in South–South cooperation

As a transnational issue, cybersecurity evokes ideas about multilateral global governance that are related to 'how to manage cyberspace for the good of all'. The same issue surfaces in our attempt to understand roles and responsibilities of countries of the Global South. The suggested unpacking of responsibility as envisaged by the Global South is an attempt to understand whether the current multilateral discussions on cybersecurity allow us to approach the issue of responsibility in a more nuanced way.

Here the 2018 G77 Ministerial Declaration, which reaffirmed that peace requires not only the absence of conflict but also a positive, dynamic participatory process in which dialogue is encouraged, and conflicts are resolved in a spirit of mutual understanding and cooperation, is useful. It reiterated that sustainable

242 Ibid.

243 18th Summit of Heads of State and Government of the Non-Aligned Movement, NAM 2019/CoB/Doc.1, paras 119 and 1133.7.

244 United Nations Office for South–South Cooperation: About South–South and Triangular Cooperation (2022).

245 Nairobi Outcome Document of the High-Level United Nations Conference on South–South Cooperation: draft resolution submitted by the President of the General Assembly (2009).

246 UN meeting coverage of High-Level United Nations Conference on South–South Cooperation (2009).

development cannot exist without peace and that peace cannot exist without sustainable development.[247] In particular, the declaration commits to international efforts directed at safeguarding cyberspace and promoting its exclusive use for the achievement of peaceful purposes and as a vehicle to contribute to both economic and social development; and affirms that international cooperation in accordance with domestic law and as far as international obligations require, as well as fully respecting human rights, is the only viable option for fostering the positive effects of ICT and preventing their potential negative effects.[248]

First, the commitment to 'safeguard cyberspace and promote its exclusive use for the achievement of peaceful purposes and as a vehicle to contribute to both economic and social development' requires that countries of the Global South refrain from adopting any policies, plans and strategies that compromise the exclusive use of cyberspace for peaceful purposes. We cannot take this for granted, as some states have made plans to develop and acquire offensive cyber capabilities that compromise it. Second, most regions are not held to the standard of compliance with national and international obligations as the only viable option for positive effects of ICT. This is a shift in the form of responsibility that transcends statements and language of countries during meetings but signifies a deeper ethos and responsibility to cyber stability than is usually perceived.

In the following sections of this chapter, we examine the statements of regional groupings at the OEWG[249]and analyse the interviews conducted with a purposefully selected group of policymakers and cyber diplomats from the Global South. The analysis is a window into how countries of the Global South experience developments in ICT and how the concept of responsibility is perceived and prioritised. Subsequently, we conceptualise 'South–South cooperation cyber dialogue'.

247 G77, 'Ministerial Declaration' (2018), available at: https://www.g77.org/doc/Declaration2018.htm; Group of 77 at the United Nations, 'About the Group of 77' (2022), available at: https://www.g77.org/doc/

248 Ministerial Declaration of the Group of 77 and China, 42nd annual meeting, New York (2018), para. 176, available at: https://www.g77.org/doc/Declaration2018.htm

249 For this study, we focused only on statements related to international law, cyber norms and cyber capacity-building. While confidence-building measures are important and relevant also for countries in the Global South, they have been omitted from this study as they have been primarily shaped by ideas and suggestions of regional groupings in the Global North.

The notion of responsibility in regional grouping statements

Various intergovernmental organisations representing developing countries and regions made remarks[250] during informal and substantive sessions of the OEWG on developments in ICT in the period 2019–2021. Due to the Covid-19 pandemic, the third and final substantive session in 2020 was cancelled, and rescheduled for 8–12 March 2021 in a hybrid manner. In order to continue its work, the Working Group held informal virtual meetings through June to December 2020 and February 2021.[251]

The effects of the shift to virtual multilateral diplomacy remain to be seen. However, on the activities of the Working Group, outcomes were positive and increased awareness beyond the formal meetings and, as can be seen in the increased number of submissions, especially among regional groups of the Global South. Furthermore, as a result of new and amplified threats to cyber stability during the pandemic, there was an urgency to discuss topics such as 'cyber stability, conflict prevention and capacity building', as was seen by the UN Security Council.[252]

Regional groups made varying proposals: some show a convergence with developed, Global North views while some diverge. Areas of convergence include regional proposals such as the consensus-based approach for developing a framework and the importance of state sovereignty. For instance, during the discussions on international law, the Association of Southeast Asian Nations (ASEAN) stressed the importance of state sovereignty and how international norms and principles derive from sovereignty applied to the conduct of ICT-related activities by states. The responsibility for a peaceful and stable ICT

250 The full list of statements of regional groupings between September 2019 and March 2022 is available at https://docs.google.com/document/d/1Yc2iI-7-upXQ5xye2bNJaLoCc2fGt4RiT0IZTW_zlmM/edit?usp=sharing. CARICOM made five statements. The PIF, an intergovernmental organisation to improve cooperation between Pacific island countries and territories, delivered six remarks. ASEAN, a political and economic union of 10 Southeast Asian states, produced one statement; and the NAM developed two statements. The Organisation of American States, an international body formed to promote solidarity and cooperation among its 34 member states in the Americas, made three comments. Regional groups were also involved in the second cycle of the OEWG, which began in 2021. In the new phase of the OEWG, the NAM has already given five statements, the ASEAN two statements, the OAS three statements, the PIF one statement and the CARICOM one statement. Iran also made a statement on behalf of Belarus, Bolivia, Cuba, Nicaragua, Syria, Venezuela and Iran.

251 Updated Procedural Report: Open-Ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security. A/AC.290/2021/L.1 (2021).

252 Security Council Report – Information and Communication Technologies (2022).

environment to achieve sustainable development objectives was evident in statements that linked an open, safe and secure cyberspace to elements of accessibility. Additionally, states emphasised the importance of a transparent, inclusive, democratic and consensus-based approach for developing a framework to address ICT-related challenges and the need to include all states in this process (CARICOM). Also, they reiterated the importance of respect for human rights and fundamental freedoms in the development of a framework for responsible state behaviour in cyberspace (PIF). These were widely held views across regions, and are reflected in the OEWG substantive report.

States in the OEWG noted how malicious ICT threats can be a threat to state sovereignty, but also the centrality of sovereignty to capacity building. Showing how vital capacity-building is for achieving the Sustainable Development Goals (PIF), almost all regional groupings referred to cybersecurity capacity-building in their interventions at the OEWG. Specifically regarding cyber capacity-building, regional groups encouraged context-specific capacity-building to assist all nations in independently developing opinions on how international law applies to cyberspace and to support the implementation of the rules (PIF). They emphasised the significance of a holistic and 'two-way street' approach to capacity-building and the pressing need to close the digital divide (PIF). They recognised the significance of practical international cooperation in cyber capacity-building and advocated for more cooperation on ICT security and use and the application of international law, norms, rules and principles to state activity in cyberspace (NAM).

During the Group of Governmental Experts (GGE) mandate, submissions of national views on the applicability of international law increased substantially[253] as new avenues of making inputs increased, beyond the UN Secretary General's annual report. The call for additional guidance on norms implementation was also heeded, through an additional layer of understanding on how to implement norms, as seen in the GGE report. In particular, regional groupings underlined the need for both formulating and operationalising cyber norms. They noted the lack of understanding of cyber norms in many parts of the world and asked for practical advice on implementing them (PIF, NAM and CARICOM). They also agreed that identifying and enforcing commonly agreed cyber norms, rules and principles can serve as a starting point for developing an enforceable legal framework (CARICOM). They appreciated that the report acknowledged that cyber norms

253 A compilation can be found on the UNODA website: https://meetings.unoda.org/meeting/oewg-ict-2021/. The UN Institute for Disarmament Research also hosts views, available at: https://cyberpolicyportal.org/

do not replace or alter states' legally enforceable obligations under international law (PIF).

A survey on the implementation of voluntary, non-binding norms was advanced by a substantial number of co-sponsor states in the OEWG. The same applies for the welcoming by consensus of what are now known as the principles of capacity-building. Similarly, they recommended that countries publish their views on applying specific principles of existing international law to cyberspace, thereby supporting the establishment of a central repository of national practices relating to the application of international law (PIF).

Some divergence can be seen in some concepts, and this will require more attention by the states advocating this position, or 'sponsors', especially as it relates to framing.

A couple of concepts have been proposed and framed as threats, namely unilateral actions and climate change. On the latter, states emphasised that cyber resilience is also linked to climate and that that member states must ensure that ICT infrastructure, particularly cybersecurity infrastructure, is climate- and disaster-resilient (PIF).The NAM has invited member states to refrain from adopting unilateral measures that do not follow the UN Charter and international law, arguing that such unilateral actions impede the full achievement of economic and social development by the affected countries' populations and impair their well-being.

The successful framing of these concepts as threats to the responsible use of ICTs in the context of international peace and security will take time and concerted and collective efforts on the part of the sponsors.

On the point of incremental use of voluntary, non-binding norms as a basis for a new legal instrument, it may be valuable to assess the ongoing Ad Hoc Committee process and also to consider other options for addressing legal gaps, should the sponsors find them and articulate them to an intergovernmental body.

Analysis of interviews

Our interviews were in line with official regional statements, and identified a strong demand for the following elements with regard to a South–South Cooperation Cyber Dialogue:

1. strengthen voice of the Global South in the UN OEWG;

2. interregional and cross-regional engagement are key to a Global South dialogue;
3. appreciation of developmental challenges addressed through solidarity and unity
4. consolidation of a Global South approach to responsibility.

The following sections elaborate on these points in detail.

Consolidate a Global South approach to responsibility

Responsibility assumes awareness of one's duties and capacity to uphold them. Our interviews highlighted the low levels of awareness of cybersecurity beyond technical and specialised fields at the national level; at regional and sub-regional levels, low awareness is compounded by competing security challenges and priorities. This was confirmed by a senior diplomat, who indicated that:

> We do now have this framework for responsible state behaviour coming from GGE and OEWG processes, but these processes are not known, it is my feeling: it is at least from my experience, in the Global South, more in specialised sectors, governmental sectors in charge of implementing this framework but not necessarily in a more broader community. (May 2022)

In the realm of international cybersecurity and international peace, interviewees expressed the view that the Global South is a linchpin to global governance of ICT. For a former senior diplomat, low levels of awareness of the governance framework and process should not be ignored, but urgently addressed:

> The Global South is a major player, so whatever norms, framework you develop, if you don't bring them on board it doesn't help, so the observation and the compliance globally, it means will be very low. (May 2022)

Responsibility to safeguard the gains of ICTs for development

Access to technology and the use of ICT for development are a priority for developing and least developed countries of the Global South. This frames the Global South countries' approach to responsibility, explored in the sections below.

> Technologies by themselves are not bad, are not necessarily to be controlled, are not necessarily to be prohibited; for instance, when discussing international security it is very common to go to very quick to the prohibition, to the control, non-proliferation approach, etc., but here with a more tech-neutral approach, where the technologies are neutral – not bad, not good, but neutral. (Senior diplomat, May 2022)

Countries of the Global South acknowledge that just as ICTs can be used for the provision of essential services and developmental goals, they are being used for political and military purposes that contravene UN obligations. It follows that countries of the Global South would call for a balanced approach. A senior diplomat stated that:

> The very first element is balance: we believe that when addressing any international security issue, any international security concerns, we do need to balance all the efforts and all the policy commitments etc. on international security through a balance between also development and human rights concerns, these three pillars; then we will have robust processes when we try to accommodate and balance. (May 2022)

Strengthen voice of the Global South in UN OEWG

The need to work together as the Global South to identify solutions fit for specific socio-economic conditions of countries in the South emerged in a few interviews. For instance, a former senior diplomat stated that:

> The general understanding within the global governance space about South–South cooperation is that countries of the Global South are predominantly under-developing countries ... the idea is that not all the solutions will come from the Global North, but within the Global South, we have to work together. (May 2022)

The NAM has been central in convening countries of the Global South for discussions on ICT governance. In the OEWG 2019–2021, the NAM working paper proposals were reflected in the chair's summary.

Among these are the Global South views on:

- balancing policy commitments with human rights and developments by ensuring resources are not pulled from one of the priorities to another;
- ensuring narratives on development and peaceful uses of ICT are reflected and mainstreamed in the discussions on all pillars of the framework.

Inter-regional and cross-regional engagement are key to a Global South dialogue

All member states of the UN agree that the participation of the multi-stakeholder community is important for the discussions of the UN OEWG. It is important to note low participation of the stakeholder community from the Global South. As a senior diplomat from Latin America stated in an interview, Global South countries have subsidiaries of the stakeholders from developed countries, who already have dialogue channels in their home base countries. Notwithstanding those multistakeholder entities that are also interested in cross-cutting issues and their impact on the Global South, our interviewees stressed the importance of supporting multi-stakeholder community in the Global South and facilitating cross-regional exchange and cooperation so that they can amplify the views of the countries where they operate.

Appreciation of developmental challenges addressed through solidarity and unity

Some countries of the Global South are leading in various sectors, including in the development and deployment of ICT. They differ in their political positions and approaches as much as in their economies. As one interviewee stated:

> To have … South–South conversations at the same level we are having this conversation in the UN, it is absolutely useful to share the plurality of experiences, how we are approaching capacity-building, not

> just taking international technical assistance but developing our own capacity and trying to share to similar countries in the regions. (Senior diplomat, May 2022)

As observed in the introduction to the chapter, Global South countries are not a homogeneous group, as acknowledged by the interviewees. During our interviews, a senior diplomat stated that:

> In the Global South we have a plurality, an absolute plurality, so we do have many countries with not necessarily the same positions but actually with very divergent positions, and we do also have in the Global South this very not necessarily engaged processes to build capacity. (May 2022)

The main development challenges raised by interviewees regarding demonstrating responsibility were policy reorientation and institutional and financial capacity:

- Policy reorientation has to do with countries of the Global South's willingness to make cybersecurity a priority across all of government.
- Institutional capacity can be understood as challenges regarding which government entities are responsible for cybersecurity, and the cybersecurity capacity of those entities to fulfil their duties.
- The financial challenges are related to governments' ability and willingness to route financial resources to those responsible entities including for capacity upskilling and building.

For countries that either are trying to develop ways to address these challenges or have addressed some of them, a South–South cooperation cyber dialogue wherein they could share their experiences is relevant.

Responsibility to maintain regional peace and security

Malicious actors and activities can arise from anywhere, and countries of the Global South are committed to countering the perception that these actors may use their territories with impunity. At the (sub)regional level, across the Global South, there are emerging discussions about how regional partners can engage,

share experiences and assist each other. Our interviews confirmed that (sub)regional bodies play an important role in this regard.

> At the regional level we are just creating right now as we speak some groups [dealing with cybersecurity], some working forces, some fora, some dialogues at the regional level, in the Global South … They have the potential to then accommodate priorities to better implement the UN Framework … by the prioritisation of what is most important for that region in that particular moment. (Senior diplomat, May 2022)

From the example above, it emerged that countries of the Global South also acknowledge that in terms of preserving international security they are not starting from scratch and can identify best practices from other multilateral security processes, and how those measures and processes may be applied to the ICT security space.

Interviewees acknowledged that the use of the CBMs developed, tried and tested for traditional security challenges may be applicable in the cyber realm. For instance, a senior diplomat indicated:

> the relevance of regional organisations when addressing and when trying to develop CBMs, not only the responsible state behaviour norms … but as an initial step to doing so, and delivering and advancing CBMs, I do believe sincerely that through regions and regional organisations we can do our best on CBMs and then more easily, later advancing on the implementation of the general framework for responsible state behaviour. (May 2022)

Responsibility to preserve international security

A rules-based system agreed upon by a multilateral system has been a consistent call from countries of the Global South, as confirmed also by the interviewees. This system has a 'perceived' enforceability to ensure malicious actors do not act with impunity; more positively it may also support countries to prioritise and to bring them to a certain benchmark level. An interviewee stated:

> If we have domesticated or internalised the norms, other countries would know there is a cybersecurity law; this is how they deal with threats, with cyber threats or malicious activities. It creates

> predictability on how we are going to act and I think for other countries in the Global South, it will definitely be the same … It shows that states are willing to prevent unnecessary tension or conflict; it adds to accountability and transparency of states, basically. (Legal adviser to OEWG delegation, May 2022)

Countries of the Global South also engage in this area to demonstrate their 'responsible user' status. As discussions in the OEWG proceed towards implementers and defaulters, countries of the Global South will take opportunities to demonstrate that they are responsible to continue to pursue their development goals and targets only if access and availability to technology advance for them.

Responsibility to implement

As emerged from the analysis of the statements, in the first, second and third substantive sessions of the 2019–2021 OEWG, there were some resounding and consistent calls:

a) for raising awareness on and working on implementation of the framework for responsible state behaviour across all pillars;
b) for a commitment to all 11 norms, rules and principles.

This commitment is echoed among countries of the Global South, for two reasons: first, as a demonstration of responsibility; and second, as a platform (i.e. the OEWG) to articulate challenges towards implementation and therefore responsibility.

> It is important to document it, to report perhaps through the OEWG, on the UNIDIR portal, just to share our policies, our general appreciation of issues, and also our national statements on the application of international law in cyberspace; it will help other countries appreciate what we are doing and be confident enough in us to say these countries are actually doing something about it, we are one, we are doing the same, we are trying to achieve a safe cyber space for everybody. (Legal adviser to OEWG delegation, May 2022)

Policy considerations and recommendations for developing a South–South cooperation cyber dialogue

In this chapter we sought to explore how countries of the Global South have been engaged with discussions on responsible state behaviour in cyberspace. Literature on the Global South and South–South cooperation provide background for us to contextualise and understand these countries' engagement with cyber governance and, parallel to this, to point to a need for a dedicated cooperation cyber dialogue platform for effective conceptualisation, articulation and implementation of the cyber governance aspirations envisaged by the Global South.

This framing of South–South intergovernmental action combined with the founding ethos and the declaration can be a model for how countries of the Global South can establish a 'South–South cooperation cyber dialogue'. The dialogue is important to answer the aforementioned questions of amplification and implementation but also to further elaborate on the proposals made in the OEWG that did not enjoy consensus and have made it into the chair's summary. These include discussions on international law and the need to take steps to avoid and refrain from taking any measures not in accordance with the Charter of the UN and international law. On norms, rules and principles, discussions could cover 'the need to encourage and support further regional efforts as well as partnerships with other stakeholders such as the private sector and the technical community on the implementation of norms'; and on capacity building, discussions could continue on how 'the use of existing platforms within the UN, its specialised agencies and in the wider international community could be used to strengthen already established coordination. These platforms could be used to share national views on capacity-building requirements, encourage the sharing of lessons and experiences from both recipients and providers of support, and facilitate access to information on capacity-building and technical assistance programmes'.[254]

254 Open-Ended Working Group on Developments in the Field of Information and Telecommunications in the Context of International Security, 'Chair's Summary', A/AC.290/2021/CRP.3*, p. 7, para 35.

As seen as the beginning of this chapter, the values of countries of the Global South on responsibility as articulated in collective and individual statements portray a more nuanced approach to responsibility and plans for global governance. That is not to say that the dynamics of economies and regions and mutual interest with regions of the Global North are irrelevant, but rather that for countries of the Global South to truly realise their aspirations they need a set-aside space, a dialogue platform, a South–South cooperation cyber dialogue.

These ideas, while not new, in the context of peace and security discussions have lacked conscious inputs from civil society, academia and the private sector of the Global South, whose work and focus can be sharply different from that of their counterparts in the North, who, while producing excellent research and best practices, lack the grassroots experience that Global South entities offer.

We recognise that analysis of a Global South approach to cybersecurity barely captures the ways in which countries of Global South address cyber concerns. Based on the analysis of the interviews, there are common areas that can be pursued within the cyber dialogue platform.

The proposed recommendations below recognise that some initiatives will need to be phased and, as such, the cyber dialogue should focus first on areas for action, based on existing policies and instruments that can be enacted immediately. Adjustments and flexibilities may be introduced as needed.

The benefits and challenges of ICT will continue to pervade almost every area of life and state conduct for years to come. Therefore, an OEWG process that will continue until 2025 is a good starting point. Nevertheless, further medium- and long-term policy proposals will be needed subsequently.

Based on our analysis, below we suggest action-oriented policy recommendations for a South–South cooperation cyber policy dialogue.

Why a South–South cooperation cyber policy dialogue?

a) The South–South cooperation cyber dialogue can be used as a platform to *continuously raise awareness* of the framework for responsible state behaviour.

b) It can serve as a platform for sharing and consolidating experiences from countries in the Global South on challenges and best practices across these countries on implementing the framework, to improve and learn from each other.

c) It can provide a capacity-building platform based on dialogue, not only between member states but also for non-state actors, including civil society organisations (CSOs), academia and the private sector, from the Global South to learn from each other.

Objectives

The main objective of the South–South cooperation cyber dialogue should be to raise awareness of the framework on responsible state behaviour in cyberspace across all countries in the Global South.

Of the eight objectives of South–South cooperation of the Buenos Aires Plan of Action, two stand out, and can be applied, for developing a South-South cooperation cyber dialogue:

1. Foster the self-reliance of developing countries by enhancing their creative capacity to find solutions to their development problems in keeping with their own aspirations, values and specific needs.
2. Create and strengthen existing technological capacities in the developing countries in order to improve the effectiveness with which such capacities are used and to improve the capacity of developing countries to absorb and adapt technology and skills to meet their specific developmental needs.

Themes

Some of the themes that should be discussed in a South-South cooperation cyber dialogue are:

a) the nexus between 'cyber' and 'development', to reinforce and explain the relationship back into the discussions at the OEWG;
b) responsibility as a development and security practice;
c) good governance, rule of law, human rights, fundamental freedoms, equal access to fair justice systems;
d) unpacking the benefits of implementing a framework on responsible state behaviour in cyberspace to achieve Sustainable Development Goals.

At an institutional level

Nationally, the dialogue should call on all countries of the Global South to continue to engage in the multilateral UN OEWG discussions.

At (sub)regional level, the dialogue should:

1. *Encourage regional organisations to establish and improve regional mechanisms and platforms to discuss international ICT* and the cross-section between development and peace and security. These fora will contribute to raising awareness at the national and (sub) regional levels of international discussions in the UN.
2. *Make efficient use of all existing mechanisms for cross-regional exchanges and multi-stakeholder engagement on international ICT*, to facilitate exchange of experiences and shared concerns.

At a global level, the dialogue should:

3. *Promote the use of existing mechanisms including NAM and G77 JCC* to make joint submissions to the OEWG.
4. Review the NAM working paper to the OEWG 2019–2021 for new inputs to be delivered during the second phase of the OEWG. This can be done during *side-events on the margins of the OEWG*.

Appendix: Interview questions

I am writing to invite you to give a short interview for a discussion paper titled 'A looking glass on South–South Cooperation to strengthen responsibility in cyberspace', which will be submitted as part of Closing the Gap 2022 | Responsibility in Cyberspace: Narratives and Practice.

The paper will consider the following questions:

a) *What are some of the pressing issues related to the framework for responsible state behaviour in cyberspace from a Global South perspective?*

b) *In your own words, how would you define South–South cooperation? And how is it relevant and beneficial for UN discussions on cybersecurity?*
c) *What can intergovernmental bodies at (sub)regional and global level do to advance the priorities of the countries of the Global South?*
d) *Our discussion paper seeks to conceptualise the concept of 'South–South Cyber Dialogue'. What contribution can a 'South–South Cyber Dialogue' provide to the activities of the OEWG, both in terms of amplifying and implementing the recommendations of the first phase, but also to further elaborate new proposals from the Global South?*
e) *How can voluntary norms help countries in the Global South preserve peace and stability in cyberspace? What are their priorities related to cyber norms for the new phase of the OEWG?*
f) *What should countries in the Global South do to build confidence in cyberspace?*
g) *How is the nexus between cyber and development understood in the countries of your region?*
h) *How can countries in the Global South enhance their cyber capacity to find local solutions to their development problems?*

Interviews are scheduled for the week of 9–13 May and will be 30 minutes each. You are also welcome to submit written inputs should you prefer.

Please do let me know if I can provide any additional information at this stage.

We look forward to your (hopefully) positive response and contribution to this study.

NATIONAL PERSPECTIVES

CHAPTER 6

Small state, loud voice

Singapore's regional leadership for norms on responsible state behaviour in cyberspace

MABDA HAERUNNISA FAJRILLA SIDIQ

Introduction

The UN Group of Governmental Experts (GGE) has become an important avenue for discussions on standards of behaviour in cyberspace. Its occasional successes in achieving consensus to publish reports are often celebrated as examples of the group's ability to bridge deep ideational divides among great powers,[255] and the evolving discussion on responsible state behaviour takes place against this backdrop. The first report (A/65/201, 2010) notes

255 See Roger Hurwitz, 'The play of states: norms and security in cyberspace', *American Foreign Policy Interests* 36 (5) (2014), 322–331; Eneken Tikk-Ringas, 'International cyber norms dialogue as an exercise of normative power', *Georgetown Journal of International Affairs* 17 (3) (2016), 47–59; Anders Henriksen, 'The end of the road for the UN GGE process: the future regulation of cyberspace', *Journal of Cybersecurity* 5 (1) (2019), tyy009.

only in passing that differences in perception on state behaviour posed risks of 'instability and misperception'. The subsequent three reports (A/68/98, 2013; A/70/174, 2015; A/76/135, 2021) point to a gradually intensified focus on the topic, especially with the success in recommending 11 voluntary norms on responsible state behaviour encapsulated in the 2015 report.

ASEAN has demonstrated an interesting enthusiasm for the 11 norms, and agreed to 'subscribe in principle' to them while Singapore was its chair in 2018.[256] S Iswaran, as Singapore's Minister for Communications and Information, noted that ASEAN was the 'first and only' regional organisation to express such a commitment to the norms,[257] and committed to refer to them while developing implementable norms.[258]

While these commitments seem declaratory—considering the non-binding character of the norms—ASEAN's decision may be considered impressive due to the highly diverse realities that member states encounter in cyberspace. This is reflected by differences in their priorities and investment in cyber capacities,[259] as well as their stance on government control over cyberspace.[260] Existing literature on responsible state behaviour notes that similar differences have impeded multilateral discussions, e.g. the failure to generate a report during the convening of the UN GGE in 2016–2017.[261] This difficulty is understandable if one considers how great power politics extends to conflicting ideas on security and responsible state behaviour in cyberspace.[262] Therefore, the agreement achieved by ASEAN is arguably significant.

Previous studies on cyber issues in ASEAN have not considered the weight of ASEAN's success in advancing dialogues on cyber norms. Most of the available

256 ASEAN Ministerial Meeting Conference on Cybersecurity, 'Chairman's Statement of the 3rd ASEAN Ministerial Conference on Cybersecurity' (Singapore, 19 September 2018), available at: https://asean.org/speechandstatement/chairmans-statement-of-the-3rd-asean-ministerial-conference-on-cybersecurity/

257 S Iswaran, 'Opening Remarks by Mr S Iswaran, Minister for Communications and Information and Minister-in-Charge of Cybersecurity at SICW Press Conference' (speech, Singapore, 2 August 2018), *CSA Singapore*, available at: https://www.csa.gov.sg/news/speeches/sicw-2019-press-conference

258 ASEAN, 'ASEAN Leaders' Statement on Cybersecurity Cooperation' (Singapore, 28 April 2018), available at: https://asean.org/asean-leaders-statement-on-cybersecurity-cooperation/

259 Candice Tran Dai and Miguel Alberto Gomez, 'Challenges and opportunities for cyber norms in ASEAN', *Journal of Cyber Policy* 3 (2) (2018), 10–13.

260 Benjamin Ang, 'Singapore, ASEAN, and international cybersecurity', in Eneken Tikk and Mika Kerttunen (eds), *Routledge Handbook of International Cybersecurity* (Abingdon: Routledge, 2020).

261 Tim Maurer, 'A dose of realism: the contestation and politics of cyber norms', *Hague Journal on the Rule of Law* 12 (2020), 285–305.

262 Caitríona Heinl, 'Cyber dynamics and world order: enhancing international cyber stability', *Irish Studies in International Affairs* 29 (2018), 53–72; Christian Pauletto, 'Information and telecommunications diplomacy in the context of international security at the United Nations', *Transforming Government: People, Process and Policy* 14 (3) (2020), 351–380.

literature is interested in understanding the landscape of regional cyberspace and how relevant dialogues have progressed,[263] suggesting that regional cybersecurity cooperation is hampered by differences in capacity or national approaches taken by member states. Two studies specifically scrutinise how ASEAN has fared in discussing cyber norms. Candice Tran Dai and Miguel Alberto Gomez argue that cyber norms dialogues in ASEAN are hampered by the absence of a common understanding on cyberspace and cyber threats, citing ontological unity as a necessary precondition.[264] Hanan Mohamed Ali further finds ASEAN practices to be 'norm subsidiary', suggesting that the inclusion of international norms is deemed secondary to the pre-existing regional norms.[265]

These studies tend to provide limited analyses on how ASEAN works around existing capacity and contextual gaps. While they have correctly identified how these gaps might materialise, none, with the exception of Ali's research, analyses the actions and consequences of ASEAN's efforts to address these gaps. Moreover, despite the general agreement on the diverse nature of the region, these studies fail to consider how this diversity is manifested in the agential forms that different actors may take, assuming that ASEAN acts as a collective and neglecting the possibility that different member states might take different roles.

I intend to further scrutinise regional cyber norms-building in ASEAN. So as not to take the question of agency for granted, I focus on Singapore's role as a norm entrepreneur in the organisation. This focus is attributed to Singapore's chairmanship in ASEAN at the time when the decision to subscribe in principle to the UN GGE norms was made. The analysis departs from literature on regional leadership to argue that Singapore has effectively and constantly committed to the role of regional leader in developing discussions and actions on cyber norms. This commitment extends beyond its chairmanship in ASEAN, as it is well integrated into its own broader cybersecurity and foreign policy agenda. Singapore displays a strong will to translate its strong cyber capacity into initiatives on cyber norms and cybersecurity, which earns its leadership legitimacy in the eyes of other ASEAN member states. This argument will be elaborated in four sections, consecutively discussing the conceptual framework and Singapore's will,

263 Khanisa Krisman, 'A secure connection: finding the form of ASEAN cyber security cooperation', *Journal of ASEAN Studies* 1 (1) (2013), 41–53; Caitríona H. Heinl, 'Regional cybersecurity: moving toward a resilient ASEAN cybersecurity regime', *Asia Policy* 18 (July 2014), 131–160.

264 Tran Dai and Gomez (see note 5 above).

265 Hanan Mohamed Ali, '"Norm subsidiarity" or "norm diffusion"? A cross-regional examination of norms in ASEAN-GCC cybersecurity governance', *Journal of Intelligence, Conflict, and Warfare* 4 (1) (2021), 122–148.

capacity, and legitimacy to exercise leadership, as well as giving a conclusion and recommendations.

Regional leadership (and the lack thereof) in ASEAN

Discussions on leadership in international relations literature mostly aim to comprehend how actors exercise agency in international institutions. Understandably, they often overlap with, or are engulfed by, conceptual debates on hegemony.[266] While hegemony is associated with the attainment of the hegemon's own goals and interests, leaders aim for actual or perceived collective goals.[267] Leadership should entail 'followership' as the audience accede to the pursuit of collective goals set by the leader.[268]

Through a similar logic, discussions on regional leadership often begin with untangling the conceptual confusion between regional leader, regional power and regional hegemon. A general conclusion is that leadership allows greater room for variety in how both the leaders and the followers display their agency. Leadership might be initiated by the leaders themselves by taking a similar role to that of norm entrepreneurs,[269] emerging out of their interest to socialise their regional neighbours into certain norms.[270] Followers are also able to initiate leadership, typically emerging from their own needs for a leader to achieve collective goals.[271] Singapore's role in advancing regional cyber norms-building illustrates its willingness to take the lead as a norm entrepreneur.

266 Charles P. Kindleberger, 'Dominance and leadership in the international economy: exploitation, public goods, and free rides', *International Studies Quarterly* 25 (2) (1981), 243.

267 Oran R. Young, 'Political leadership and regime formation: on the development of institutions in international society', *International Organization* 45 (3) (1991), 285; Jens Heibach, 'Public diplomacy and regional leadership struggles: the case of Saudi Arabia', *International Politics* (2021), available at: https://doi.org/10.1057/s41311-021-00310-7

268 Stefan A. Schirm, 'Leaders in need of followers: emerging powers in global governance', *European Journal of International Relations* 16 (2) (2010), 197–221.

269 See Amitav Acharya, 'How ideas spread: whose norms matter? Norm localization and institutional change in Asian regionalism', *International Organization* 58 (Spring 2004), 244–250.

270 Sandra Destradi, 'Regional powers and their strategies: empire, hegemony, and leadership', *Review of International Studies* 36 (4) (2010), 921–925.

271 Destradi (see note 16 above), 924–925.

A regional leader might be inclined to take the leading position when it is willing, capable of doing so, and legitimated by its followers. First, the willingness to lead comes from certain norms or values that shape the leader's interest in leading. Subsequently, the leader might make claims to represent their self-identification as a regional leader. Second, the leader should have the capacity to translate its material and/or ideational prowess to gain followership from its neighbours. Lastly, leadership should be acknowledged by other regional and/or external actors.[272] These three components form an important frame to comprehend how Singapore's leadership takes shape and is acknowledged in the region.

Within the context of internal ASEAN relations, pinpointing which states take the leading position necessitates clarity on whether formal or informal forms of regional leadership are in question.[273] Formal leadership is usually conferred by a member state's position within ASEAN's institutional arrangements, such as the chairman position. Alternatively, informal leadership is grounded on the leader's influence within the region. The latter form of leadership has received greater scrutiny. Conventional readings of regional leadership in ASEAN overwhelmingly argue that Indonesia's material prowess grants it the leverage to take the leading position.[274] This argument has been challenged, as no ASEAN member state is able to single-handedly influence regional initiatives across all the issues that the organisation covers.

Instead, many scholars have turned to the idea of 'sectoral leadership' to describe regional leaders in ASEAN. Regional leaders only possess specific material and ideational resources that would allow them to legitimately exercise leadership in a specific issue area.[275] I suggest that Singapore's leadership in promoting cyber norms corresponds to both formal and informal forms of leadership. While its formal leadership as the ASEAN chair in 2018 afforded Singapore room for agenda-setting, its ability to maintain and intensify dialogues on the issue demonstrated its willingness and ability to legitimise its sectoral leadership.

272 Jinsoo Park, 'Regional leadership dynamics and the evolution of East Asian regionalism', *Pacific Focus* XXVII (2) (2012), 293–295; Marieke Zwartjes, Luk Van Langenhove, Stephen Kingah and L. Maes, 'Determinants of regional leadership: is the European Union a leading regional actor in peace and security?', *Southeast European and Black Sea Studies* 12 (3) (2012), 395–396.

273 Pattharapong Rattanasevee, 'Leadership in ASEAN: the role of Indonesia reconsidered', *Asian Journal of Political Science* 22 (2) (2014), 118–120; Ralf Emmers and Huong Le Thu, 'Vietnam and the search for security leadership in ASEAN', *Asian Security* 17 (1) (2021), 65–66.

274 Anthony Smith, 'Indonesia's role in ASEAN: the end of leadership?', *Contemporary Southeast Asia* 21 (2) (1999), 238–260; Linda Quayle, 'Indonesia, the ASEAN socio-cultural community, and the contingent profile of regional "great-power management"', *Pacific Review* 31 (2) (2018), 131–150.

275 Ralf Emmers, 'Indonesia's role in ASEAN: a case of incomplete and sectorial leadership', *Pacific Review* 27 (4) (2014), 543–562; Rattanasevee (see note 19 above), 118; Emmers and Thu (see note 19 above).

Norms on responsible state behaviour and Singapore's foreign policy

Singapore's cyber policy has long emphasised the importance of norms on responsible state behaviour, often directly linking regional initiatives with the UN GGE's endeavours to formulate these norms. Its high level of enthusiasm represents its willingness to exercise regional leadership to promote norms encapsulated in the 2015 UN GGE report. To date, Singapore has released two documents, in 2016 and 2021, to outline its cybersecurity strategy. Both present international partnerships or cooperation as one of their pillars, aiming to establish 'a rules-based multilateral order' in the long run.[276]

Both strategies similarly sketch three forms of actions that Singapore pursues in building international partnerships. First, Singapore seeks to address cyber threats and cybercrime so as to account for their transnational character. It also strives to facilitate capacity-building activities, mostly catering to ASEAN member states. Lastly, Singapore expresses its commitment to greatly contribute to dialogues on cyber norms.[277]

Juxtaposing how the 2016 and 2021 strategies address cyber norms reveals that Singapore has expanded its scope of activities and targeted more precise goals in building partnerships. The 2016 strategy focuses on fostering dialogues with other ASEAN member states.[278] The 2021 strategy states that Singapore seeks to simultaneously participate in dialogues under the UN. It cultivates avenues to build on the 2015 UN GGE report and develop implementable norms, including a partnership with the UN Office of Disarmament Affairs (UNODA) and Malaysia to formulate a regional action plan on cyber norms.[279]

276 CSA Singapore, *Singapore's Cybersecurity Strategy* (2016), 4–5, available at: https://www.csa.gov.sg/-/media/Csa/Documents/Publications/SingaporeCybersecurityStrategy.pdf; CSA Singapore, *The Singapore Cybersecurity Strategy 2021* (2021), 4, available at: https://www.csa.gov.sg/News/Publications/singapore-cybersecurity-strategy-2021

277 CSA Singapore, *Singapore's Cybersecurity Strategy*, 42–47; CSA Singapore, *The Singapore Cybersecurity Strategy 2021*, 31–40

278 CSA Singapore, *Singapore's Cybersecurity Strategy*, 42–47.

279 CSA Singapore, *The Singapore Cybersecurity Strategy 2021*, 31–40.

Singapore's push to implement the 11 norms on responsible state behaviour, as well as its fervent engagement in discussions relevant to these norms, departs from its belief that cyber norms are imperative for regional and international cybersecurity. Singapore frames cybersecurity as an 'enabler' towards 'social development and economic progress'.[280] The linkages drawn between cybersecurity and the economy indicate the urgency of cyber norms within Singapore's foreign policy posture, due to its identity as a small and 'smart' state.[281] As a small state, Singapore is wary of potential consequences from great power conflicts,[282] while it is very much dependent on ICTs as a financial and transportation hub in Southeast Asia and even Asia-Pacific.[283] Moreover, cybersecurity is central to Singapore's 'smart nation' initiative, embracing digital technologies to facilitate a substantial part of its public and non-public services.[284] Given these characteristics, it is reasonable that cybersecurity may be viewed as pivotal for Singapore's 'future survival and growth', motivating the city state to invest in collective initiatives within ASEAN.[285]

Singapore views regional cybersecurity as a 'team sport', meaning that all ASEAN member states should partake in ensuring cybersecurity.[286] Acting collectively to develop cyber norms allows the organisation to display a unified posture as a region, enhancing its ability to have its regional perspectives accounted for in international platforms.[287] In this respect, it is also important to consider that great power rivalry often permeates into discussions on cyber norms, as

280 Yaacob Ibrahim, 'Intervention by Dr Yaacob Ibrahim, Minister For Communications and Information and Minister-In-Charge of Cyber Security at the 15th ASEAN TELMIN' (speech, Da Nang, 27 November 2015), *CSA Singapore*, available at: https://www.csa.gov.sg/news/speeches/intervention-by-minister-yaacob-ibrahim-at-the-15th-asean-telmin

281 CSA Singapore, *The Singapore Cybersecurity Strategy 2021*; United Nations General Assembly, *Developments in the field of information and telecommunications in the context of international security: Report of the Secretary-General*, A/72/315 (11 August 2017), available at: https://daccess-ods.un.org/access.nsf/Get?OpenAgent&DS=A/72/315&Lang=E

282 Ang (see note 6 above).

283 David Koh, 'Keynote address by Mr David Koh, Chief Executive, Cyber Security Agency of Singapore, at the 3rd Annual Billington International Cybersecurity Summit' (speech, Washington, DC, 21 March 2018), *CSA Singapore*, available at: https://www.csa.gov.sg/news/speeches/the-3rd-annual-billington-international-cybersecurity-summit-keynote-address-by-ce

284 Smart Nation and Digital Government Office, 'Smart Nation: the way forward' (executive summary), available at: https://www.smartnation.gov.sg/files/publications/smart-nation-strategy-nov2018.pdf

285 Ang (see note 6 above).

286 Ibrahim (see note 26 above).

287 Yaacob Ibrahim, 'Opening Speech by Dr Yaacob Ibrahim, Minister for Communications and Information and Minister-In-Charge of Cyber Security, at the ASEAN Ministerial Conference on Cybersecurity' (speech, Singapore, 18 September 2017), *CSA Singapore*, available at: https://www.csa.gov.sg/news/speeches/asean-ministerial-conference-on-cybersecurity-2017

exemplified by the failure to reach consensus during the 2016–2017 UN GGE.[288] Bearing in mind the collectivity of ASEAN would have improved Singapore's standing in international dialogues, especially during its membership in the 2019–2021 UN GGE and chairing of the 2021–2025 Open-Ended Working Group (OEWG). While it is difficult to draw direct correlation between a particular member's contribution and the final UN GGE report, the 2021 UN GGE dialogue embodies a remarkable achievement as it improves the operability of the voluntary norms in the 2015 report.

Singapore's identity as a small and smart nation, as well as its framing of regional cybersecurity as a 'team sport', clearly motivates its will to lead activities on regional cyber norms. Coupled with this will is Singapore's own capacity as a global digital powerhouse, which allows it to invest its own cyber capacity into its leadership endeavours. Looking back at Singapore's cybersecurity strategy, both the 2016 and 2021 versions clearly express that Singapore is interested in investing substantial capital in facilitating dialogues and capacity-building.

Many Singapore-led regional initiatives are designed and conducted by its Cyber Security Agency (CSA). Central to the agency's work in regional dialogues is its annual inauguration of the ASEAN Ministerial Meeting on Cybersecurity (AMCC) since 2016. Prior to this, ASEAN did not have a platform specifically dedicated to cyber issues, which were mostly discussed within wider dialogues on ICTs under the ASEAN Telecommunications and Information Technology Ministers Meeting (TELMIN)[289] or other meetings under the political–security pillars. The AMCC itself is held as an event of the Singapore International Cyber Week, indicating that to date, the meetings have always been hosted by Singapore. Along with the AMCC, CSA has invited the UN to participate in regional cyber norms-building through the UN–Singapore Cyber Program. Established in 2018, the programme is designed to support ASEAN member states in collaboratively developing national-level 'policy, strategy and operational practice' in cybersecurity.[290]

288 Alex Grigsby, 'The end of cyber norms', *Survival* 59 (6) (2017), 109–122; Daniëlle Flonk, Markus Jachtenfuchs and Anke S. Obendiek, 'Authority conflicts in internet governance: liberals vs. sovereigntists?', *Global Constitutionalism* 9 (2) (2020), 364–386.

289 TELMIN was renamed ASEAN Digital Ministers Meeting (ADGMIN) in 2019 to represent its expanded scope of dialogue to cover digital issues.

290 CSA Singapore, 'United Nations–Singapore Cyber Programme: Senior Executives Cyber Fellowship and Workshop on Implementation of Norms and Confidence Building Measures' (fact sheet, 2019), available at: https://www.csa.gov.sg/-/media/csa/documents/sicw_2019/amcc/factsheet---unscp.pdf

Singapore also invests greatly in regional cyber capacity building. In 2016, Singapore spent a total of SGD10 million[291] to launch the ASEAN Cyber Capacity Program (ACCP), providing training on technical, policy and strategy development for ASEAN member states. Singapore has also engaged external partners, as the ACCP receives support from Singapore's cybersecurity cooperation with the United States, the United Kingdom, Canada and Cisco Systems.[292] The programme was extended through the establishment of the ASEAN–Singapore Cybersecurity Center of Excellence (ASCCE) in 2018. The centre was supported by a funding commitment amounting to SGD30 million. Aside from holding programmes to support training, exercises and best practices, the centre also conducts research on cyber norms and related policy issues.[293] A training facility for the ASCCE was built in Singapore and has been operational since 2019. These initiatives speak to Singapore's strong will to lead and utilise its cyber capacity in regional cyber norms building.

Gaining legitimacy in ASEAN: fitting the frames into the locale

It may seem obvious that Singapore's seemingly one-sided efforts grant legitimacy to its leadership. Moreover, the monumental decision to subscribe to the norms on responsible state behaviour was achieved during Singapore's time as chairmanship. However, Tran Dai and Gomez's study observes that while decision-making processes in member states such as Singapore and Malaysia have taken cyber threats very much into consideration, other states have only been

291 CSA Singapore, 'Factsheet on ASEAN Cyber Capacity Programme', available at: https://www.csa.gov.sg/-/media/csa/documents/sicw2016/amcc/factsheet_accp_final.pdf

292 See 'Singapore and the United States Sign Declaration of Intent on Cybersecurity Technical Assistance Programme', *CSA Singapore* (16 November 2018), available at: https://www.csa.gov.sg/news/press-releases/singapore-and-the-us-sign-doi-on-cybersecurity-technical-assistance-programme; 'Singapore Signs Memorandum of Cooperation on Cybersecurity Capacity Building with the United Kingdom', *CSA Singapore* (17 April 2018), available at: https://www.csa.gov.sg/news/press-releases/singapore-signs-memorandum-of-cooperation-on-cybersecurity-capacity-building-with-the-united-kingdom; 'Singapore Signs Memorandum of Understanding with Canada on Cybersecurity Cooperation', *CSA Singapore* (14 November 2018), available at: https://www.csa.gov.sg/news/press-releases/singapore-signs-memorandum-of-understanding-with-canada-on-cybersecurity-cooperation; 'CSA and Cisco Systems Sign Memorandum of Collaboration to Establish a Framework for Cybersecurity Cooperation', *CSA Singapore* (29 November 2018), available at: https://www.csa.gov.sg/news/press-releases/csa-and-cisco-systems-sign-memorandum-of-collaboration

293 CSA Singapore, 'ASEAN–Singapore Cybersecurity Centre of Excellence (ASCCE)' (fact sheet, 2019), available at: https://www.csa.gov.sg/-/media/csa/documents/sicw_2019/amcc/factsheet-ascce-2019.pdf

able to partially address cyber threats, or have yet to recognise them as threats due to their limited use of ICTs.[294] Not all member states deem cyber norms to be necessary. Taking Singapore's leadership for granted precludes further scrutiny of how Singapore attempts to mould its cyber norms agenda to cater to the needs of ASEAN member states, thus giving its leadership the necessary legitimacy.

While initiatives noted in the previous sections are primarily led and hosted by Singapore, much of their progress is extended into actions taken under the purview of the ASEAN institutional framework. This suggests that the predominantly Singapore-led initiatives have now been absorbed into more formal and less ad-hoc actions taken directly under the auspices of ASEAN's collective works as a regional organisation, and points to ASEAN member states' acceptance of Singapore's leadership. One of the most notable successes is the release of the ASEAN Leaders' Statement on Cybersecurity Cooperation, agreed during the 32nd ASEAN Summit in 2018. The ASEAN Summit is conducted annually as 'the highest policy-making body in ASEAN'.[295] The statement instructed ministers in charge of several areas, such as ICTs and transnational crime, to formulate implementable norms out of the 2015 UN GGE report.[296]

Another important success is the inclusion of norms implementation as one of the key initiatives outlined in the most recent draft of the ASEAN Cybersecurity Cooperation Strategy (2021–2025), with the objective of enhancing regional-level collaboration in cyber policy. This initiative is built on the agreement to 'subscribe in principle' to the UN GGE norms on responsible state behaviour. Malaysia and Singapore co-proposed the ASEAN Regional Action Plan on the Implementation of Norms of Responsible State Behavior in Cyberspace. The plan towards implementation starts with initiatives that all member states will certainly be able to accept, such as capacity building, and will be periodically updated in the future.[297] Taking an incremental approach is necessary to ensure that no member states are left behind. While more ambitious goals seem to be compromised, ASEAN has traditionally preferred to opt for activities that

294 Tran Dai and Gomez (see note 5 above), 12.

295 See 'ASEAN Summit', ASEAN, available at: https://asean.org/about-us/asean-summit/

296 ASEAN, ASEAN Leaders' Statement on Cybersecurity Cooperation.

297 ASEAN, ASEAN Cybersecurity Cooperation Strategy (2021–2025) (draft, 2022), available at: https://asean.org/wp-content/uploads/2022/02/01-ASEAN-Cybersecurity-Cooperation-Paper-2021-2025_final-23-0122.pdf

accommodate all member states[298] and take the 'lowest common denominator'.[299] This approach is indicative of Singapore's consideration of the diverse nature of the region's digital landscape.

The previous two examples are important reminders that Singapore does not diverge from pre-existing institutional mechanisms and cultures available in ASEAN. The AMCC itself was designed to build on works under the TELMIN,[300] a ministerial dialogue that predominantly addresses the use of ICTs within the region's economy. Central to TELMIN's cooperation are initiatives to foster the region's digital sector by promoting the digital economy and preparing adequate resources for digitalisation.[301] Singaporean representatives repeatedly referred to the great potential of ASEAN's digital market in several past addresses in the AMCC, claiming that the regional digital market was predicted to grow to US$200 billion in 2025.[302] Going back to Singapore's attempt to frame cybersecurity as 'an enabler' for economic development, this points to an attempt on Singapore's part to build direct issue linkages between cyber norms and the digital economy to attract support from ASEAN member states.

Weaving cyber norms into ASEAN's digital economy agenda also allows Singapore to tap into the region's pursuit of integration, of which connectivity is an important component. The bulk of the work on digital integration is concerned with narrowing the digital divide among ASEAN member states, aiming to improve access to and adoption of ICTs across the region.[303] Singapore's choice to focus on capacity-building activities to advance cyber norms-building is very much aligned with the direction that ASEAN's cooperation in digital economy has taken. Building direct issue linkages between the digital economy and cyber norms emphasises the contingent nature of both issues, which might have

298 See Gillian Goh, 'The "ASEAN Way": non-intervention and ASEAN's role in conflict management', *Stanford Journal of East Asian Affairs* 3 (1) (2003), 114–115; Kei Koga, 'The normative power of the "ASEAN Way"', *Stanford Journal of East Asian Affairs* 10 (1) (2010), 81.

299 See Bilson Kurus, 'The ASEAN triad: national interest, consensus-seeking, and economic co-operation', *Contemporary Southeast Asia* 16 (4) (1995), 406.

300 'ASEAN member states call for tighter cybersecurity coordination in ASEAN', *CSA Singapore* (11 October 2016), available at: https://www.csa.gov.sg/news/press-releases/asean-member-states-call-for-tighter-cybersecurity-coordination-in-asean

301 'ASEAN digital sector', ASEAN, available at: https://asean.org/our-communities/economic-community/asean-digital-sector

302 Ibrahim (see note 33 above); S. Iswaran, 'Opening Speech by Mr S Iswaran, Minister for Communications and Information, Minister-in-Charge of Cybersecurity, at the ASEAN Ministerial Conference on Cybersecurity 2019' (speech, Singapore, 2 October 2019), *CSA Singapore*, available at: https://www.csa.gov.sg/news/speeches/asean-ministerial-conference-on-cybersecurity-2019

303 ASEAN, *ASEAN ICT Masterplan 2015* (Jakarta: ASEAN Secretariat, 2011), 9; ASEAN, *The ASEAN ICT Masterplan 2020* (Jakarta: ASEAN Secretariat, 2015), 5; ASEAN, *ASEAN Digital Masterplan 2025* (Jakarta: ASEAN Secretariat, 2021), 11.

contributed to the acceptance of Singapore's leadership by other ASEAN member states.

Beyond the confines of ASEAN dialogues, Singapore's chairmanship of the 2021–2025 OEWG may have contributed to its legitimacy. This is demonstrated by statements made by ASEAN member states in the OEWG, promoting measures such as confidence-building measures (CBMs) and capacity-building programmes conducted under ASEAN.[304] Moreover, ASEAN collectively issued a statement, delivered by Brunei Darussalam, the 2021 chair of ASEAN, to introduce its general views in the OEWG in December 2021 The statement outlined ASEAN's general approach to cyber issues and promoted its Regional Action Plan to implement cyber norms.[305] These statements are instructive of Singapore's success in its efforts to 'amplify ASEAN's voice'[306] in international forums.

Although these successes are laudable, Singapore's one-sided approach has limitations. One of the most obvious is its highly central position in pioneering discussions or initiatives on cyber issues within the region, meaning that to date, Singapore's leadership has almost certainly been necessary to initiate cyber cooperation. Aside from Thailand headquartering the ASEAN–Japan Cybersecurity Capacity Building Centre and Malaysia co-proposing the Regional Action Plan, no initiatives are funded or led by other ASEAN member states. This trend might put a strain on the sustainability of Singapore's support for cyber cooperation in ASEAN.

Moreover, in amplifying ASEAN's voice, Singapore meets challenges in generating collective ASEAN-level support for initiatives beyond the ASEAN purview. For instance, EU Member States were able to submit the 2020 proposal to establish a Programme of Action (PoA) to end the dual GGE–OEWG dialogues as a collective, while Singapore was the only ASEAN member state to join this submission.[307] While the absence of other ASEAN member states might be for various reasons, it may be inferred that the support Singapore receives at the

304 'Thailand's Intervention on Confidence-building Measures at the First Substantive Session of the Open-ended working group (OEWG) on security of and in the use of information and communications technologies 2021–2025' (New York, 16 December 2021), available at: https://documents.unoda.org/wp-content/uploads/2021/12/Thailand-CBMs-OEWG.pdf; K.M.M. Malang, 'Intervention: On how international law applies to the use of information and communications technologies by States' (New York, 16 December 2021), available at: https://documents.unoda.org/wp-content/uploads/2022/02/PHILIPPINE-INTERVENTION_Capacity-building-REV.pdf

305 Noor Qamar Sulaiman, 'Statement on Behalf of the Association of Southeast Asian Nations (ASEAN)' (New York, 13 December 2021), available at: https://documents.unoda.org/wp-content/uploads/2021/12/ASEAN-Statement-OEWG-First-Substantive-131221.pdf

306 Ibrahim (see note 33 above).

307 United Nations, 'The future of discussions on ICTs and cyberspace at the UN' (proposal, 10 August 2020), available at: https://front.un-arm.org/wp-content/uploads/2020/10/joint-contribution-poa-future-of-cyber-discussions-at-un-10-08-2020.pdf

regional level might not be easily extended to international dialogues. Speaking collectively as ASEAN has been limited to promoting regional initiatives taken under ASEAN.

Conclusion

Singapore has demonstrated great willingness, capacity and legitimacy to lead regional norms-building on responsible state behaviour in cyberspace. Its motivation is rooted in Singapore's identity as a small state that relies on cyberspace to ensure its survival and livelihood. It reasonably views cyber norms as highly relevant to its economic concerns. In pursuit of its interest, Singapore has thus designed multiple initiatives to ensure that cyber norms will be on ASEAN leaders' radar. To guarantee acceptance by other ASEAN member states, Singapore has accommodated the needs and interests of its regional neighbours. By highlighting the entanglement between cyber norms and digital economy, coupled with a great emphasis on capacity-building programmes, Singapore seems to have succeeded in gaining the interest of other ASEAN member states.

Many important lessons can be drawn from Singapore's advances as both a regional leader and a small state. First, Singapore's regional leadership demonstrates that a willingness to capitalise on strong cyber capacity is necessary for cyber norm entrepreneurship. While its strong motivation might have stemmed from its interests as a small and smart state, it is also important to account for its effectiveness in moulding these interests into a more palatable narrative that other ASEAN member states can easily accept as their own. Second, crucial to Singapore's leadership is its great awareness of the regional cyber landscape, demonstrating its willingness to develop implementable norms at a pace that all member states can accept. This is coupled with its choice to situate the initiative within prior ASEAN initiatives and agendas. Both of these features are translated into Singapore's inclusive and incremental approach. This approach should be considered if other regional leaders are interested in advancing regional norms-building, especially in regions with diverse cyberspace profiles.

An important caveat in Singapore's leadership is its strongly one-sided approach. While the regional context necessitated this approach, Singapore should consider whether it will be sustainable in the long run. Singapore needs to foster genuine interest from ASEAN member states to address cyber issues. More crucially, it might be in Singapore's interest to ensure that other member states will also invest in cyber cooperation, most crucially by providing material support.

CHAPTER 7

What does Nigeria's national identity server downtime suggest about accountability and cyber norms in local CERTs?

An exploratory study

BABATUNDE OKUNOYE

Overview

Backed by UN Sustainable Development Goal (SDG) 16.9 ('By 2030 provide legal identity for all, including birth registration'), many governments across the world have begun implementing national (digital) identity projects. Set within the context of UN SDG 16 ('Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable, and inclusive institutions at all levels'), national identity projects are designed to confer legal identities on those who lack

identities, including 1 billion globally and 400 million in Africa. This enables participation in the national and global digital economies as millions are empowered to complete transactions with governmental, public and private organisations. Nigeria's national identity project, administered by the National Identity Management Commission (NIMC), is Africa's biggest, catering for a population of over 200 million people, and had enrolled 77.1 million individuals as of March 2022. Upon enrolment for national identity, individuals are assigned a national identity number (NIN), which became mandatory in January 2019 for critical transactions including, among others, banking services, receiving government social grants, obtaining an international passport or driver's licence, registering a SIM card and participating in national health insurance.

National identity programmes are basically very large information systems that in many nations are classified as critical national infrastructure (CNI) and critical information infrastructure (CII), and their administration can teach important lessons on best practices in responsible state behaviour, accountability and cyber norms. In February 2022, the server hosting the NIN verification platform experienced a lengthy downtime of over 10 days. As service providers were unable to verify the NINs of people needing the services for which the NIN was made compulsory, many people across the country were stranded with no effective alternatives. The critical sectors that witnessed interrupted services included the financial and telecommunications sectors, which, alongside the NIMC, are mandated to have sectoral-level computer emergency response teams (CERTs), according to Nigeria's revised (2021) Cybersecurity Policy. The telecommunications sector was reported to have incurred revenue losses occasioned by the server downtime. Through analysis of research literature and expert interviews with cybersecurity experts in the country knowledgeable about the case of the NIN server downtime, this chapter reviews a significant cyber-related incident in Nigeria. It shows how gaps in the implementation of specific cyber norms and obligations relating to the working of sectoral CERTs might have contributed to the unusual length of recovery (over 10 days) for a critical national infrastructure for which there were no effective alternatives, and suggests avenues for strengthening cybsersecurity capacity and cooperation.

The context: Nigeria's national identity database as critical national infrastructure and critical information infrastructure

Nigeria's current digital identity project commenced in 2007 with the establishment of the National Identity Management Commission (NIMC) and the passage of the National Identity Management Commission (NIMC) Act of 2007.[308] Alongside the NIMC Act, several policies were created to support the administration of the nation's identity project, including on privacy, biometric standards and a revised national identity policy.[309] Nigeria's identity project was backed by international donors such as the World Bank and the European Investment Bank.

Thus began the third iteration of a national identity project that had failed in two previous attempts.[310] Nigeria has Africa's biggest population, numbering 200 million people, and the digital identity project was planned to correct some of the systemic gaps in civil registration data capture such as low birth registration rates in the country specifically, but also generally in Africa.[311] For example, the possession of a birth certificate is a precious thing in sub-Saharan Africa, where 120 million children under the age of five do not have one. In Nigeria, more than half of the births of children under the age of five remain unregistered,[312] a fact that restricts access to education, social services and financial inclusion in later life. The national identity project was also targeted as a solution to the rising spate of insecurity in the country,[313] which manifested in its most

308 B. Okunoye, 'Digital identity in Nigeria: case study conducted as part of a ten-country exploration of socio-digital ID systems in parts of Africa (Towards the Evaluation of Digital ID Ecosystems in Africa: Findings from Ten Countries)' (2021), available at: https://researchictafrica.net/publication/digital-identity-in-nigeria-case-study-conducted-as-part-of-a-ten-country-exploration-of-socio-digital-id-systems-in-parts-of-africa/

309 NIMC (n.d.), available at: https://nimc.gov.ng/policies/ World Bank, 'Policies', available at: https://nimc.gov.ng/policies/

310 World Bank, 'ID4D Country Diagnostic: Nigeria' (Washington, DC: The World Bank, 2016).

311 UNICEF, 'A Snapshot of Civil Registration in Sub-Saharan Africa' (2017), available at: https://data.unicef.org/resources/snapshot-civil-registration-sub-saharan-africa/

312 UNICEF, 'Only 43 per cent of Nigerian Children's Births Registered' (2021), available at: https://www.unicef.org/nigeria/press-releases/only-43-cent-nigerian-childrens-births-registered-unicef

313 T. Daka, 'NIN registration crucial to combating insecurity, says Buhari' (2021), available at: https://guardian.ng/news/nin-registration-crucial-to-combating-insecurity-says-buhari/

extreme form as the deadly insurgency in the country's north-east border with Cameroon, Niger Republic and Chad. It was thought that the nation's porous borders contributed to the insecurity and efficient identification was needed as a tool in a suite of measures to arrest the situation. The identity project was also positioned as the 'single source of truth'[314] about Nigerians and legal residents in the country, and it is envisaged that national identity data will be harmonised with all other identity databases in the country, such as immigration, SIM cards, driver's licences and financial records.[315]

The NIN is the central feature of the national identity scheme. It is an 11-digit number given upon recording of citizens' data, including personal information (for example, current and previous name(s), date of birth, place of birth and place of origin, marital status, education level, telephone number), residence information (address of residence, town of residence, country of residence, etc.), physical features (gender, tribal marks, hair colour, height, etc.), supporting documentation (national passport, insurance number, driver's licence, etc.) and biometric information (face and fingerprints) in at least 7000 enrolment centres across the country, and in 40 countries across the world. These enrolment data are captured in an enrolment form.[316] Enrolment for the digital identity commenced in February 2012,[317] and became compulsory in January 2019[318] as a prerequisite to obtaining public and private services such as international passports, driver's licences and pension plans. Enrolment for the national identity reached 77.1 million in March 2022.

The national identity database and supporting infrastructure is listed among Nigeria's CII in the nation's revised Cybersecurity Policy,[319] alongside national assets in the following sectors identified as CNI sectors in the Cybersecurity Policy: power and energy, water, information, communications, science and technology, banking/finance and Insurance, health, public administration, education, defence and security, transport, food and agriculture, safety and emergency

314 K. Breckenridge, 'The failure of the "single source of truth about Kenyans": the NDRS, collateral mysteries and the Safaricom monopoly', *African Studies* 78 (1) (2019), 91–111.

315 NIMC (n.d.), available at: https://nimc.gov.ng/policies/

316 NIMC, 'Enrolment form' (2021), available at: https://nimc.gov.ng/enrolment-form/

317 World Bank (see note 4 above).

318 O. Awojulugbe, 'FG: Mandatory use of national ID number begins in January 2019' (2018), available at: https://www.thecable.ng/fg-mandatory-use-of-national-id-number-begins-in-january-2019

319 Office of the National Security Adviser, 'National Cybersecurity Policy and Strategy' (2021), available at: http://ctc.gov.ng/wp-content/uploads/2021/02/NATIONAL-CYBERSECURITY-POLICY-AND-STRATEGY-2021_E-COPY_24223825.pdf

services, industrial and manufacturing, and mines and steel. The process of formalising CNI identification in Nigeria followed some of the most important principles, including having a strong mandate from national leadership, technical and policy competence, having clear and transparent policy development processes, leveraging existing laws and organizations, and building public–private relationships to facilitate critical infrastructure identification.[320]

A strong mandate from national leadership in identifying the national identity database as CII is reflected in the office of the National Security Adviser, domiciled in the presidency of the country. Technical competence is reflected for instance in the constellation of agencies that manage Nigeria's identity ecosystem and exemplified by NIMC; a transparent policy development process was demonstrated in the inclusive process that brought about the nation's Cybersecurity Policy, the strong policy and legal landscape governing cyberspace, and the multi-stakeholder approach which leveraged inputs from the private sector and civil society in a public-sector-led initiative.

Critical national infrastructure protection and incidence response in Nigeria's Cybersecurity Policy are laid out by the Office of the National Security Adviser.[321] Its policy recommendations are robust in detailing the governance structure and operating procedures that cater for adequate protection of the nation's CNI and CII. Nigeria's cybersecurity strategy and policy also identify eight pillars of national cybersecurity strategy, which include:

1. strengthening cybersecurity governance and coordination
2. fostering protection of critical national information infrastructure
3. enhancing cybersecurity incident management.

To achieve these objectives, the policy is explicit in many sections and paragraphs on the need for intersectoral cooperation and collaboration between the many stakeholders that constitute the cybersecurity ecosystem. These include government agencies and private sector operators that are in many instances the owners of CNI and CII. For instance, in its foreword the policy states:

320 Global Forum on Cyber Expertise, 'Towards Identifying Critical National Infrastructures in the National Cybersecurity Strategy Process' (2021), available at: https://cybilportal.org/publications/towards-identifying-critical-national-infrastructures-in-the-national-cybersecurity-strategy-process/

321 Office of the National Security Adviser (see note 12 above), Chapters 4 and 5 respectively.

> Our approach to national cybersecurity is the development of a robust and adaptive digital ecosystem based on mutual collaboration and synergy of a triad of government, academia, and industry, reinforced by strong regional and international alliances.

The Cybersecurity Policy provides guidelines on how this intersectoral co-operation might work in practice. For example, it mandates the creation of a 'Trusted Information Sharing Network' (TISN) where the owners and operators of critical infrastructure in the public and private sectors collaborate and share information on threats and vulnerabilities, as well as developing strategies and solutions to mitigate risks to infrastructure.[322]

In Chapter 4, 'Fostering protection of Critical National Information Infrastructure', it states:

> This strategy acknowledges that protection of CNII is a shared responsibility across government, private sector (owners and operators of CNI infrastructure) and the entire populace.

It also places a premium on identifying and managing cross-sectoral dependencies in a sub-heading in Chapter 4 which extensively treats this topic and responsibility in cyber norms and obligations. Empowered by section 41 of the Cybercrime Act,[323] the Nigerian Computer Emergency Response Team (ngCERT) is the focal point of national incident management domiciled under the National Cybersecurity Coordination Centre (NCCCC) and coordinates activities of sectoral incident response teams (CSIRTs) and other CSIRTs in the private sector.

In February 2022, the server hosting NIMC's verification service experienced a downtime of over 10 days. This incident was not due to a cybersecurity breach but is significant because it restricted access to the many services for which the NIN was made a prerequisite, such as SIM card registration, opening of bank accounts and obtaining international passports and driver's licences. In Chapter 4, the Cybersecurity Policy includes CNI protection to include incidents of unintentional disruption such as this incident with the NIMC servers.[324]

322 Office of the National Security Adviser (see note 12 above).

323 Computer Emergency and Response Team, Cybercrime Act 2015, available at: https://www.cert.gov.ng/ngcert/resources/CyberCrime__Prohibition_Prevention_etc__Act__2015.pdf

324 Office of the National Security Adviser (see note 12 above).

This chapter seeks to understand the nature of incident response in the cybersecurity practice around Nigeria's digital identity project (a CII), and its relationship with existing cybersecurity norms/obligations, particularly how these relate to maturity of the nation's cybersecurity capacity. The chapter contributes to literature working towards understanding public–private partnerships in CII protection.[325] This includes understanding differences in objectives, operating models and incentive structures, which are important hurdles but are underestimated and not well considered when CII protection is evaluated.

This work draws from semi-structured interviews with five cybersecurity experts who are knowledgeable about the case of the NIN server downtime to illustrate how implementation gaps in cyber obligations enshrined in the national Cybersecurity Policy relating to intersectoral collaboration might have contributed to the lengthy downtime of the NIN, a CNI/CII for which there were no effective alternatives. The interviewees were selected by snowball sampling from the cybersecurity community.

The cyber incident

In January 2019, the NIN was made a prerequisite for Nigerians and legal residents to access services such as SIM card issuance and retrieval of lost SIM cards, opening of bank accounts, and issuance of international passports and driver's licences.[326] An integral part of this process is verification of the NINs of applicants for these services by service providers. However, for more than 10 days in February 2022, the NIN verification platform hosted by the government's ICT service provider, Galaxy Backbone, experienced a downtime, leaving citizens unable to access the critical services listed above. Service providers first had to have their NINs verified by the NIMC: a function the technical glitch did not permit. A workaround using the virtual NIN (vNIN) also presented difficulties, which caused delays and frustrated both service providers and customers, and resulted

325 Global Forum on Cyber Expertise (see note 13 above).

326 Awojulugbe (see note 11 above).

in economic consequences for telecommunications service providers and financial institutions that lost revenue due to business interruptions.[327]

Galaxy Backbone was set up by the federal government of Nigeria to maintain a nationwide IP-based network providing connectivity and other information technology services for all ministries, departments and agencies (MDAs) of the government. Its functions, obtained from its website, include:[328]

1. providing shared ICT infrastructure, applications and services to all MDAs and institutions of the Federal government;
2. building and operating a single nationwide IP broadband network to provide network services to all Federal Government MDAs and institutions;
3. deploying and maintaining all national database management systems and transversal applications in government. These include government-wide messaging and collaboration, federal public service personnel and payroll system, government gateway and national portal;
4. setting standards and guidelines for the support of government MDAs in the acquisition and acceptable usage of ICT infrastructure, applications and services across different agencies and government institutions;
5. providing wide area networks (WANs) and metropolitan area networks (MANs) to connect all government entities;
6. providing technical support to the Ministry of Communication Technology for end-to-end quality assurance of ICT projects and capacity-building for ICT professionals in government.

A downtime in the Galaxy Backbone-maintained server that supports the NIN verification service persisted for over 10 days in February 2022, an unusually long time for a single, no-alternative, mandatory identity provider and CII. Besides the difficulty numerous individuals seeking services were confronted with in engagements with service providers—including immigration, who issued passports; the Federal Road Safety authorities, who issue driver's licences;

327 B. Okunoye, '#GoodID lessons: Why Nigeria needs more than the NIN' (2022), available at: https://www.africaportal.org/features/goodid-lessons-why-nigeria-needs-more-nin/; Daily Trust, 'Collapse of NIMC server' (2022), available at: https://dailytrust.com/collapse-of-nimc-server

328 Galaxy Backbone, 'Our Mandate' (2019), available at: https://galaxybackbone.com.ng/mandate/

financial service providers who could not open new accounts; and telecommunications service providers who could not issue new SIM cards or retrieve lost SIMs—the lengthy downtime resulted in economic losses, particularly for the telecommunications sector.[329] Feedback from interviews with cybersecurity experts in Nigeria conversant with this incident suggested that the lengthy downtime had connections with gaps in intersectoral cooperation in responding to it.

Although the server downtime affected many critical sectors of the economy, some of which were mandated to maintain sectoral CERTs by the cybersecurity strategy and policy, many were not set up and operational at the time.[330] Those that were set up were reluctant to intervene, even though the cybersecurity strategy[331] mandated intersectoral cooperation, for example by the sharing of information by agencies. Despite their sectors and businesses being affected by the downtime, they were restrained by the unspoken relational protocols that discouraged units from intervening in units not in their MDAs. Not only does this run against the best cyber norms, it also runs against best international practice in inter-agency and intersectoral cooperation.[332]

329 B. Okunoye, 'Digital identity for development should keep pace with national cybersecurity capacity: Nigeria in focus', *Journal of Cyber Policy* 7 (1) (2022), 24–37.

330 Research by the author suggested that some financial institutions and the Nigerian Communications Commission (NCC) were among the organisations that had functional CERTs at the time. As explained in the introduction, the sectors identified as CNI sectors in the Cybersecurity Policy (power and energy, water, information, communications, science and technology, banking/finance and Insurance, health, public administration, education, defence and security, transport, food and agriculture, safety and emergency services, industrial and manufacturing, and mines and steel) are also expected to have sectoral CERTs or CSIRTs. It was not clear whether certain sectoral CERTs were in operation.

331 Office of the National Security Adviser (see note 12 above).

332 Cybsersecurity and Infrastructure Security Agency (2022), available at: https://twitter.com/CISAgov/status/1501972494357606404

Cyber norms and cyberculture as key components of national cyber capacity/maturity

A national government's structure for dealing with CNI/CII should be clearly identified in a declarative policy, ideally within a national cybersecurity strategy (NCS) that is developed through a multi-stakeholder/consensus process.[333] This structure, when it is identified in a policy such as a national cybersecurity strategy and an Act of Parliament (a legal mandate) like the Cybercrime Act 2015—both applicable to Nigeria—can be held to be stronger than cyber norms, which are not always adhered to, and instead viewed as binding obligations given the backing of law that the document has. Whereas norms are political agreements that do not impose legal obligations on parties, the same cannot be said of obligations encoded in national policies and legislation.[334]

Intersectoral cooperation and collaboration in CNI/CII is a critical cyber obligation that has important consequences for the security of CNI/CII. The lengthy downtime of Nigeria's NIN verification platform and the CII incident response described above suggests strengthening of the aspect 'incident response' under the dimension 'cybersecurity policy and strategy' in the Cybersecurity Maturity Model for Nigeria.[335]

In 2018, the Global Cyber Security Capacity Centre (GCSCC) reviewed the maturity of cybersecurity capacity in Nigeria at the request of the government.[336] The GCSCC is an international centre of excellence on cybersecurity research at the Department of Computer Science, University of Oxford. A major research focus of the GCSCC is on cybersecurity capacity-building across the world. In furtherance of its aims to assess the maturity of cybersecurity, the GCSCC developed the GCSCC Cybersecurity Maturity Model (CMM) model, an assessment tool, to review cybersecurity capacity of nations and help them 'to self-assess,

333 Global Forum on Cyber Expertise (see note 13 above).

334 Australian Strategic Policy Institute, 'The UN norms of responsible state behaviour in cyberspace: Guidance on implementation for Member States of ASEAN', Australian Strategic Policy Institute: International Cyber Policy Centre (2022), available at: https://www.aspi.org.au/report/un-norms-responsible-state-behaviour-cyberspace

335 Global Cyber Security Capacity Centre, 'Cybersecurity Capacity Review: Nigeria 2019', available on request from the Global Cyber Security Capacity Centre, Oxford.

336 Ibid.

benchmark, better plan investments and national cybersecurity strategies, as well as set priorities for capacity development'.[337] The CMM assesses the maturity of nations across five dimensions of cybersecurity capacity: cybersecurity policy and strategy; cyber culture and society; cybersecurity education, training and skills; legal and regulatory frameworks; and standards, organisations and technologies. Since inception, the CMM has been deployed more than 120 times in over 87 countries.

CMM dimensions have factors that describe the cybersecurity capacity of nations. Each factor consists of aspects, and for each aspect there are indicators, which describe steps and actions that, once implemented, define the state of maturity of that aspect.[338] The CMM describes five stages of maturity listed by the Global Cyber Security and Capacity Centre: start-up, formative, established, strategic and dynamic, with start-up being the most basic and dynamic the most advanced.[339]

Across five dimensions, Nigeria's CMM report suggested a maturity stage above 'formative' for only three factors: 'trust and confidence on the internet', 'framework for education' and 'formal and informal cooperation frameworks to combat cybercrime'. All the other 21 factors across five dimensions of maturity were adjudged to be at either start-up or formative stage.

337 GCSCC, 'Cybersecurity Capacity Maturity Model for Nations (CMM)', 2021 edition, available at: https://gcscc.ox.ac.uk/files/cmm2021editiondocpdf

338 Ibid.

339 Ibid.

TABLE 1 | CMM dimensions and factors that measure a nation's cybersecurity capacity and maturity.

Dimensions				
Cybersecurity policy and strategy	Cyber culture and society	Cybersecurity education, training and skills	Legal and regulatory frameworks	Standards, organisations and technologies
Factors				
National cybersecurity strategy	Cybersecurity mindset	Awareness-raising	Formal and informal cooperation frameworks to combat cybercrime	Adherence to standards
Incident response	Trust and confidence on the internet	Framework for education	Criminal justice system	Internet infrastructure resilience
Critical infrastructure protection	User understanding of personal information protection online	Framework for professional raining	Legal frameworks	Software quality
Crisis management	Reporting mechanisms			Technical security controls
Cyber defence consideration	Media and social media			Cryptographic controls
Communications redundancy				Cybersecurity marketplace
				Responsible disclosure

Central to this chapter are the factors 'incident response' and 'critical infrastructure protection', under the dimension 'cybersecurity policy and strategy'. Both are adjudged 'formative', suggesting a maturity stage where 'some features of the Aspect have begun to grow and be formulated, but may be ad hoc, disorganised, poorly defined or simply new. However, evidence of this activity can be clearly demonstrated'.[340]

Nigeria's CMM report describes the factor 'incident response' as addressing the capacity of the government to 'identify and determine characteristics of

340 GCSCC, 2019 (see note 28 above).

national-level incidents in a systematic way. It also reviews the government's capacity to organise, coordinate and operationalise incident response.'[341] The CMM assessment in Nigeria on the factor 'incident response' captures some of the feedback from expert interviews and demonstrated in the NIN verification server downtime scenario:

> On the other hand, the delineation between the roles and responsibilities of the ngCERT and critical national infrastructure operators, for example, in an event of an incident affecting the latter is much less clear; the representatives of the critical national infrastructure operators are unaware of the protocols for escalating an incident to ngCERT, with whom and how to establish appropriate communication in an event of cyber incident, etc. Certainly, no formally documented processes and protocols for cyber incident response exist. This is not to say that collaboration between various incident handlers does not exist.[342]

Analysis and conclusion

This chapter examines the incident of the lengthy downtime of Nigeria's national identity server, a critical information infrastructure, in February 2022, and how the incident response highlighted the role of cyber norms and obligations in a specific cyber context. It describes how although specific obligations regarding intersectoral cooperation are encoded in the national Cybersecurity Policy and Cybercrime Act 2015, they were not specifically followed in the incident response. Feedback from expert interviews suggests that this could be linked to relational protocols that preclude agencies from intervening in other agencies' operational situations. A previous research report by the author highlighted similar challenges with the NIMC app and the e-Naira app, both of which had lengthy times before stable versions could be operational, in a national cybersecurity context where sectoral partners (e.g. financial institutions) had launched similar

341 Ibid.

342 Ibid., p. 27.

apps with great technical efficiency.[343] Interview data from cybersecurity experts contacted for this research suggest that public cyber projects do not sufficiently take inputs from experts outside their circles.

This chapter also links this cyber context (NIN) with the maturity of Nigeria's cybersecurity capacity as evaluated in its 2018 cybersecurity capacity report by the GCSCC.[344] Policy recommendations arising from this might include the following:

1. *Strengthening of incidence response protocols as mandated in the nation's Cybersecurity Policy and the GCSCC report for Nigeria:* In particular, there seems to be a need for top-down guidance from leadership that reorients cultures within some organisations in the public and private sectors towards conceiving cooperation and receiving assistance as strength-building and not as threatening or adversarial. Strong policy leadership is implied here, to ensure that the recommendations in the nation's Cybersecurity Policy are adhered to in practice in future cybersecurity incident handling.
2. *Instituting legal recourse mechanisms and/or penalties against organisations involved in the sub-par handling of cybersecurity incidents:* In this incident, the downtime in the NIN verification server cost individuals and businesses time and money, yet there are no avenues for recourse in the nation's Cybercrime Act for affected individuals or organisations. Provisions in the relevant legislation and policy such as the Cybercrime Act,[345] Nigerian Data Protection Regulation[346] and Draft Data Protection Bill[347] tend to address issues of cybercrime or data breaches, but do not sufficiently address scenarios of cybersecurity incident handling by instituting mechanisms whereby affected parties might gain relief. Nigeria's Cybercrime Act's planned amendment should include such mechanisms. It must be said, however, that much progress has been made in the legal and regulatory environment around national identity in

343 Okinuye, 2022 (see note 22 above).

344 GCSCC, 2019 (see note 28 above).

345 CERT, 2015 (see note 16 above).

346 NDPB, 'Nigeria Data Protection Regulation 2019', available at: https://ndpb.gov.ng/Files/NigeriaDataProtectionRegulation.pdf

347 NCC, 'Data Protection Bill 2020', available at: https://www.ncc.gov.ng/documents/911-data-protection-bill-draft-2020/file

Nigeria, which has seen landmark legal cases brought against NIMC by civil society for two cases of privacy violations.[348]

3. *Provision of alternatives/redundancy mechanisms for CNI:* This cybersecurity incident was worsened because there were no alternatives to the NIN, which in January 2019 had been made the sole identification required to access numerous public and private services.[349] A recommendation from this chapter might be that instead of a single, mandatory identification as a prerequisite for accessing services, a system of trusted identification providers be approved for accessing services. This would ensure some degree of redundancy within the system.
4. *Closer public sector synergy with the private sector in cybersecurity*: Responses from cybersecurity experts interviewed for this research suggested that private sector participation in public technical projects could be strengthened. Both public and private sector professionals bring unique skills and experiences that enhance public sector projects. In line with international best practices, cybersecurity policy leaders in Nigeria can benefit from closer cooperation with the private sector in the planning and execution of projects.

348 Andersen Tax, 'Federal High Court Affirms the Data Privacy Rights of Nigerian Citizens' (2019), available at: https://ng.andersen.com/federal-high-court-affirms-the-data-privacy-rights-of-nigerian-citizens/; H. Abiola, 'National Digital Identity Card: NGO Seeks Injunction against NIMC for Data Breach and Omission to Conduct Data Protection Impact Assessment' (2020), available at: http://loyalnigerianlawyer.com/national-digital-identity-card-ngo-seeks-injunction-against-nimc-for-data-breach-and-omission-to-conduct-data-protection-impact-assessment/

349 Awojulugbe, 2018 (see note 11 above).

CHAPTER 8

The role of state–civil society relations in shaping cyber norms in South Korea

SOFIYA SAYANKINA

Introduction

Although cyberspace started simply as a tool for faster and more convenient communication and a platform for opinion-sharing, free speech and self-expression, its technological evolution and the rapid growth of internet users have changed it into a place of competitiveness and opportunism. The internet has opened up new ways for individuals and civil society groups to engage in the deliberation process through digital public spheres, and even to instigate change in power structures. Such developments led to the discussion of the applicability of norms to civil society's behaviour in cyberspace.

The concept of the public sphere was famously introduced in 1989 by Habermas, who laid out its three characteristics: rational–critical argument as the only measurement of contribution judgement, topics relating to the issue

'common concern',[350] and 'openness to all members of the public'.[351] This concept was developed by Dahlberg, who established six attributes of a public sphere discussion: reflexivity, ideal role taking, sincerity, formal inclusion, discursive equality and autonomy.[352] Habermas later identified market-based motivations and low level of awareness among members of the public as threats to this deliberative democracy model.[353]

While there are fair criticisms that address poor quality of debates in the digital public spheres, potential underrepresentation of some demographics due to lack of access to the internet and a tendency of the participants to gravitate to like-minded discussion spaces that resemble echo chambers, optimistic attitudes attribute the diminished power of mass media to the rise of online public spaces.[354] According to Benkler, the simplicity of online communication can transform the passive members of the public into active participants in the cyberdemocracy process.[355] Finally, cyberspace can provide a better opportunity for constituents to control their representatives.[356]

The cyberdemocracy-building process is facilitated by debates in the online public spheres that continue to challenge state power structures in a variety of ways. The question posed in this research is how norms in South Korea's segment of cyberspace are shaped through the interaction of the state with the digital public sphere. South Korea makes a good case study as one of the most tech-savvy societies in the world, with developed internet infrastructure and high level of internet penetration. According to the recent data, 97% of the country's population aged 3 years or older regularly use the internet,[357] while the area

350 Jürgen Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society* (Cambridge: Polity, 1989).

351 Deen G. Freelon, 'Analyzing online political discussion using three models of democratic communication', *New Media & Society* 12 (7) (2010), 1172–1190.

352 Lincoln Dahlberg, 'The Habermasian public sphere: taking difference seriously?', *Theory and Society* 34 (2) (2005): 111–136.

353 Jürgen Habermas, 'Political communication in media society: does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research', *Communication Theory* 16 (4) (2006), 411–426.

354 Kyle Lorenzano, Miles Sari, Colin Storm, Samuel Rhodes and Porismita Borah, 'Challenges for an SNS-based public sphere in 2016', *Online Information Review* 42 (7) (2018), 1106–1123.

355 Ibid.

356 Esther Dyson, George Gilder, George Keyworth and Alvin Toffler, 'Cyberspace and the American Dream: a Magna Carta for the knowledge age', *Future Insight* (August 1994), available at: http://www.pff.org/issues-pubs/futureinsights/fi1.2magnacarta.html

357 World Bank, 'Individuals Using the Internet (% of Population) – Korea, Rep. Data' (2020), available at: https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=KR

coverage rate is second in Asia after Brunei, with 96.37%.[358] Internet coverage currently reaches 99.2% of households in South Korea.[359] 87% of the population are active social media users, and the number of mobile connections in the country is equivalent to 118% of the population.[360] Korea also boasts the fastest mobile speed in the world at 113.01 Mbps.[361]

In order to answer the research question, the state's role in building an online democratic sphere in South Korea will first be analysed. Next, this chapter will review online political activity in South Korea and three ways it engages with the state power structures in cyberspace. Lastly, conclusions will be drawn on how the state's policies contributed to the norm-building process in South Korea's digital public sphere.

The establishment of the online democratic sphere in South Korea

To fully understand the dynamics of the norm-building process in South Korea's segment of cyberspace, it is necessary to briefly look at how the development of Korea's civil society has coincided with the government's intention to establish its place in the world. Although South Korea's legacy of economic guidance, especially in the infrastructure sector, has always been strong, it is important to underline here that the government's initial decision to promote cyber technology and involve different government agencies in procurement of the ICT sector was done as part of its globalisation strategy.[362] After the country democratised in 1987, the authorities were looking for ways to incorporate the globalisation trend into South Korea's political, economic and social agendas, which resulted in the proclamation of *segyehwa* (globalisation). Even in the aftermath of the

358 Nina Jobst, 'Topic: Internet usage in South Korea', *Statista* (11 August 2021), available at: https://www.statista.com/topics/2230/internet-usage-in-south-korea/#dossierKeyfigures

359 Elaine Ramirez, 'Nearly 100% of households in South Korea now have internet access, thanks to seniors', *Forbes* (31 January 2017), available at: https://www.forbes.com/sites/elaineramirez/2017/01/31/nearly-100-of-households-in-south-korea-now-have-internet-access-thanks-to-seniors/?sh=793feb8d5572

360 Jobst (see note 9 above)

361 Ibid.

362 Benjamin Gosnell Bartlett, 'Institutional determinants of cyber security promotion policies: lessons from Japan, the U.S., and South Korea' (PhD dissertation, UC Berkeley, 2018); Samuel S. Kim (ed.), *Korea's Globalization* (Cambridge: Cambridge University Press, 2000), 1–29.

Asian Financial Crisis, when the IMF imposed structural reforms as part of the 1997 financial crisis agreement that liberalised the financial sector in the country and advised the government to slash expenditure, South Korea's government continued with its control over economic development, in part balancing against business interests by coordinating with labour.[363]

Already experiencing growing tensions with the *chaebol* companies (traditionally powerful industrial conglomerates), the government was committed to preventing the commercialisation of the internet, which happened to portions of the web in 1995 in the US. The goal in Korea was to preserve the domestic segment of the internet for 'grassroots political and social purposes', as an indispensable part of further democratisation of the country, which is why the government relied heavily on civil society. Civil society at the time also opposed any attempt to introduce a telecommunication policy aimed towards privatisation, and insisted on keeping the internet 'for non-market use of the electronic communications technology'.[364]

Furthermore, Korea did not invest only in tangible assets, but also in public awareness and human resources. Through the government's support of public cyber-literacy, a number of training programmes were implemented to provide internet expertise to different groups of people. Bartlett singled out one such programme that was named 'Ten Million People Internet Education': a project to provide computer and internet skills to 10 million people by 2002.[365] A target population for this programme was housewives: the reports indicate that around 1 million of them were provided with courses among the total 4.1 million people who participated in government-initiated programmes.[366] Another successful initiative was providing primary and secondary schools with high-speed internet access. Internet cafes with high-speed access, known in Korea as PC-bangs ('personal computer rooms'), spread widely, providing inexpensive access to the internet and quickly becoming the youth's favorite pastime. Such resource allocation by the administration contributed to public awareness of cyber issues and a high level of cyber-hygiene as well as improved computer skills of

363 Dong-Myeon Shin, *Social and Economic Policies in Korea: Ideas, Networks, and Linkages* (London: RoutledgeCurzon, 2003), 80.

364 Myong Koo Kang, 'The grassroots online movement and changes in Korean civil society', *Review of Media, Information and Society* (3) (1998), 107–127.

365 Bartlett (see note 13 above).

366 Ibid.

the general population, which in turn strengthened the government's ability to combat cybercrime.

Thus, the high number of internet users in South Korea is a direct result of government initiatives that prompted a fast spread of the internet country-wide: while in 1998 only about 6% of the population was actively using the internet in Korea, by 2001 this figure jumped to 56%.[367] In the 18 months from May 1999 to November 2000, the number of registered domain names in South Korea grew from 100,000 to more than 500,000.[368] As Ahn et al. note, it was the amount of emerging content in Korea's domestic cyber segment that drew so many new users in.[369]

Here it is necessary to point out that, although there are no information flow restrictions, Korean national cyberspace is somewhat isolated from global cyberspace, for two reasons. First is the use of the Korean language: the Korean cyber domain is fully available only to Korean language speakers, which includes the majority of residents in South Korea and some members of the Korean diaspora. Since the majority of North Korean citizens do not have access to the global internet, only a tiny percentage can access the Korean-language segment of the internet. The language barrier here carries the function of sovereign state borders, which lets the Korean-language segment develop authentically without significant information campaigns or fake news operations from abroad. The second reason is the strict registration rules of online discussion forums: in order to participate, users need to confirm their real identity through the Korean phone number verification process, which effectively limits the potential pool of discussants to South Korea's residents.

Korean civil society has been gaining prominence since the country's democratisation in 1987, and the spread of advanced cyber technology has largely facilitated its development by providing accessible and safe platforms for discussion and engagement in political discourse. Korea's digital society uses the digital public sphere to exercise its power by either opposing the government and its policies or supplementing the state's responsibilities with its own initiatives.

It is important to mention again here, that, unlike the US, where the government's pursuit of economic profit encouraged companies to quickly privatise and

367 World Bank (see note 8 above).

368 Korea Network Information Center (KRNIC), 'IPv4', KRNIC (KISA, January 2022), available at: https://xn--3e0bx5euxnjje69i70af08bea817g.xn--3e0b707e/jsp/eng/ipas/statistics/ipV4.jsp

369 Joong-ho Ahn et al., *History of Internet Companies in Korea: Experimentation of the Network Economy* 한국 인터넷 기업의 변천사 : 네트워크 경제의 실험과 형성 (Seoul: Seoul National University, 2006), 36.

commercialise the internet in the 1990s, the internet in Korea has always been developed with public purposes in mind. It was chosen as a means to support democratisation in South Korea, and since as early as 1995 digital society has been actively opposing the push towards privatisation and promoting not-for-profit use of electronic communications technology.[370]

These days much political discussion in South Korea happens on social networks such as Twitter and Facebook, comments sections of the news-generating websites, and politically oriented podcasts and YouTube channels. The popularity of online news can in part be attributed to the trend that emerged during the presidencies of Lee Myung-bak and Park Geun-hye, when media freedom suffered growing restrictions: critical journalists quit the mainstream services and started individual reporting via social media, with the majority of the public following them.[371] As social media platforms provide a much-needed space for direct involvement and discussions, the individual reporters also managed to create content that was a lot more entertaining and engaging than that of the mainstream sources; this also contributed to their popularity.[372] This development indicates that a core criticism—dependence of the online democratic sphere on traditional media for spreading the political message[373]—does not apply in South Korea, where more than 90% of news consumption comes from online sources: an extremely high figure compared to an average of 50% in other countries.[374]

One more example of digital society in South Korea defying the state is presented in the work of Kwon and Rao on cyber-rumour sharing.[375] Contrary to the expected decrease in cyber-rumour volume under tightened surveillance, the research has shown that the concerns of the cyber public sphere of South Korea about the state's surveillance in cyberspace has 'increased their willingness to engage in cyber-rumor sharing, and this tendency was particularly strong

370 Ronda Hauben, 'The rise of netizen democracy: a case study of netizens' impact on democracy in South Korea' (Columbia University, 2005), available at: http://www.columbia.edu/~rh120/other/misc/korean-democracy.txt

371 Hyejin Kim, 'Online activism and South Korea's Candlelight Movement', in *Dog Days: Made in China Yearbook 2018* (Canberra: ANU Press, 2018), 224–227.

372 Ibid.

373 Andrew Chadwick, *The Hybrid Media System: Politics and Power* (New York: Oxford University Press, 2013); Kevin M. Carragee and Wim Roefs, 'The neglect of power in recent framing research', *Journal of Communication* 54 (2) (2004), 214–233.

374 Milos Djordjevic, '25+ remarkable news consumption statistics [2021 edition]' (17 February 2021), available at: https://letter.ly/news-consumption-statistics/

375 K. Hazel Kwon and H. Raghav Rao, 'Cyber-rumor sharing under a homeland security threat in the context of government internet surveillance: the case of South–North Korea conflict', *Government Information Quarterly* 34 (2) (2017), 307–316.

when the homeland security was on alert'.[376] The spread of government-related rumours in cyberspace despite high chances of being subject to state surveillance also points to the capability of the online democratic sphere to withstand attempts at government control.

While the high level of digitalisation in South Korea is indisputable, the literature has underlined that the level of political apathy was also high prior to the 2010s: the turnout for the 2007 presidential election was 63%, but the three subsequent elections have recorded an over-75% rate.[377] The turnout for parliamentary elections was even lower: only 46% in 2008, the lowest among OECD countries, with a big increase to 66.2% in 2020.[378] While the data clearly show the increase in the level of political engagement from the public, it is necessary to look further into how the digital public sphere could have facilitated that change.

The individual and the state in South Korea's cyberspace

Internet activism in South Korea, mainly driven by individuals, has its roots in 2002, when an internet user named AngMa (앙마) proposed a candlelight vigil for two girls who were killed by a US military vehicle in the Yangju highway incident. AngMa's post attracted a lot of attention online and was useful in mobilising ordinary people (especially Korean youth) to demand SOFA (Status of Forces Agreement) reform and an apology from President George W. Bush through peaceful candlelight protests in Gwanghwamun Square.[379] Although the 2002 candlelight protests were still mainly coordinated by the civil society groups, it gave them a chance to directly propose a political agenda and get involved in renegotiations.[380] On the other hand, the 2008 anti-US beef protests were facilitated almost exclusively by the citizens themselves via cyber platforms, with no

376 Ibid.

377 International IDEA, 'Voter Turnout by Election Type: Korea, Republic of', International Institute for Democracy and Electoral Assistance (2022), available at: https://www.idea.int/data-tools/country-view/163/40

378 Ibid.

379 Hee-Yeon Cho, Lawrence Surendra and Hyo-Je Cho, *Contemporary South Korean Society : A Critical Perspective* (London: Routledge, 2015), 145.

380 Whasun Jho, 'The transformation of cyberactivism and democratic governance in Korea: the role of technology, civil society, and institutions', *Korea Observer* 40 (2) (2009), 337–368.

guidance from civil organisations, which was uncharacteristic of a traditional protest or a rally.[381] The result of the 2008 protests was the replacement of seven out of eight senior presidential secretaries.

Since then, the internet has grown even more and has become the primary space for active discussion of political issues and activism in South Korea, culminating in 2016–2017 when a few million people self-mobilised exclusively via online sources for the country-wide candlelight protests that ultimately resulted in the impeachment of the then president Park Geun-hye. As the mainstream media in South Korea initially did not report on the protest movement, it was the news and videos from ordinary citizens posted on social media that mobilised digital society. The decision of the National Assembly to impeach South Korea's president and the constitutional court's subsequent dismissal of the president through legal deliberation can be considered a direct result of processes facilitated by the digital public spheres and a part of the mature constitutional democracy.[382]

In the aftermath of the 2016–2017 protests, South Korean citizens were given a more or less official way to state their opinion directly to the president's administration and the government via cyber means. The government established an official online outlet on the Blue House website where citizens could file petitions. This attempt to coordinate the activities of the online public sphere had simple mechanics: the petition needed to gather more than 200,000 signatures within 30 days in order for a relevant high-level official to issue a response within the next 30 days, expressing the government's view and providing some possible solutions.[383] Petitions, divided into 17 categories, ranged from requests for a severe sentence in prosecution cases to requests to enact an anti-discrimination law, while the three most popular sections were political matters, miscellanious and human rights.

The total number of petitions filed reached 1.1 million in February 2022, averaging about 670 per day. Visitors aged 18–24 constituted the largest proportion of people signing the petitions, amounting to 29.3%.[384] During this time,

381 Ibid.

382 Jongcheol Kim, 'The Impeachment of South Korean President Park Geun-Hye: Constitution, Politics and Democracy', EAF Policy Debates (East Asia Foundation, 21 March 2017), available at: http://www.keaf.org/book/EAF_Policy_Debate_The_Impeachment_of_South_Korean_President_Park_Geun-Hye:_Constitution_Politics_and_Democracy

383 Tae-jun Kang, 'What is next for South Korea's official online petition channel?', *The Diplomat* (10 November 2018), available at: https://thediplomat.com/2018/11/what-is-next-for-south-koreas-official-online-petition-channel/

384 Arin Kim, 'Millennials dominate Blue House petition board', *Korea Herald* (7 November 2019), available at: http://www.koreaherald.com/view.php?ud=20191107000693

293 petitions reached the necessary 200,000 signatures and received a response from a member of the administration, usually in video or text format. Originally taking the 'We the People' petition channel of the US White House as its prototype, the Blue House set up a much higher threshold for the number of signatures (200,000 vs 100,000) required in order to produce an official response. Although some expressed concerns that such a channel might decrease the responsibility of the public as it would tend to rely on the power of the state, or that it could become a platform for further polarisation of Korean society, the petition website helped administration focus on citizens' concerns as it increased the sense of involvement on the part of the public, as well as bringing attention to issues that the government would not traditionally be emphasising, such as removal of abandoned dog shelters or taking measures against high levels of fine dust.

One other indicator of the Blue House petition channel playing an important role in streamlining the public's responsibility in cyberspace was the emergence of similar initiatives. The high level of interest that digital society expressed in the option to directly call for government action has contributed to the popularity of petition services offered by other governmental organisations. According to Kang, the number of applications received by the Anti-Corruption and Civil Rights Commission petition channel, which existed even prior to the Blue House petition website, has increased, with a total of 350,000 civil petitions filed through the commission's online service from January to August 2017 (when the Blue House's petition platform launched), subsequently increasing to 440,000 from January to September 2018.[385]

There also are individuals who are able to pursue some of the traditional state functions. In 2020, after online anger over high-profile sexual harrassment cases that were not followed up with a proper prosecution process under the traditional state judicial system reached its peak, the Digital Prison website was launched. The website was supposed to assume a sovereign function of traditional prosecution services by revealing private information of alleged perpetrators (including sex offenders and pedophiles). However, the issue of revealing the identities of perpetrators was deemed controversial, and due to the rise of concerns about false accusations, public access to the Digital Prison was shut down later in 2020. The Korea Communications Standards Commission expressed the view that 'While freedom of expression needs to be fully protected, acts that undermine the legal system should not be allowed. Publishing personal

385 Kang (see note 34 above).

information on Digital Prison could lead to double punishment or falsely accused victims.'[386] The alleged operator was arrested on the charges of privacy law violation in Ho Chi Minh City, Vietnam, and was sentenced to 3.5 years in prison for sharing personal data online.[387] Although the public's discontent with the legal system was extremely high and the Blue House petition calling for revealing the identities of the offenders received the most signatures in the service's history (2,715,626), the case of the Digital Prison indicates where the line for civil society's responsibility in cyberspace is crossed in the view of the state, as such actions by individuals threaten the legitimacy of digital public spheres. However, in a similar case the court acquitted the operators of the 'Bad Fathers' website, which exposed personal information of fathers who avoid paying child support, explaining the decision as 'beneficial to public interest'.

Online civil society in South Korea also creates initiatives in line with the state's purposes, such as public diplomacy and national branding strategy. In his study of the Korean civil society group VANK (Voluntary Agency Network of Korea), which functions exclusively online, Ayhan Kadir highlights the successful efforts of this NGO in improving the image of South Korea abroad.[388] VANK's activity featuring cyber promotion of Korea is two-fold: first, forming the idea of Korea as 'my friend's country' by establishing personal relationships with people from other countries using social media platforms, and second, correcting factual mistakes about South Korea found on the internet.[389] As the primary goal of public diplomacy is achieving foreign policy changes of a state in favour of the other state,[390] for South Korea as a sovereign state it is especially important to be positively viewed by the global public due to its reliance on soft power, in which case it is possible that the online activities of VANK do indeed contribute to improved perceptions of the country abroad.

The three types of activities of the digital society described above represent three ways in which the public self-mobilises through the digital online sphere in South Korea. First, by filing and signing petitions through the officially provided channel, digital society builds its presence in a way that requires feedback

386 Yun-hwan Chae, 'S. Korea blocks access to "digital prison"', Yonhap News Agency (24 September 2020), available at: https://en.yna.co.kr/view/AEN20200924012300320

387 Ibid.

388 Kadir Ayhan, 'Branding Korea as "My Friend's Country": The case of VANK's cyber public diplomats', *Korea Observer – Institute of Korean Studies* 49 (1) (19 January 2018), 51–81.

389 Ibid.

390 Ibid.

from the state officials.[391] Second, it challenges the power structure through online means (such as alternative prosecution of criminals by revealing their personal information) when official feedback is deemed insufficient. Third, by participating in online debates, civil society contributes to improving the state's image via its own public diplomacy activities. The government in its turn provides civil society with opportunities for action through direct communication and engagement with the power structures. This interaction further facilitates the norm-building process in which both state and civil society are participating.

The state's policies shaping the norm-building process in South Korea's digital public sphere

The South Korean government has played an important role in the development of the country's cyberspace from the late 1980s, which explains the fast advancement of domestic cyber technology. Korea has rather successfully implemented elements of multi-stakeholderism into its normative framework, not only by cooperation between the government, the military and the tech sector in its cybersecurity, but also by safeguarding online public spheres from criminal activity, raising cyber awareness among its population and opening ways to engage digital society in e-governance practices. The government introduces various cyber policies that support a safe and functional online democratic sphere.

Two of the core pillars in Korea's 2019 National Cybersecurity Strategy are increasing participation in establishing universally accepted international rules on cybersecurity and taking the lead in disseminating UN Norms of Responsible State Behaviour in Cyberspace.[392] Similar to its earlier initiatives like *segyehwa* and Ten Million People Internet Education, the Korean government's idea of promoting international norms and best practices usually means applying such norms in the domestic context first. A very high level of overall digital

391 Paul Ferber, Franz Foltz and Rudy Pugliese, 'Cyberdemocracy and online politics: a new model of interactivity', *Bulletin of Science, Technology & Society* 27 (5) (2007), 391–400.

392 Korea Internet and Security Agency, '2019 National Cybersecurity Strategy from the Republic of Korea', KISA (20 June 2019), available at: https://www.kisa.or.kr/EN/302/form?postSeq=4&page=7

development, internet skills of the population and its willingness to engage in democratic deliberation make the government account for the country's digital society in its cyber norm-building process.

Rather than looking at cyber norms as products, this chapter will adopt Finnemore and Hollis' approach, which considers norms as processes through which a preferred behaviour of certain actors is formed.[393] Norm entrepreneurs can construct and support norms through *incentives* (either positive inducements or coercive measures), *persuasion* (asking, arguing or giving reasons for actors to adopt prefered pattern of behaviour) and *socialisation* (incorporating newcomers into existing patterns).[394] All the policies and measures taken by the South Korean government in order to secure a safe cyber environment for its citizens will be divided into these three categories.

First, an extremely fast increase in internet users among South Korean citizens, prompted by the government's strategic investment in cyber technology and network infrastructure, has also become a reason for a greater number of cybercrimes. Despite pre-emptively introducing a few legal mechanisms including amendments to the criminal law back in 1995 (such as Article 347-2 Fraud by the Use of Computer) and establishing a cyber unit of its national police force in 2000 to deal with cyber-mediated crimes, the authorities have failed to prevent the spread in criminal activity in the country's segment of cyberspace; thus, they had to turn to more coercive strategies of regulation involving the user self-identification mechanisms.

The law for self-identification was first enacted in the early 2000s when cybercrime was on the rise due to a significant jump in the number of users and services after the Asian Financial Crisis, which was followed by easing of regulations and requiring self-identification only for services with more than 300,000 daily users.[395] However, the policy was later changed to a self-regulation mechanism applied by the internet service providers due to the new type of cybercrimes resulting in data theft and/or leaks. Through the years, South Korea has retained the real-name authentication process. To register for the majority of popular services or Korean social networks, at least a local phone number is required, or, in some cases, a full identity verification, including name, phone number that

393 Martha Finnemore and Duncan B. Hollis, 'Constructing norms for global cybersecurity', *American Journal of International Law* 110 (3) (2016), 425–479.

394 Ibid.

395 Sung Eun Cho and Sang Hoon Ahn, 'A policy change effected through a change in the meaning of target populations: a case study of South Korea's limited internet user self-identification policy', *Korean Journal of Policy Studies* 30 (3) (2015), 63–89.

corresponds to the name, gender and date of birth. The process is supposed to help combat cybercrime by providing safer spaces for user interaction since the identity information of the offender would be easily available, and at the same time it restricts the use of online public spheres to Korean nationals.

After an infamous 'nth room' case, the Telecommunications Business Act was revised in 2020 and the revision came into effect in December 2021. Under the revision, 'social media, online community operators and big portals – those which generate 1 billion won (850,000 USD) in annual revenue or attract more than 100,000 visitors per day – must check for and filter out illegal videos and photos in their public and group chat rooms'.[396] Earlier in 2020 another law was proposed that would ban teenagers from using messaging and chat room applications that do not require identity verification.[397] The requirement for the platform operators to check for illegal content is also supposed to prevent the spread of cyber-mediated crimes by protecting citizens against participating in illegal group chats, and, although these laws could target domestic applications but could not be enforced on platforms such as Telegram, it did lead to the online debate on the necessity of stricter terms for cyber sex crimes that resulted in the Digital Prison initiative.

The South Korean government quite often uses persuasion to polish the norms of behaviour in online space. The range of existing cyber policies and legal mechanisms is very wide in South Korea, as the government is able to adapt quickly to new developments, partly due to the constant response from the public. According to Chung et al., South Korea's infrastructure was complete in the early 2000s, along with a major portion of the population acquiring access to it, so, regardless of the government in power, there was no major case for opposition by the public on the topic of cybersecurity and, thus, little need for persuasion tactics.[398] The government also views protection of personal information as a priority due to overall increased online activity during the Covid-19 pandemic; hence, it has used persuasion mechanisms to promote decentralised identity data storage with blockchain technology by successfully testing it as a vaccination certificate application.

396 Editorial, 'Dispute over "Nth room" law', *Korea Herald* (14 December 2021), available at: http://www.koreaherald.com/view.php?ud=20211213000597

397 *Korea Times*, 'Gov't to block teenagers from using certain random chat apps' (13 May 2020), available at: https://www.koreatimes.co.kr/www/nation/2021/02/113_289481.html

398 Choong-Sik Chung, Hanbyul Choi and Youngmin Cho, 'Analysis of digital governance transition in South Korea: focusing on the leadership of the president for government innovation', *Journal of Open Innovation: Technology, Market, and Complexity* 8 (1) (2022), 1–28.

Finally, the socialisation initiated by the government is conducted through continuous education of the public on matters of cybersecurity and cyber-hygiene. This process contributes to strengthening overall national security by increasing public awareness. Continuing this trend, among the most recent policies of the Korean government is Digital New Deal 2.0, which is heavily focused on digital education, contactless services and social overhead capital.[399] The policy ensures all levels of cyber-based education from kindergarten to job training centres and nurtures cyber-mediated services in day-to-day lives in the aftermath of the Covid-19 pandemic. In it, the Korean government combines its commitment to continuous development of cyberspace infrastructure with increasing convenience of its use for all people in the country, thus constructing an environment for online debate with high levels of awareness.

To reiterate, the high number of the frequent internet users in South Korea is a direct consequence of late 1990s policies of procurement and expanding the internet infrastructure that made cyberspace available to a large number of people. Civil society in Korea can act against the state by staging online and cyber-mediated campaigns, or can supplement the state's initiatives. The state-enabled online channels that let civil society express its opinion and/or take actions, on the one hand, put pressure on the state to focus on the relevant issue, but, on the other, let the presidential administration officially oversee the predominant narrative and apply necessary regulatory measures if appropriate.

The democratisation process that culminated in the recent history of successful large-scale protests that originated online has created a sense within Korean society that the increased influence it has on government affairs also means increased responsibility in cyberspace. This has resulted in tensions between the state and its citizens when, for example, members of the public tried to establish their own prosecution process of alleged perpetrators by revealing their identities online: an initiative that was largely welcomed in the digital public sphere even though it was a criminal act.

To prevent the members of online spaces from taking initiatives that harm both fellow citizens and the power structure of the state, the South Korean government proceeded with construction of norms for the digital public spheres in the form of incentives, persuasion and socialisation. Incentives consist of identity verification requirements and content monitoring that online platforms and communication providers are obliged to do in some cases. Socialisation includes

399 Ministry of Economy and Finance, 'Government announces Korean New Deal' (14 July 2021), available at: https://english.moef.go.kr/pc/selectTbPressCenterDtl.do?boardCd=N0001&seq=5173

creating official communication channels with the public, which help streamline the opinion of the public into an outlet that lets the government control the response and issue feedback. These policies both contribute to the creation of state-supported norms and support norms that arise naturally from the digital public sphere, such as self-mobilisation tactics of online political activism in South Korea. As cyberspace and the digital public spheres are constantly evolving, the norm-building has indeed become an ongoing process, with frequent updates both to the state's policies and to online democratic initiatives.

The policies and initiatives adopted by the government focused mainly on regulation through a verification and monitoring process, but also on education and information-sharing activities. The government's desire to transform a recently democratised country through digitisation has created an environment in which the digital public spheres, although facing some strict identification and verification regulations, can be actively engaged in the norm-building process thanks to a high level of cyber-literacy achieved through the government-initiated programmes and self-mobilisation routines that evolved from civil society's participation in cyberdemocracy practice.

CAPACITY BUILDING AND PUBLIC-PRIVATE PARTNERSHIPS

CHAPTER 9

Closing the cyber-capacity gap in digital financial inclusion

A critical analysis of prevailing narratives and approaches

NANJIRA SAMBULI AND ADITI BAWA

Introduction

In discussions about governing cyberspace, cyber diplomacy and digitalisation more broadly, it is common to find capacity-building and/or development highlighted as a challenge, opportunity or recommendation. This is especially so regarding developing countries. Capacity-building as a definitional concept in international development traces its origins to the mid-1990s, when shortcomings were identified in the prevailing approaches to development aid and technical assistance, in place since the 1950s. The lack of domestic ownership, shortcomings in tailoring aid delivery to local demand signals, poor

coordination, and the inability to effect sustainable change are some of the issues identified as perceived failings of the 'traditional' approach.[400]

Capacity-building is sometimes conflated with, and other times distinguished from, capacity development.[401] The Organisation for Economic Cooperation and Development (OECD) defines capacity development as 'the process whereby people, organizations and society as a whole unleash, strengthen, create, adapt and maintain capacity over time'. The United Nations Development Program's (UNDP)[402] definition is similar to OECD's and distinguishes capacity development from capacity building by noting that the latter is a 'process that supports only the initial stages of building or creating capacities and assumes there are no existing capacities to start from'. While this is an important distinction, the terminology is less vital than the concept itself. For the purposes of this chapter, we will be using cyber capacity-building and development interchangeably.

As the internet and ICT for development space emerged in the early 2000s, international conversations naturally turned to capacity-building. Cyber capacity-building is often approached from a development angle, so in the mid-2000s actors from technologically advanced countries initiated cross-border mechanisms to assist other countries and organisations in maintaining safe, secure and open use of the digital environment.[403] For international organisations and partnerships to prioritise cyber capacity development concurrent to the digital space's creation itself shows a clear initiative to expand the benefits of cybersecurity. According to an Open-ended Working Group[404] (OEWG) report, cyber capacity-building/development is useful to help develop the necessary social capacities—skills, human resources, policies and institutions—that enable a more secure, stable and resilient ICT environment. The report further recognises that the capacity of each state to prepare for and respond to emerging cyber threats informs the international community's collective ability to do the same.

400 *Understanding Capacity-Building/Capacity Development: A Core Concept of Development Policy* (Strasbourg: European Parliament, 2017), available at: https://www.europarl.europa.eu/RegData/etudes/BRIE/2017/599411/EPRS_BRI(2017)599411_EN.pdf

401 *Perspectives Note: The Enabling Environment for Capacity Development* (Paris: OECD 2008), available at: https://www.oecd.org/development/accountable-effective-institutions/48315248.pdf

402 *Capacity Development: A UNDP Primer* (New York: United Nations Development Program, 2015), available at: https://www.undp.org/publications/capacity-development-undp-primer

403 Robert Collett, 'Understanding cybersecurity capacity building and its relationship to norms and confidence building measures', *Journal of Cyber Policy* 6 (3) (2021), 298–317.

404 OEWG, *Open Ended Working Group on Development in the Field of Information and Telecommunications in the Context of International Security* (Geneva: UN General Assembly, 2021), available at: https://front.un-arm.org/wp-content/uploads/2021/03/Final-report-A-AC.290-2021-CRP.2.pdf

Capacity-building/development recommendations are typically centred around training efforts for financial regulators, technical individuals, not-for-profit institutions and end users. But other needs may be equal to or greater than training: for example, sustained institutional funding. And, within training programmes, there might be an under-appreciated need to include underrepresented groups. A critical reading of these calls for capacity-building highlights that it is not entirely clear for whom capacity is being developed, if it is demand-driven or based on contextual needs analysis, or even what exactly counts as successful capacity-building. Furthermore, the literature has yet to show how cyber capacity-building works in practice; it also does not explore unintended consequences of prevailing approaches or areas where changes in approach could be beneficial. For instance, it is difficult to ascertain which training approaches work well in environments where technological leapfrogging presents novel cybersecurity challenges.

Increasingly, there are also geopolitical connotations to consider when thinking about cyber capacity-building. For example, the US and the EU recently created a plan to support critical infrastructure technology in developing countries. This initiative is being framed as assistance to 'counter China'[405] and dissuade countries from accepting China's support, as well as to defend digital democracy.[406] Elsewhere, cyber capacity-building initiatives supported by the US government speak of 'helping establish the U.S. as the cyber development partner-of-choice in areas contested by China and Russia'.[407] This framing further complicates how cyber capacity-building supply matches local demand, as technological great power competition may not be a priority for developing countries.[408]

This chapter aims to help proponents and providers of capacity-building to better understand what capacities and resources various constituencies in developing countries most need. This may not be what supply-driven capacity-building programmes often provide. The chapter is informed by the authors' ongoing work to coordinate global efforts to advance cybersecurity as a priority consideration in digital financial systems, with a special focus on cybersecurity

405 Catherine Stupp, 'U.S., EU plan joint foreign aid for cybersecurity to counter China', *Wall Street Journal* (15 June 2022).

406 Matthew Gooding, 'US and EU could fund cybersecurity improvements in developing countries', *TechMonitor*, available at: https://techmonitor.ai/policy/geopolitics/us-eu-cybersecurity-china-russia

407 Bill Eidson, 'MITRE strengthens cyber capacity of developing nations', *MITRE* (December 2019), available at: https://www.mitre.org/publications/project-stories/mitre-strengthens-cyber-capacity-of-developing-nations

408 David Ehl, 'Africa embraces Huawei technology despite security concerns', *DW* (8 February 2022), available at: https://www.dw.com/en/africa-embraces-huawei-technology-despite-security-concerns/a-60665700

considerations for digital financial ecosystems across Africa. Digital financial inclusion is a key driver of digital technology adoption across the continent. Through experience gained by leapfrogging legacy infrastructure and systems and contextually adapting varied technologies to connect the unconnected, Africa has a lot to contribute to prevailing discourses on capacity-building based on the fast-growing fintech sectors in several countries. Furthermore, fintech and digital financial inclusion developments intersect uniquely with other digital development goals such as access to education, health and reliable energy, therefore gains made in clarifying what capacity-building entails in these dynamic digital finance environments could have far-reaching benefits across other domains. Enhanced and nuanced understanding of capacity-building challenges in developing and emerging markets can help elevate everyone's contributions to the global governance of cyberspace and digital technologies.

Cyber capacity and the financial system: insights from the FinCyber Strategy

The financial sector has been particularly attuned to cyberspace's opportunities and threats. Following a series of cyber-attacks that laid bare systemic risks that cyberspace poses to financial stability, stakeholders worked together towards a strategy for the international community to better protect itself against cyber threats.[409] Capacity building was identified—alongside cyber resilience, cyber workforce challenges and international norms and collective response mechanisms—as a priority in the resulting 'FinCyber' strategy.[410] As noted in the strategy report, 'Cybersecurity capacity-building has therefore become a growing priority, especially considering the rising numbers of state-sponsored attacks and the increase in fraud during the coronavirus pandemic. At the same time,

409 Tim Maurer and Arthur Nelson, *International Strategy to Better Protect the Financial System Against Cyber Threats* (Washington, DC: Carnegie Endowment for International Peace, 2020), available at: https://carnegieendowment.org/2020/11/18/priority-5-capacity-building-pub-83113

410 The project was led by Carnegie and comprised an international advisory group as well as inputs by over 200 stakeholders. Carnegie Endowment for International Peace, 'FinCyber Strategy Project: Cybersecurity and Financial Inclusion', available at: https://carnegieendowment.org/specialprojects/fincyber/financialinclusion/

"capacity-building" is an amorphous term and requires clarification before anyone can progress from concept to action.'[411]

Existing international cyber capacity-building recommendations and initiatives for the financial system vary. They encompass efforts to increase financial institutions' cyber resilience and strengthen law-enforcement, supervisory and regulatory capacity.[412] Others include providing resources and coordination centres to support information sharing and cybersecurity coordination to bolster cybersecurity norms,[413] as well as private sector-driven capacity-building through training and education for clients.[414] These efforts are increasingly targeted at low-income and developing countries, recognising that regions like Africa have also seen a surge in cybercrime activities.[415] Furthermore, given the variance of financial systems in Africa, where mobile technology has been a key driver of financial inclusion, the cybersecurity and cyber-resilience framing expands beyond traditional financial systems such as banks.[416] The nexus of financial inclusion and cybersecurity requires capacity-builders with expertise on financial inclusion as well as cybersecurity to ideally coordinate their efforts in ways that will bolster cyber capabilities for digital financial inclusion.

Cyber capacity-building efforts for digital financial inclusion—which the authors are observing through an ongoing project dubbed 'CyberFI'[417]—typically comprise 'top-down' development sector approaches to supporting developing countries. That is, capacity is considered to come from the developed markets, to be passed along to developing countries. This is noted in the concentration of trainings as a focus for capacity-building, as well as the overwhelming reference to capacity-building rather than capacity development. The latter would proceed more from the ground up and incorporate the insights and practical expertise

411 Maurer and Nelson (see note 10 above).

412 'Cyber Resilience and Financial Organizations: A Capacity Building Toolbox', Carnegie Endowment for International Peace (2021), available at: https://carnegieendowment.org/specialprojects/fincyber/guides

413 Silvia Baur-Yazbeck and Jean-Louis Perrier, 'Regional center can help low-income countries build cyber resilience', *CGAP* (8 July 2020), available at: https://www.cgap.org/blog/regional-centers-can-help-low-income-countries-build-cyber-resilience

414 'Helping customers strengthen their cyber defences', SWIFT, available at: https://www.swift.com/myswift/customer-security-programme-csp

415 'There are more cyberattacks in Africa than anywhere else', *WeeTracker* (12 January 2022), available at: https://weetracker.com/2022/01/12/there-are-more-cyberattacks-in-africa-than-anywhere-else/

416 Nanjira Sambuli and Taylor Grossman, 'Introducing CyberFi: perspectives on cybersecurity, capacity development, and financial inclusion in Africa', Carnegie Endowment for International Peace (2 May 2022), available at: https://carnegieendowment.org/2022/05/02/introducing-cyberfi-perspectives-on-cybersecurity-capacity-development-and-financial-inclusion-in-africa-pub-87001

417 Carnegie Endowment for International Peace, 'Securing Digital Financial Inclusion', available at: https://carnegieendowment.org/programs/technology/securingDigitalFinancialInclusion

of intended beneficiaries. Furthermore, cyber capacity-building for developing countries would factor in the reality that many are primarily focused on advancing a domestic agenda of digital inclusion before prioritising cybersecurity, even as technological leapfrogging introduces vulnerabilities.[418]

There are ongoing efforts on knowledge dissemination and resource mapping with the aim of improving coordination and collaboration between disparate cybersecurity and digital financial inclusion stakeholder groups. Typical stakeholder groups targeted by cyber capacity-building endeavours include developing country governments and their cybersecurity capacity at the national level; financial institutions and their organisational and clients' cybersecurity; and small businesses, which are increasingly vulnerable to cyber threats.[419] Products range from toolkits to technical assistance, consultancy and advisory services,[420] research methodologies and maturity assessments,[421] as well as opt-in operational assessments of existing cybersecurity processes for small and medium-size enterprises.[422] A portal managed by a core coordinating body, the Global Forum for Cyber Expertise (GFCE), serves as one repository of projects, tools, publications and other resources pertaining to cyber capacity-building for financial inclusion,[423] signalling an appreciation for mapping existing and upcoming initiatives in a sector gaining interest and focus from the international community and its support for low-income and developing countries' financial systems.

CyberFI has also noted the importance of focusing on gender and cybersecurity if the overall goals of secure development are to be achieved. Identified efforts entail a focus on gender disparities in the cyber workforce[424] as well as the gendered impacts of cybersecurity threats, whereby women in developing

418 Melissa Hathaway and Francesca Spidalieri, *Integrating Cyber Capacity in the Digital Development Agenda* (Global Forum on Cyber Expertise, 2021), available at: https://thegfce.org/wp-content/uploads/2021/11/Integrating-Cybersecurity-into-Digital-Development_compressed.pdf

419 Cybil Portal, 'Design of a Cyber Security Capacity Building Tool Kit for Governments', Global Forum for Cyber Expertise (April 2018), available at: https://cybilportal.org/projects/design-of-a-cyber-security-capacity-building-tool-kit-for-governments/

420 Eidson (see note 8 above).

421 Cybersecurity Multi-Donor Trust Fund, 'The World Bank Announces the Launch of a New Trust Fund on Cybersecurity', The World Bank, available at: https://www.worldbank.org/en/programs/cybersecurity-trust-fund/overview

422 'Cylab – Africa Operational Assessment Research', Carnegie Mellon University Africa, available at: https://cylab.africa.cmu.edu/

423 Cybil Portal, 'Financial inclusion', available at: https://cybilportal.org/projects-by?page=tag&_sft_post_tag=financial-inclusion

424 Muhammad Khurram Khan, 'Overcoming gender disparity in cybersecurity profession', *G20 Insights, Global Foundation for Cyber Studies and Research* (December 2020) available at: https://www.g20-insights.org/policy_briefs/overcoming-gender-disparity-in-cybersecurity-profession/

countries are more susceptible to cyber fraud because of existing gender inequalities.[425] Gender mainstreaming in cyber capacity-building for digital financial inclusion is welcome progress, and signals the influence the development community can wield in shaping capacity resources at this intersection.

Towards effective and sustainable cyber capacity-building and digital financial inclusion: tensions and emerging questions

Cybersecurity will continue to gain momentum as a digital development priority. As stakeholders work to boost cyber capacity within the digital financial ecosystem, we note the following tensions at this intersection that warrant more debate, reflection and deliberation among development practitioners.

Firstly, there is an insufficient or unclear delineation of intended beneficiaries of cyber capacity-building within the digital financial ecosystem. For instance, initiatives targeting government officials would do well to specify further which arms of government are targeted, as well as the specific objectives informing the endeavour. For instance, ministry officials, sector regulators, legislators, judicial officers and law enforcement are all government stakeholders, yet their cyber-capacity needs will vary. Generalisations such as 'cyber capacity-building for government officials' could inadvertently lead to a skewed focus on some stakeholder groups more than others. This could be further complicated by 'scaling', where what is seen to work in one context is then supported for replication in another context. This assumes that a subset of approaches, such as support in developing legislative frameworks, and stakeholder groups such as regulators, are the main functions and constituencies in need of capacity support. A specificity of objectives, and of which actors are targeted or reached by cyber capacity-building, will improve collective understanding on whether efforts undertaken are one-off or continuous engagements.

425 Michael Wechsler and Samikshya Siwakoti, 'Gender, Cybersecurity & Fraud' (Spring 2022), available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4103747

Secondly, it is not evident whether cyber capacity-building initiatives for digital financial inclusion are driven by the people and institutions who seek assistance or by those who seek to supply it. Do programmes follow from locally driven needs-based assessments or from external actors' assumptions about what should be needed? If the latter, to what extent are those assumptions rooted in the local digital finance ecosystem—including sometimes idiosyncratic fintech dynamics—versus presumptions of more traditional, formal financial systems? For instance, cyber capacity-building efforts often aim to address a lack of appropriate regulation or 'indigenous expertise' in developing countries by drawing on experience and expertise gained in technologically advanced countries. However, this approach may fail to account for the disproportionate adoption rates of innovative financial technology—for example, the ubiquitous M-PESA system in Kenya—and therefore incorrectly account for the unique challenges and opportunities for capacity development in such a financial landscape.

The unique capacities among regulators, service providers or consumers ought to be factored into framing how cyber capacity-building for an ecosystem like Kenya's, for example, is conceptualised as well as deployed and monitored for impact. In this case, the most pressing capacity needs for Kenya's regulators—who have otherwise been trailblazers in mobile money regulation[426]—might be not training but strengthening coordination between line ministries, for example. Other developing markets in Africa could also benefit from capacity-building efforts that feature Kenyan regulatory actors' experience and expertise as a knowledge resource or as a peer exchange mechanism.

Third, as cybersecurity in digital financial inclusion gains further credence, a significant consideration is how to foster coherence and synergies in cyber capacity-building efforts for all stakeholders in the ecosystem. There is the risk of duplication of initiatives as more actors become interested in building cyber capacity for digital financial inclusion. As mentioned previously, there tends to be a strong focus—on the supply side of cyber capacity-building for digital financial inclusion—on training and accreditation, knowledge repositories and mapping of stakeholders. However, in some jurisdictions the most impactful cyber-capacity support could be sustained financial resources to implement good practices, such as national or financial sector-specific cyber emergency response teams (CERTs). Fitting demand- and context-driven capacity needs to already defined

426 Njuguna Ndung'u, 'A digital financial services revolution in Kenya: the M-PESA case study', *African Economic Research Consortium* (November 2021), available at: https://aercafrica.org/african-governments-challenged-to-rethink-fiscal-policy-as-part-of-economic-recovery/

capacity-support mechanisms risks undermining local ecosystems' incentives to critically assess where they fall short and to develop suitable approaches to good practices such as information sharing, and even for recipient governments to bypass consulting local experts and stakeholders in favour of the 'international expertise' that shapes capacity-building support. This creates a risk compounding over time, in which knowledge and financial investments may not lead to more, improved or sustainable cyber capacity. Related to this is the perennial question of how to go beyond developed–developing, donor–beneficiary dichotomies that assign expertise and capacity needs. How can bottom-up insights be incorporated into cyber capacity-building or development endeavours, especially in the case of developing countries whose digitalisation trajectories have entailed aspects of leapfrogging and agility, both in infrastructure and of personnel?

Fourth, although there is a plethora of literature on cyber capacity-building, there aren't readily available assessment frameworks for what counts and does not count as a successful cyber capacity-building measure, even in a sub-sector like digital financial inclusion. It is widely accepted in the development sector that 'trainings' and education are the key to building capacity. We contend, however, that the priority placed on training and education, often conducted in one-sided and non-collaborative formats, assumes that recipient countries lack vital knowledge. This model may diminish opportunities for co-learning and regional or community-specific solutions to cyber-capacity needs. It also may overlook the contributions of self-taught practitioners.

What will count as successful cyber capacity-building?

Many aspects could be classified as cyber capacity-building/development for and in digital financial inclusion—from digital forensics skills for law enforcement and cybercrime incident response teams to digital financial literacy tools and cybersecurity awareness training for end users. Arguably, even physical infrastructure that connects people to digital financial products and services can be considered capacity-building. All are key components to advancing inclusive participation in cyberspace. However, for many important stakeholders, such as NGOs, international organisations and funding institutions, there remain fragmented definitions and distribution of resources to increase capacity-building/development. To improve the outcomes of cyber capacity-building in digital

development, and digital financial inclusion more specifically, we recommend more debate and reflection on the tensions discussed above. To this end, we propose an analytical framework to assess successful cyber capacity-building measures. An important aim would be to create guidelines for capacity-building measures that could mitigate unintended consequences and problematic 'solutions'. The points outlined below are framed within the context of digital financial inclusion but may be of relevance to broader digital transformation cyber capacity-building endeavours.

Context-rooted training as a cyber capacity-building measure

Training, evidently, is a favoured approach for enhancing cyber capacity. To help stakeholders get a better sense of what works—be it to impart technical skills or to develop legislative frameworks—we propose that those conducting such trainings outline if they refer to or conduct one-off or continuous trainings. Additionally, post-training assessments should be conducted to evaluate which approaches work and under what conditions. Such assessments can enable iterative improvements and replication of the most effective models.

Self-training modules can further complement time-bound cyber capacity training such as through workshops, to allow for in-person engagements to be more interactive and engaging for participants. Existing mechanisms such as the Global Forum on Cyber Expertise could facilitate vibrant exchanges and mapping of the cyber capacities most needed in given development contexts. This mapping could identify types of training that would be more demand-driven as well as those that are already being provided and by whom. This will augment international coordination efforts to better identify gaps and redirect resources from crowded domains.

For instance, MITRE's National Cyber Strategy Development & Implementation (NCSDI) framework[427]—through consultation with intended beneficiaries—can serve as a starting point for mapping different aspects of cyber capacity needs and demands in a government context. It outlines eight strategic areas across two analytical levels in identifying existing capacity and aspirational needs.

427 MITRE, 'National Cyber Strategy Development & Implementation Framework – Assessment Phase', Global Forum for Cyber Expertise (May 2020), available at: https://cybilportal.org/wp-content/uploads/2020/05/Cyber-Capacity-Assessment-Phase-Overview1.pdf

Other frameworks—overarching or niche—can be developed by interested parties to help take stock of the overall cyber capacity-building efforts under way, or that need undertaking, so that training supply better meets demand.

Demand-driven and contextual cyber capacity building efforts

We posit that cyber capacity-building efforts ought to be implemented following a comprehensive assessment of the needs (demands) of the intended beneficiaries. This necessarily includes the establishment of success metrics. The onus is on the development community to test their assumptions before designing programmes, financial resources or assistance mechanisms. One-size-fits-all approaches that tend to be popular in development assistance can fail to accommodate the complexities of the different ecosystems, a particularly important consideration as pertains to digital financial inclusion.

It is important to factor in the different politico-economic dynamics that will inform the unique determinants of what ends up being considered successful and thus scalable.[428] An insistence on scaling of top-down approaches risks undermining the agency of recipient countries in assessing and articulating their capacity needs. This can disincentivise commitment to the prevailing capacity-building approaches, which could undermine motivation—a crucial success determinant. Development practitioners ought instead to frame scaling as an interoperability of diverse and sustainable approaches and modalities, rather than a cut-and-paste from one context to another.

Interdisciplinary approaches

Neither cybersecurity nor capacity-building is a single issue, and nor is cyber capacity-building. Interdisciplinary approaches are key to addressing cyber capacity-building challenges. Interdisciplinarity in general is one of the most vital concepts that can be applied to sustainable global development. Social, political and economic dynamics are critical in closing the very gaps that technologies may create. We contend that it is important to ensure that cyber capacity training

428 Sebastian Pfotenhauer, Brice Laurent, Kyriaki Papageorgiou and Jack Stilgoe, 'The politics of scaling', *Social Studies of Science* 52 (1) (2021), available at: https://journals.sagepub.com/doi/full/10.1177/03063127211048945

is not framed only as an imparting of technical skills: psychology, behavioural analysis and other non-technological fields also have a lot to offer to the cyber workforce.[429] For instance, social engineering is a common cybercrime tactic[430] with regard to digital financial ecosystems. Effective countermeasures demand more than technical capability to not only redress but also mitigate future digital finance-related cybercrimes. Configuring training to accommodate the interests and insights of these disciplines will likely vary depending on the local context. Attending to these local variations will help ensure that cyber-capacity efforts are impactful beyond the training event.

A plethora of institutions and specialisations are involved in cyber capacity-building. Resource management, cyber resilience and organisational change in implementing bodies are among the functions that capacity-building will need to address. Recently, the US and the EU have introduced a wave of initiatives to 'fund improvements to the cybersecurity of critical infrastructure in developing countries aiming to help these nations better withstand attacks and improve the international community's overall online resilience'.[431] This seems promising; however, it is important for these efforts to be synergised and implemented in a manner that is not duplicative or siloed. These cybersecurity and cyber capacity-building efforts must also be contextually appropriate to where, when and how they are deployed.

Gender and cyber capacity-building

It is laudable that gender, as discussed earlier, is an early emphasis in discussions and resourcing for cyber capacity-building. Creating inclusive cyber workforces is one important goal. The aim, however, should be not only to train, mentor and support more women, but also to investigate and address the systemic issues that perpetuate a gendered divide in cybersecurity and technology workforces more broadly.[432] Another important priority is mainstreaming approaches to

429 Joanne Hall and Asha Rao, 'Non-technical Skills Needed by Cyber Security Graduates', IEEE, available at: https://ieeexplore.ieee.org/document/9125105

430 Silvia Baur-Yazbeck, Judith Frickenstein and David Medine, 'Cyber Security in Financial Sector Development', *CGAP* (2019), available at: https://www.findevgateway.org/sites/default/files/publications/files/cyber_security_paper_november2019.pdf

431 Gooding (see note 7 above).

432 Nanjira Sambuli, 'Reflection on "Women in Tech" Narratives', Observer Research Foundation (October 2021), available at: https://www.orfonline.org/expert-speak/reflection-on-women-in-tech-narratives/

cybersecurity in which an appreciation can be cultivated of how gender shapes identities, roles and expectations within society and even cybersecurity. Gender informs social structures and attendant hierarchies, often attributing technical expertise to masculinity and earmarking skills such as communication, or initiatives promoting diversity, equity and inclusion, as concerning women or femininity.[433] Gendered perspectives can sharpen cybersecurity design, defence and response mechanisms to mirror the reality that neither technology broadly, nor cybersecurity more specifically, is gender neutral.

Complementarity of efforts by countries and institutions providing cyber-capacity assistance

One of the main goals of forums such as the CyberFI process is to identify and avoid gaps and duplications in cyber capacity-building efforts. This can be done in many ways. The Cybil portal is one example of how initiatives can be mapped in a broad domain like cyber capacity-building, with specific focus on niche areas such as digital financial inclusion. If implementers continue to use and support such a tool both as a reference and as an active contributor, duplication can perhaps be minimised and resources better coordinated through strategic complementarities. Honest and authentic communication between funders, implementers and beneficiaries may also help mitigate some of the barriers that can arise from inadvertent duplications of cyber-capacity efforts.

Critical evaluations by funders and implementers on successes and gaps

Sustainable cyber capacity-building goes well beyond initial investments in the workforce and distribution of training and other resources. Continuous and retrospective evaluations of what has and has not been successful is key to future beneficial cyber-capacity efforts. Sharing among implementers what works and, perhaps more importantly, what doesn't work—and drawing honest assessments from intended beneficiaries—is a crucial part of such evaluations.

433 UNIDIR, 'Gender Approaches to Cybersecurity: Design, Defence and Response', Association for Progressive Communications (February 2021), available at: https://www.apc.org/en/pubs/gender-approaches-cybersecurity-design-defence-and-response

It can contribute to a vibrant information-sharing culture within the cyber capacity-building domain. Such evaluation exercises, however, ought to be careful not to place an inadvertent burden on support recipients to continually explain the challenges and opportunities faced without deliberate and continual improvement in how support is tailored to address the identified pain points. Intermediary institutions, such as think-tanks and other specialised nonprofits, can help facilitate dialogue between stakeholders and lend an outside analytical perspective.

Sustained resources for institutions and programmes

One-off or time-bound cyber capacity-building projects may serve as useful catalysts. However, and especially in development contexts, a bigger determinant of success and entrenchment of good practices is often tied to the continued availability of resources for programmes that deliver, as well as the institutions that the capacity efforts are supposed to benefit. It is one thing to train a CERT workforce; it is another for a national or sector-specific CERT to be adequately resourced to achieve the stated objectives. Thus, coupling resourcing needs with training and other capacity building modalities such as cyber strategies is imperative to secure short- and long-term success in identifying, meeting and sustaining cyber capacity.

Conclusion

As digital transformation continues to pervade nearly every sector, building and maintaining capacity to enable resilience against the inevitable cyber threats while simultaneously creating equitable and inclusive digital systems is no small feat. There are large differences in cultures of information sharing and overall receptivity to technological shifts across regions. The digital finance system is a perfect example of this—informal cash-based banking ruled much of the infrastructure in African countries and the Indian subcontinent. Now, as there is a conspicuous shift towards and rapid uptake of mobile money banking, digital payment systems and even cryptocurrency, cyber vulnerabilities are on the rise and will persist. It is one thing to address these threats; it is another to do so in

a manner that does not inadvertently undercut the people that the systems are working to serve.

Increasing cyber capacity is one way to work to enable resiliency of systems while investing in resources. However, without clear metrics or measures of success in the cyber capacity-building community outlining what is and is not successful, even in a niche subset such as the digital financial inclusion sector, efforts to bolster capacity may be shots in the dark. Within the digital financial space, it is still unclear not only what exactly the cyber capacity-building measures are, but also who the intended beneficiaries are and what drives certain capacity-building efforts over others. This is not to say that capacity building does not exist—rather that digital financial inclusion spaces do not have a clear picture of how cyber capacity-building is measured and how successful capacity efforts are sustained to promote long-term cyber resilience.

We propose in this chapter a few measures of cyber capacity-building that stakeholders can look towards to assess whether capacity-building measures are successful, inclusive and sustainable. We posit that cyber capacity-building initiatives—for digital financial inclusion and digital development in general—will likely be more successful when (1) training-based measures are coupled with other efforts to lessen solely 'educate-first' narratives; (2) an interdisciplinary approach to cyber capacity is applied, encompassing intersectionality and unique efforts; (3) efforts are demand-driven and not based on an assumed or presumptive assessment of the cyber capacity needs of a given region; (4) capacity-building efforts operate in alignment with the reality that technology is not gender-neutral and therefore apply gender-responsive capacity approaches; (5) cyber-capacity efforts are not merely duplicates of prior or parallel and ongoing efforts; (6) communication between funders, implementers and beneficiaries is honest and authentic; and (7) there is regular feedback from implementers following cyber-capacity efforts about what has and has not been successful in practice.

Cyber capacity-building is a broad term. As digital expansion and transformation accelerates at a historic pace, its effects are not distributed equally. This is clear in the digital financial sector. It is up to the organisations, funders, implementers and institutions working on these issues to enable demand-driven efforts, as well as to establish clear measurements of success in capacity-building so that digital financial inclusion advances as a cyber-secure and cyber-resilient undertaking. The framework proposed above can hopefully serve as a starting point to create fluency with regard to cyber capacity-building or development as we advance further into the digital age.

CHAPTER 10

Shaping platform governance in Central Asia

Challenges and opportunities for human rights defenders and journalists

PAVLINA PAVLOVA

Introduction

Social media shapes the way societies communicate, mobilise and engage in politics. Online platforms constitute a vital but contested space in Central Asia, where people express discontent against a backdrop of censorship and surveillance. With the growing presence and influence of social media and messaging platforms, human rights defenders (HRDs) and journalists alike face new challenges and threats related to their engagement in digital space. Social networks are developed and maintained by private entities, and users have minimal influence over the rules, policies and practices with which they are requested to comply. State interference presents another layer of restrictions to citizens' access to information and ways of engagement—often in the form of imposed state controls on accessing and disseminating information. Governments employ

several restrictive tactics, including blocking online content and throttling or shutting down networks to prevent politically charged assemblies and public expressions of discontent, particularly around elections and protests.

Information control in Central Asia builds on historical, political and economic ties with Russia, while China increasingly engages in the region.[434] Being at the forefront of digital authoritarianism, the two countries not only influence policies and practices of other states' behaviour in cyberspace on a country level but exercise an influence on platform governance in regional terms. Amid the rising tensions in the aftermath of Russia's invasion of Ukraine, the approach with which Central Asian governments shape the online narratives and engage with stakeholders can indicate the emerging dominant form of platform governance in the region. In the meantime, the growing pressure on media and online spaces to provide more 'neutral' coverage of the war creates an uneasy landscape for HRDs, journalists, and others who are left to navigate the tightening net of repression.

The challenges and opportunities pertaining to the use of social media by HRDs and journalists in Central Asia are outlined below in five sections. The first section introduces the information control landscape and underlying trends in the region relevant to social media platforms. The second section provides a framework for understanding domestic, regional and international stimuli for internet policy development, emphasising potential interdependencies between the practices of neighbouring states, and the mechanisms by which information control spreads. The section indicates the emergence of a set of shared characteristics and assesses freedom of expression, information and assembly in the region. The following section outlines the impact of the war in Ukraine on online freedom in Central Asia, which is set to deteriorate further in the aftermath of the conflict. The fourth section looks at the responses by social media companies. The chapter concludes with recommendations for closing the gaps in the participation of non-state actors to reconcile the asymmetric relations between actors in platform governance.

The focus is on national-level, state-mandated control of online information and the related underlying trends in Central Asia, but the arguments raised extend to other censorship and surveillance practices and regions, especially those with similar characteristics in the post-Soviet region. The chapter aims to contribute to the body of knowledge about platform governance and online freedom

434 Janko Šćepanović, 'Can Russia still be a dependable "sheriff" for Eurasia?', The Diplomat (30 September 2022), available at: https://thediplomat.com/2022/09/can-russia-still-be-a-dependable-sheriff-for-eurasia

in the region, where social media emerged as important spaces for communication but contextual knowledge about these dynamics remains limited. The thesis of internet policy diffusion and coordination in the region is examined.[435] It is observed that countries of Central Asia, each to a different degree, have developed and growingly relied on regulating the flow of online information. Furthermore, while demonstrating a variety of policies and methods, the approaches to information control adopted in Central Asia have shown similar patterns—generally diffusing from Russia or being imitated from the example of the neighbouring country. By analysing and comparing case studies from across the region, the chapter aims to enrich the analytical framework of networked authoritarianism, defined as 'a form of internet control common in former Soviet states where manipulation over digitally mediated social networks is used more than outright censorship' and previously applied to countries in the Caucasus.[436] Since case studies come with a risk of empirical generalisation, more research is needed to test the outlined hypothesis.

Internet freedom and the social media landscape in Central Asia

Internet expansion in the Central Asian region, while necessary for economic viability and development, posed challenges for states seeking control over the impact of technology on society and political processes. By facilitating open and accessible spaces for public discourse, information exchange and engagement, social media networks introduced risks of increased political dissent, association and mobilisation. These challenges prompted a period of experimentation and adaptation across authoritarian-leaning countries[437] as active online users and the potential for political dissent and instability rose. Freedom House ranking considers both Uzbekistan and Kazakhstan as 'not free.' Tajikistan and Turkmenistan do not even rank, and Kyrgyzstan is ranked 'partly free' for

435 Jaclyn A. Kerr, 'Information, security, and authoritarian stability: internet policy diffusion and coordination in the former Soviet region', International Journal of Communication 12 (2018), 3814–3834.

436 Katy E. Pearce and Sarah Kendzior, 'Networked authoritarianism and social media in Azerbaijan', Journal of Communication 62 (2) (2012), 283–298.

437 Kyrgyzstan is the only Central Asian country that has experienced protest-driven regime change and, according to its constitution, is a parliamentary democracy.

internet openness, while it is noted that the restrictions in the country remain significant.[438] Low internet connectivity, unreliability of the internet connection and high costs are still prohibitive for many users, especially in Tajikistan and Turkmenistan, where the combination of poor infrastructure and political repression hinders online participation.[439]

The social media landscape comprises mainly Russian and US-based platforms, but also Chinese video-sharing service TikTok.[440] Russian platforms Odnoklassniki, Vkontakte and and Moi Mir remain popular in Central Asia. Telegram is the most widely used instant messaging service with channels as a tool that enables users to broadcast public messages to subscribers.[441] It entered the market early and caters to an audience for which the Russian language continues to be the lingua franca of online content despite state promotion of native languages across the region.[442] Ethnic Russians constitute a notable percentage of the local population[443] and Russia accommodates a sizeable Central Asian diaspora, and labour migrants in particular—making Russian media a popular source of news. US-based social networks such as Facebook, Instagram, YouTube and Twitter have emerged as more popular among the urban society and the educated middle class as a forum for free expression and socio-political debates.[444]

Popular digital platforms can foster interest in civic involvement, while also serving as a source of alternative information. In a government-influenced media environment, online content offers independent journalists a space for news and public debate. For instance, scores of Kazakhstanis have turned to digital platforms to conduct journalistic investigations, discuss and analyse events in the country, report on political protests, and contradict the narrative served by

438 'Freedom on the Net report', Freedom House, available at: https://freedomhouse.org/countries/freedom-net/scores

439 Colleen Wood, 'Can social media change governance in Central Asia?', The Diplomat (25 April 2019), available at: https://thediplomat.com/2019/04/can-social-media-change-governance-in-central-asia

440 'Uzbekistan unblocks, re-blocks popular social media amid TikTok talks', Eurasianet (17 March 2022), available at: https://eurasianet.org/uzbekistan-unblocks-re-blocks-popular-social-media-amid-tiktok-talks

441 Murodjon Tuhtasinov, 'How Uzbeks learned to love (and live on) the Telegram messenger app', Global Voice (12 April 2019), available at: https://globalvoices.org/2019/04/12/how-uzbeks-learned-to-love-and-live-on-the-telegram-messenger-app

442 Mohammad Reyaz, 'Cyberspace in the Post-Soviet States: assessing the role of new media in Central Asia', Jadavpur Journal of International Relations 24 (1) (2020), 7–27.

443 Percentage of Russians in the total population as recorded by CIA Factbook: Kazakhstan, 19.3%; Kyrgyzstan, 5.1%; Turkmenistan, 4%; Uzbekistan, 2.3%; Tajikistan, less than 2%, available at: https://www.cia.gov/the-world-factbook

444 Bruce Pannier, 'Understanding Central Asia's cautious approach to Russia's invasion of Ukraine', Foreign Policy Research Institute (25 March 2022), available at: https://www.fpri.org/article/2022/03/understanding-central-asias-cautious-approach-to-russias-invasion-of-ukraine

traditional media. Kazakhstan's youth in particular are the target audience of the country's non-traditional media projects. This is in a country where, according to Reporters without Borders, journalism is viewed with widespread suspicion in society, but citizens tend to rely on bloggers or anonymous posts on social media.[445] Such dependence has both positive and negative aspects, as social media presents fertile ground for misinformation and disinformation. This is especially the case during critical periods such as the pandemic, when risks were growing due to the lack of timely and reliable information from the governments.[446]

Trends towards social media-based journalism, activism and debates can be also observed in Uzbekistan, where a strong youth base popularised online groups on social media platforms that allow for exchanging information on corruption issues, which the official media barely cover.[447] In Kyrgyzstan too, the high level of corruption leads to public demand for investigative work on these issues. At the same time, average users do not disseminate critical opinions and political ideas circulated by journalists.[448] These trends are set to continue and intensify with the growing numbers of people in the region actively using social media platforms.

Information control: legislation and practice

Central Asian governments have pursued several strategies to control the flow of and access to online information, experimenting with both repression and cooptation. Case studies will primarily cover Kazakhstan and Uzbekistan, the region's largest and most populated states, but similar trends have been observed throughout the region. To a large degree, the authorities have been playing catch-up—crafting and copying laws on online content after it emerged as problematic by amending their respective media laws. For example, the law on mass

445 Sher Khashimov, 'Kazakhstan's alternative media is thriving—and in danger', Foreign Policy (12 July 2021), available at: https://foreignpolicy.com/2021/07/12/kazakhstan-alternative-media-thriving-danger

446 Anastassiya Fershtey, 'Misinformation and conspiracies spread while Kazakhstan reimposes lockdown', The Diplomat (10 July 2020), available at: https://thediplomat.com/2020/07/misinformation-and-conspiracies-spread-while-kazakhstan-reimposes-lockdown

447 'Uzbekistan', Reporters Without Borders, available at: https://rsf.org/en/country/uzbekistan

448 'Kyrgyzstan', Reporters Without Borders, available at: https://rsf.org/en/country/kyrgyzstan

media[449] in Kazakhstan considers online platforms a type of mass media outlet, making companies legally responsible for online content on their platforms and subject to suspension or ban. In the same vein, authors of online content can be legally accountable for violating the law alongside journalists. The Ministry of Information and Social Development monitors content published online within the framework of an 'automated system of monitoring the national information space' that checks for content deemed illegal under the referenced law. Concerns arise around the system's potential misuse to monitor public discontent on social media and track dissent.[450]

The diffusion dynamics in the region were evident during the global outbreak of coronavirus. In the early stages of the pandemic, Central Asian countries followed the lead of Russia, where media outlets charged with deliberately spreading 'false information' about public safety risked heavy fines. Within weeks, the Russian media regulatory agency began using the updated rules to block, censor and fine online media critically reporting on the handling of the health crisis.[451] Governments in Central Asia also failed to uphold human rights obligations in their responses to the public health emergency. Coronavirus-related measures were accompanied by censorship of access to information about the spread of the virus and implementation of restrictions in discriminatory or arbitrary ways, often targeting members of civil society, HRDs and journalists.[452] Uzbekistan passed amendments criminalising the spread of 'false information' in April 2020.[453] Those found guilty of publishing 'fake news' could face fines or up to three years in prison under the temporary rules. Similar provisions came into force in Tajikistan in July 2020[454]—making it illegal to distribute 'inaccurate' and 'untruthful' information about Covid-19 through the press, social networks 'or

449 Law of the Republic of Kazakhstan dated 23 July 1999, No. 451-I. About the media (with amendments and additions as of 5 March 2022).

450 'Freedom on the Net 2019: Kazakhstan', Freedom House, available at: https://freedomhouse.org/country/kazakhstan/freedom-net/2019

451 'New "fake news" law stifles independent reporting in Russia on COVID-19', International Press Institute (8 May 2020), available at: http://ipi.media/new-fake-news-law-stifles-independent-reporting-in-russia-on-covid-19.

452 'Central Asia: Respect Rights in Covid-19 Responses', Human Rights Watch (23 April 2020), available at: https://www.hrw.org/news/2020/04/23/central-asia-respect-rights-covid-19-responses

453 Agnieszka Pikulicka-Wilczewska, 'Is Uzbekistan using coronavirus to curtail civil liberties?', Al Jazeera (3 April 2020), available at: https://www.aljazeera.com/news/2020/04/uzbekistan-coronavirus-curtail-civil-liberties-200403074921162.html

454 Daria Litvinova, 'Fake news or the truth? Russia cracks down on virus postings', AP News (1 April 2020), available at: https://apnews.com/article/health-ap-top-news-international-news-moscow-virus-outbreak-dbbf02a747b11d8ffe3b07d5e33ff129

other electronic means'. However, the vague terms were instead misused to cover up the scale of Tajikistan's coronavirus outbreak.[455]

Prominent examples of suppression of freedom of expression online often occur around elections. For example, Uzbekistan changed its criminal code to make insults to the president illegal, outlining further penalties when the offences are committed in the online space. The provisions were signed by President Mirziyoyev in March 2021, prior to the presidential elections in October of that year. Restrictions have also been introduced in terms of barriers to online anonymity. The controversial 'false information' bill in Kyrgyzstan that came into power in August 2021 compels internet service providers to register their clients in a unified identification system and provides authorities with full information related to users if a court or a state agency requests such data. [456] The law also stipulates that the owners of social media accounts must have their personal data publicly available, while anonymous internet users would be located and removed.[457] While Kyrgyz authorities can find the implementation difficult in practice—lacking the necessary resources to monitor the online information and communicate their orders to social media providers—it retains its chilling effect on HRDs, journalists and activists.[458]

One side of the coin is the adopted legislation, which solely rests on the powers of the state authorities. But implementing the decisions requires cooperation with other stakeholders, especially private companies such as internet service providers, exporters of surveillance technology, and social media platform owners. For example, cooperation between social media companies and governments on removing content was prominently discussed in November 2021 when the Kazakh government announced it was granted access to Facebook's internal 'content reporting system' (CRS). The system would enable the Ministry of Information to promptly report content containing violations of Facebook's

455 'Rush to pass "fake news" laws during Covid-19 intensifying global media freedom challenges', International Press Institute (3 October 2020), available at: https://ipi.media/rush-to-pass-fake-news-laws-during-covid-19-intensifying-global-media-freedom-challenges; 'Tajikistan: journalists silenced, media under pressure', International Press Institute (3 May 2020), available at: http://ipi.media/tajikistan-passes-coronavirus-fake-news-law; 'Tajikistan: COVID-19 outbreak offers cover for fresh assault on free press', Eurasianet (12 June 2020), available at: https://eurasianet.org/tajikistan-covid-19-outbreak-offers-cover-for-fresh-assault-on-free-press

456 'How are the authorities in Central Asia trying to control the internet?', Human Rights Watch (18 November 2021), available at: https://www.hrw.org/news/2021/11/18/how-are-authorities-central-asia-trying-control-internet

457 Sher Khashimov and Colleen Wood, 'As press freedom shrinks in Kazakhstan, journalists are standing up for civil liberties', Waging Nonviolence (6 November 2021), available at: https://wagingnonviolence.org/2021/11/press-freedom-shrinks-kazakhstan-journalists-stand-up

458 Human Rights Watch (see note 23 above).

global content policy and local laws of the Republic of Kazakhstan.[459] This type of agreement would have been unprecedented in the region. Facebook parent company Meta denied the claim.[460] Facebook has long faced criticism from rights groups for being too compliant with government censorship requests, and the company received a backlash on the censorship attempt. Kazakhstan's presentation in calling the access 'exclusive' seems to be misleading, following Meta's statement that the company follows a single global process that is 'independent from any government' to assess content in line with Facebook's policies, local laws and international human rights standards. The company further clarified that it has a dedicated online channel for governments to report content that they believe violates local law. The announcement that the Kazakh government asserted was issued jointly has since been considered a subject of miscommunication, but the incident may result in a further chilling effect on politically charged criticism online.[461]

In May 2022, Kazakh President Kassym-Jomat Tokayev signed into law a controversial bill that requires foreign social media companies to set up a local presence. While labelled as an accountability measure to step up the fight against cyberbullying, the law introduced a potential vehicle for the authorities to exercise influence over private actors. Taking place in a region where vague terminology that allows for the abuse of restrictions on online resources on legal grounds is a common trend, implementation of staff or data localisation laws can be potentially used to exercise pressure on content that is seen as politically problematic.[462] Having in-country representatives and staff opens the door for the government to coerce the companies to comply with arbitrary censorship

459 'Kazakhstan granted access to Facebook's content system to flag "harmful content"', RFE/RL (1 November 2021), available at: https://www.rferl.org/a/kazakhstan-access-facebook-content/31539818.html

460 'Facebook caught up in Kazakhstan internet crackdown', Eurasianet (2 November 2021), available at: https://eurasianet.org/facebook-caught-up-in-kazakhstan-internet-crackdown; 'Meta denies Kazakh claim of exclusive access to Facebook's content reporting system', Reuters (3 November 2021), available at: https://www.reuters.com/world/asia-pacific/facebook-lets-kazakh-govt-directly-flag-harmful-content-joint-statement-says-2021-11-01

461 Catherine Putz, 'Meta pushes back against Kazakh claims of "exclusive" access to Facebook's content reporting system', The Diplomat (2 November 2021), available at: https://thediplomat.com/2021/11/meta-pushes-back-against-kazakh-claims-of-exclusive-access-to-facebooks-content-reporting-system

462 Human Rights Watch (see note 23 above); 'Kazakh president signs bill allowing social media to be shut down', RFE/RL (3 May 2022), available at: https://www.rferl.org/a/kazakhstan-law-social-media/31832653.html; Catherine Putz, 'Kazakh president signs controversial law aiming to control social media companies', The Diplomat (4 May 2022), available at: https://thediplomat.com/2022/05/kazakh-president-signs-controversial-law-aiming-to-control-social-media-companies

requests,[463] for example, under threats of imprisonment.[464] This was the case in the Russian parliamentary elections in September 2021, when the authorities used the threat of prosecuting employees to gain leverage against Apple and Alphabet's Google to remove the tactical voting online app by opposition leader Aleksei Navalny from their online stores.[465]

Platform governance can be further addressed at the level of internet infrastructure. Shutting down or throttling networks and service restrictions is a common practice in the region in response to critical events—demonstrating the technical means and political leverage of countries attempting to control access to and flow of information online. Access Now and the #KeepItOn coalition highlight the trend toward deepening digital authoritarianism globally. Their 2021 report recorded shutdowns affecting both broadband and mobile networks in all five Central Asian countries in response to protests while national security and public order were cited as justification.[466] A nation-scale internet blackout took place in Kazakhstan in early January 2022 as a reaction to mass protests erupting in the country.[467] The authorities first throttled the internet and imposed targeted blocks but later resorted to cutting off both broadband and mobile internet access almost completely in an attempt to curb the unrest.[468] Internet blackouts helped the regime to stifle the crowds at a decisive moment but led to a week of information chaos. Regulating information that spread across and outside of the country halted real-time reporting and organising through online channels that could otherwise trigger or enable even larger gatherings.[469] The Kazakh government justified the move with legal provisions on anti-terrorism and public

463 'Turkey: YouTube precedent threatens free expression', Human Rights Watch (18 December 2020), available at: https://www.hrw.org/news/2020/12/19/turkey-youtube-precedent-threatens-free-expression

464 Deborah Brown, 'US tech companies bow to Russian government', Human Rights Watch (21 September 2022), available at: https://www.hrw.org/news/2021/09/21/us-tech-companies-bow-russian-government

465 Anton Troianovski and Adam Satariano, 'Google and Apple, under pressure from Russia, remove voting app', New York Times (23 September 2021), available at: https://www.nytimes.com/2021/09/17/world/europe/russia-navalny-app-election.html

466 'The return of digital authoritarianism: internet shutdowns in 2021', #KeepItOn coalition, AccessNow (April 2022), available at: https://www.accessnow.org/cms/assets/uploads/2022/05/2021-KIO-Report-May-24-2022.pdf

467 'Behind the unrest in Kazakhstan', International Crisis Group (14 January 2022), available at: https://www.crisisgroup.org/europe-central-asia/central-asia/kazakhstan/behind-unrest-kazakhstan

468 Pavlina Pavlova, 'How Kazakhstan's control of information can turn into a regime weakness', Open Global Rights (31 January 2022), available at: https://www.openglobalrights.org/how-kazakhstans-control-of-information-can-turn-into-a-regime-weakness

469 International Crisis Group (see note 34 above).

security. On the technical level, shutdowns were enabled by state control over large segments of the country's telecommunication infrastructure.[470]

State authorities in the region have long tried to control or limit access to information also by throttling, blocking and filtering online content.[471] Depending on the goal, these practices are enabled mainly through distributed denial of service (DDoS) attacks, deep package inspection (DPI) methods or man-in-the-middle attacks. During their respective elections, both Kazakhstan and Tajikistan initially obstructed access to social media platforms for a few hours and then resorted to complete blocking of Instagram, Twitter and Facebook.[472] Apart from tactics that focus on limiting the availability of information, there is a growing tendency of weaponising online harassment against targeted HRDs and journalists. There have been recorded examples of gender-based harassment of female journalists in Central Asia, who were attacked by fake accounts originating in troll factories—with a significant impact on their work and security.[473] These cases and many other instances outline a worsening situation in the context of online freedom in Central Asia. The level of fulfilment of state obligations towards freedom of expression and access to information varies from country to country, but never achieves a conducive environment for journalists and HRDs to work freely and independently on sensitive topics, particularly those connected to political leaders and corruption.

Though approaches adopted throughout the region are not identical, there are notable common trends. Information control legislation and practices in Central Asia indicate the emergence of a set of shared characteristics that point to intraregional diffusion or coordination dynamics. At the forefront of the phenomenon of digital authoritarianism stand Russia and China, which developed

470 'Freedom on the Net 2020: Kazakhstan', Freedom House, available at:
https://freedomhouse.org/country/kazakhstan/freedom-net/2020; Katrina Keegan, 'Information chaos in Kazakhstan', The Diplomat (10 January 2022), available at: https://thediplomat.com/2022/01/information-chaos-in-kazakhstan

471 Catherine Putz, 'Uzbekistan unblocks Twitter, TikTok still restricted', The Diplomat (4 August 2022), available at: https://thediplomat.com/2022/08/uzbekistan-unblocks-twitter-tiktok-still-restricted

472 Human Rights Watch (see note 23 above).

473 'Kyrgyzstan: surveillance, marginalisation and targeting of LGBT defenders', UN Special Rapporteur on Human Rights Defenders (13 September 2021), available at: https://srdefenders.org/kyrgyzstan-surveillance-marginalisation-and-targeting-of-lgbt-defenders-joint-communication; 'Reinforcing media freedom and the safety of journalists in the digital age: report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan', Human Rights Council (3 June 2022), available at: https://reliefweb.int/report/world/reinforcing-media-freedom-and-safety-journalists-digital-age-report-special-rapporteur-promotion-and-protection-right-freedom-opinion-and-expression-irene-khan-ahrc5029-enarruzh

and exported distinct technology-driven playbooks.[474] China's influence has been exercised through digitalisation incorporated in the Belt and Road Initiative and exports of high-tech devices. Russia exercises a significant normative influence in the region, building on historical, cultural and socio-political ties leveraged by political, security and economic dependencies. Moscow's digital censorship model has proved more adaptable in Central Asia than China's high-tech model. The flexibility of the ad hoc model that utilised a combination of political, administrative, legal and technical means makes it well positioned to diffuse across a region that is in proximity to Russia in terms of power structures, legal systems and economic resources.[475]

The impact of Russia's war in Ukraine

Considering the strong presence Russia has in the region, the full-scale invasion of Ukraine on 24 February 2022 has unnerved the political elites in Central Asia. The combination of dependency and wariness of Russia's territorial ambitions causes the states to manoeuvre in an uneasy terrain. Regional elites largely opposed the invasion,[476] which they feared could be a pretext to turn on their territories as well. At the same time, as Russia faces a long period of isolation and sanctions, governments aim to reduce their dependence. They pursue a multi-vector foreign policy[477] to minimise the extent of collateral damage on domestic economies caused by sanctions while preventing negative ramifications from

474 Anna Gusarova, 'Culture of protecting personal data: from online freedom to digital surveillance?', Central Asian Bureau for Analytical Reporting (13 April 2020), available at: https://cabar.asia/en/culture-of-protecting-personal-data-from-online-freedom-to-digital-surveillance

475 Alina Polyakova and Chris Meserole, 'Exporting digital authoritarianism: the Russian and Chinese models', Brookings (27 August 2019), available at: https://www.brookings.edu/wp-content/uploads/2019/08/FP_20190827_digital_authoritarianism_polyakova_meserole.pdf

476 At the UN, none of the five Central Asian countries supported Russia in the 2 March resolution condemning the Ukraine invasion.

477 Rachel Vanderhill, Sandra F. Joireman and Roza Tulepbayeva, 'Between the bear and the dragon: multivectorism in Kazakhstan as a model strategy for secondary powers', International Affairs 96 (4) (2020), 975–993.

Moscow.[478] The early general circumspection on the issue, muted expressions of concern and statements of neutrality in the conflict have been superseded on occasion in Kazakhstan, Kyrgyzstan and Uzbekistan by open criticism and reiteration of Ukraine's territorial integrity.[479] Tajikistan and Turkmenistan, which are both going through a domestic power transition process, remained silent on the issue, with the latter publicly ignoring the war altogether.[480]

Russia sealed its information space in an attempt to control the narrative on the war in Ukraine. The government introduced sanctions and bans on social media platforms and adopted laws criminalising media coverage and online content for contradicting state views.[481] The severe social media bans have been described as a 'digital iron curtain' to maintain the state's hold on the dissemination of information coming from abroad as well as circulation of information inside the country.[482] Central Asia avoided Russia's scenario, but the authorities have exercised pressure on the media to provide 'neutral' coverage of the events amid domestic public opinion that is divided by the conflict.[483] Russia's influence through the soft power of its media and as the provider of financial means for Central Asian workers is tangible. Given the high share of remittances in the countries' GDP, labour migrants and their relatives find their standard of living and economic future tied to Russia. The financial and physiological dependency

478 Jeffrey Mankoff, 'Central Asia is keeping a nervous eye on Russia's war in Ukraine', World Politics Review (26 April 2022), available at: https://www.worldpoliticsreview.com/articles/30491/in-central-asia-russia-s-war-in-ukraine-is-raising-anxieties; Wilder Alejandro Sánchez and Kamila Auyezova, 'Kazakhstan cancels Victory Day in protest over Putin's Ukraine War', Atlantic Council (11 May 2022), available at: https://www.atlanticcouncil.org/blogs/ukrainealert/kazakhstan-cancels-victory-day-in-protest-over-putins-ukraine-war; Bradley Jardine, 'Russia's war in Ukraine spells disaster for neighboring Central Asia', Time (10 March 2022), available at: https://time.com/6156524/russia-ukraine-central-asia-impact

479 Paul Stronski, 'The common theme in Central Asia's response to Russia's invasion of Ukraine', Carnegie Endowment for International Peace (30 March 2022), available at: https://carnegieendowment.org/2022/03/30/common-theme-in-central-asia-s-response-to-russia-s-invasion-of-ukraine-pub-86764

480 Kirill Nourzhanov, 'Uneasy neutrality: Central Asia's response to the Ukraine crisis', Australian Institute of International Affairs (17 March 2022), available at: https://www.internationalaffairs.org.au/australianoutlook/uneasy-neutrality-central-asias-response-to-the-ukraine-crisis; 'Europe–Central Asia: polarisation to the west, war & propaganda to the east', Reporters Without Borders, available at: https://rsf.org/en/region/europe-central-asia

481 'Russia, Ukraine, and social media and messaging apps', Human Rights Watch, 16 March 2022, available at: https://www.hrw.org/news/2022/03/16/russia-ukraine-and-social-media-and-messaging-apps; 'Putin signs law introducing jail terms for "fake news" on army', Moscow Times (4 March 2022), available at: https://www.themoscowtimes.com/2022/03/04/putin-signs-law-introducing-jail-terms-for-fake-news-on-army-a76768; 'Russian media watchdog blocks Facebook after limiting access to multiple other sites', RFE/RL (4 March 2022), available at: https://www.rferl.org/a/russia-rferl-bbc-facebook-google-twitter-blocked/31735597.html

482 'Russian Instagrammers face uncertain future as government tightens control over social media', RFE/RL (22 March 2022), available at: https://www.rferl.org/a/russia-instagram-facebook-ban-impact/31765511.html

483 'Stop pressuring journalists in Central Asia over Ukraine war coverage, RSF says', Reporters Without Borders (19 April 2022), available at: https://rsf.org/en/stop-pressuring-journalists-central-asia-over-ukraine-war-coverage-rsf-says; Joanna Lillis and Ayzirek Imanaliyeva, 'Ukraine war inspires rival passions in Central Asia', Eurasianet (7 March 2022), available at: https://eurasianet.org/ukraine-war-inspires-rival-passions-in-central-asia

on Moscow [484] is coupled with the threat of possible repercussions for nationals living and working in Russia.[485] For example, a Tajik journalist, Negmatullo Mirsaidov, wrote on his Facebook page: 'Do not do your compatriots in Russia a disservice. Maintain your neutrality in the Russia-Ukraine conflict.'[486] Others fear that Central Asian countries may face a similar fate to Ukraine's.[487]

The invasion of Ukraine has been accompanied by information operations and propaganda sponsored by Russia, aiming to justify the war both to domestic audiences and abroad. This is especially relevant in the post-Soviet countries which the Kremlin considers its sphere of influence, and that present to a large degree a Russophone information space.[488] According to the UK's Foreign Office, such online content generally follows the Kremlin narrative, such as that Russia is combating a Nazi regime in Ukraine to liberate Ukraine's oppressed Russian-speaking citizens. Evidence of coordinated information operations has been detected across major online platforms, including Telegram, Twitter and Facebook, and has been particularly concentrated on Instagram, YouTube, and TikTok. Central Asia has been part of the global information war, with propaganda and disinformation having a higher resonance on the local populations due to the interconnectedness and interdependence of the regions. At the same time, Central Asians who express pro-Ukraine attitudes on social media are targeted by trolls and bots who fulfil a dual function of spreading propaganda to audiences and attacking critical voices.[489]

The external forces are accompanied by country-level censorship. The state authorities in the region reportedly caution journalists, bloggers and activists to exercise self-restraint on both sides when writing about the war.[490] Largely in response to pro-war content on social media, Kazakh law enforcement warned against succumbing to provocative statements and appeals in the media

484 Parviz Mullojonov, 'Official Dushanbe silent as Tajik society deeply divided on Ukraine war', RFE/RL (21 May 2022), available at: https://www.rferl.org/a/tajikistan-public-divided-war-ukraine/31861484.html

485 Nourzhanov (see note 47 above).

486 'Central Asian leaders mute on Ukraine, but markets and public reel', Eurasianet (24 February 2022), available at: https://eurasianet.org/central-asian-leaders-mute-on-ukraine-but-markets-and-public-reel

487 Stronski (see note 46 above)

488 Navbahor Imamova, 'Central Asian countries tread cautiously on Russia's war in Ukraine', Voice of America (1 March 2022), available at: https://www.voanews.com/a/central-asian-states-tread-cautiously-on-russia-s-war-in-ukraine/6465144.html

489 'UK exposes sick Russian troll factory plaguing social media with Kremlin propaganda', UK Foreign, Commonwealth & Development Office (1 May 2022), available at: https://www.gov.uk/government/news/uk-exposes-sick-russian-troll-factory-plaguing-social-media-with-kremlin-propaganda

490 Nourzhanov (see note 47 above).

and on social media platforms, and against inciting ethnic tension or questioning the territorial integrity of Kazakhstan.[491] In Uzbekistan, journalists and bloggers who wrote about the Russian invasion were warned by the authorities to cover the war in 'very neutral' terms. Some of them spoke about being interrogated by intelligence officers, while others said they were ordered to delete their work.[492] Several journalists and bloggers were reportedly summoned to the investigative department of the State Security Service. Nearly a dozen people were called in because of their coverage of the war. Among those who were called by the security agency were editors and managers of Uzbek kun.uz – an online news publication.[493] Meanwhile, traditional media avoid using words such as 'invasion' or 'aggression'.[494] Government officials claimed such measures were necessary to combat misinformation and disinformation but deny that independent media were silenced. 'Uzbek media are covering Ukraine,' said the chief country media regulator. 'No one is banned from touching the topic, but we must be neutral and unbiased.'[495] The lack of reliable information drives people to seek information from digital and foreign media, which are often represented by popular Kremlin-funded outlets and accounts on Russian social media platforms.

Responses by social media platforms

Private companies are the owners of online platforms enabling people to access and share information. It is therefore important to see how the companies behind popular social media reacted to the war and whether such steps met their responsibility to respect human rights. The UN Guiding Principles on Business

491 'Address of the Deputy Prosecutor General of the Republic of Kazakhstan Dembaev B.B. in accordance with Article 31 of the Law "On the Prosecutor's Office"', General Prosecutor's Office of the Republic of Kazakhstan (28 March 2022), available at: https://www.gov.kz/memleket/entities/prokuror/press/news/details/346474?lang=ru

492 Khurmat Babadjanov '"Very neutral": Uzbek journalists pressured over their coverage of Russian War in Ukraine', RFE/RL (9 March 2022), available at: https://www.rferl.org/a/uzbekistan-journaliists-pressured-ukraine-war/31741826.html

493 'In Uzbekistan, journalists are summoned to the special services for "incorrect coverage" of the war in Ukraine', Radio Azattyq (6 March 2022), available at: https://rus.azattyq.org/a/31738517.html

494 Reporters Without Borders (see note 50 above).

495 Navbahor Imamova, 'Fear of Russia drives Central Asian response to Ukraine war', Voice of America (27 April 2022), available at: https://www.voanews.com/a/fear-of-russia-drives-central-asian-response-to-ukraine-war-/6547957.html

and Human Rights[496] include provisions that require private actors to take steps towards addressing adverse human rights impacts that can be facilitated by the extent of their operations and to take actions in line with international human rights standards and in a transparent and accountable manner.[497] While the principles are not binding international law, they present an authoritative international statement on the responsibilities of business in regard to human rights. However, as Human Rights Watch[498] and other watchdogs[499] documented, online platform providers have been chronically failing in responding to human rights challenges in critical instances. In their reaction to Russia's invasion of Ukraine, companies have taken a wide range of steps to counter harmful disinformation, label or block Russia's state-sponsored or affiliated media and introduce additional safety measures.[500]

As pointed out by Natalia Krapiva, the tech legal counsel of Access Now, 'major tech companies have a responsibility to their Ukrainian and Russian users to respect their rights to freedom of expression and access to information, especially in the time of war and political crisis. They do, however, also have a responsibility to keep their users safe and identify and respond to any campaigns of disinformation that may result in violence and abuse.' Social media companies such as Twitter and Meta have tried to address a rise in war-related disinformation. By labelling posts from Russian state-controlled media, they have also added friction to potentially harmful content, resulting in limited appearances in online spaces, searches or automatic recommendations.[501] Facebook, Instagram, Twitter and YouTube have also played a major role in spurring global support for Ukraine. Viral images and videos reporting on the devastating effects of the war

496 'Guiding principles on business and human rights – implementing the United Nations "Protect, Respect and Remedy" Framework', United Nations (2011), available at: https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

497 Human Rights Watch (see note 48 above).

498 'Big Tech's heavy hand around the globe', Human Rights Watch (8 September 2020), available at: https://www.hrw.org/news/2020/09/08/big-techs-heavy-hand-around-globe; 'Social media's moral reckoning', Human Rights Watch (21 December 2018), available at: https://www.hrw.org/world-report/2019/country-chapters/global-6

499 'An open letter to Mark Zuckerberg', The Santa Clara Principles (n.d.), available at: https://santaclaraprinciples.org/open-letter

500 Human Rights Watch (see note 48 above).

501 David Klepper, 'New Twitter policy aims to pierce fog of war misinformation', AP News (19 May 2022), available at: https://apnews.com/article/russia-ukraine-twitter-inc-technology-humanitarian-crises-cb2ff8c5572bbf0a3ba894f3d0318627

in Ukraine and the impact on civilians have captured the world and helped to galvanise humanitarian, political and military support.[502]

Private actors have further exercised pressure on Moscow-backed information channels. Google blocked RT, Sputnik and other Russian state-sponsored channels on YouTube and discontinued their ad revenue. Facebook, which was banned by the Russian authorities after being declared an 'extremist' organisation, took similar steps against state media outlets. Apple stopped selling its devices on the Russian market and removed RT and Sputnik from its app store outside the country.[503] The war has also amplified intolerant, inciting and hateful online content, which forced social media platforms to adapt their content moderation policies and practice in real-time.[504] Still, companies' systematic responses to the spill-over effect of the Russia–Ukraine information war, which resulted in collateral damage, remain unclear. Yaroslav Tartykov, the editor of Factcheck.kg,[505] indicated that unchecked information 'flows through the internet'. According to his statements, unverified information comes both from the supporters of the Kremlin and those who support Kyiv.[506] It remains unclear how the social media platforms improve content moderation beyond their immediate responses and whether the particular intricacies and risks in the countries of Central Asia have been considered—for example by increasing the number of language-specific and context-aware content moderators or strengthening cooperation with civil society organisations to fact-check online content.

The actions outlined above, however, follow a track record of tech companies' compliance with authoritative governments, which in its effect undermined activist groups, impaired access to reliable information and, essentially, became untenable in the wake of the invasion.[507] Cooperation between private companies

502 John Thornhill, 'War in Ukraine underlines the need for Telegram to protect its users', Financial Times (24 March 2022), available at: https://www.ft.com/content/bb4ff22c-ac64-4423-a679-f893d3a1d117

503 Elizabeth Dwoskin, Gerrit De Vynck and Taylor Telford, 'Silicon Valley companies have been rewriting their rules during the war in Ukraine. Russia is retaliating', Washington Post (11 March 2022), available at: https://www.washingtonpost.com/technology/2022/03/11/russian-prosecutor-general-seeks-ban-instagram-declare-meta-an-extremist-organization/?itid=lk_inline_manual_27

504 Parviz Mullojonov, 'Official Dushanbe silent as tajik society deeply divided on Ukraine war', RFE/RL (21 May 2022), available at: https://www.rferl.org/a/tajikistan-public-divided-war-ukraine/31861484.html

505 An independent online platform in Kyrgyzstan with the main purpose of checking and refuting false information, manipulation and propaganda. More information is available at: https://factcheck.kg/category/about/

506 'It will never be the same again: how does war in Ukraine affect people's relations in Kyrgyzstan?' Central Asian Bureau for Analytical Reporting (19 April 2022), available at: https://cabar.asia/en/it-will-never-be-the-same-again-how-does-war-in-ukraine-affect-people-s-relations-in-kyrgyzstan

507 Greg Miller and Joseph Menn, 'Putin's prewar moves against U.S. tech giants laid groundwork for crackdown on free expression', Washington Post (12 March 2022), available at: https://www.washingtonpost.com/world/2022/03/12/russia-putin-google-apple-navalny

and governments grew in prominence with the increasing interest of authorities in controlling access to and availability of information. Close cooperation has been established particularly in Kazakhstan and Uzbekistan, which have demonstrated multiple efforts to build relations and develop ties with social media platforms. In contrast, journalists, HRDs and civil society organisations do not have established contacts with private actors. At the same time, cooperation between governments and civil society is limited to programmes and training developing skills for digital transformation.[508] To correct this imbalance, tech companies and authorities alike should contribute to opening information channels in a multi-stakeholder manner and build communication bridges between state and non-state actors to meaningfully support platform governance in the region, particularly bearing in mind the consequences of the war and their far-reaching negative impacts on populations in Central Asia.

Recommendations

States are the main guarantors of human rights. To fulfil their international obligations and uphold human rights commitments protecting journalists and HRDs, governments must respect and protect freedom of expression, association and assembly online and the right to seek and impart information freely for all citizens.[509] The authorities should secure the legislative framework against abuse and exploitation for political ends, ensure the independence of internet service providers and online platforms, and refrain from instituting internet shutdowns, throttling, blocking or other means of censorship or surveillance. Facing an unprecedented amount of propaganda, misinformation, disinformation, and hateful and inciting online content, governments should strengthen the ability of journalists, HRDs and civil society to use digital technologies for communication and reporting that flags harmful content, debunks fake news and provides timely and verified information and evidence-based reporting, which is critical in times of

508 'Digital Technology Center opens in Andijan', UZ Daily (17 May 2019), available at: https://www.uzdaily.uz/en/post/49770.

509 'Reinforcing media freedom and the safety of journalists in the digital age: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan', Human Rights Council (3 June 2022), available at: https://reliefweb.int/report/world/reinforcing-media-freedom-and-safety-journalists-digital-age-report-special-rapporteur-promotion-and-protection-right-freedom-opinion-and-expression-irene-khan-ahrc5029-enarruzh

conflict.[510] Further education should be aimed at the public to promote critical thinking and media literacy as well as to improve digital security practices and raise awareness about risks associated with online spaces.

Private actors should exercise corporate responsibility, protect the rights of users on their platforms and challenge attempts that aim to limit internet freedom. Facing mounting pressure from authoritative governments, intermediary services need to preserve their core function as means for facilitating communication, receiving and imparting information, and civic organising. Social media platforms can ease political influence by increasing transparency about their operations and practices and by disclosing information regarding potential threats to user privacy and freedom of expression for people navigating their platforms. Transparent reporting contributes to the understanding of the scope and scale of online surveillance, service disruptions, content removal and other practices impacting users' rights and security. Social media providers can empower journalists and HRDs by increasing opportunities and mechanisms supporting and protecting their work. Priority areas include guaranteeing end-to-end encryption and data protection safeguards against intrusive data collection, improving fact-checking mechanisms that include trusted and verified partners in the loop of accuracy verification and labelling misinformation, and countering cyberbullying of journalists and HRDs by improving verification techniques to address the use of trolls and bots.[511] Particular attention should be given to the segments of society who are exposed to a higher amount of harassment and intimidation online, such as female and LGBTI journalists, HRDs and activists, whose needs should be considered a litmus test for online safety.[512] As an underlying effort, it is important to strengthen communication and cooperation between targeted groups from among media and civil society and private actors and to establish direct contact, to meaningfully inform the practices and advocate for improvements to platforms' features. At the same time, journalists and HRDs need to proactively seek communication channels to social media platforms, especially in regard to providing evidence-based information about the challenges they face when using online platforms in their respective countries.

510 Elira Turdubaeva, 'Media landscape in Kyrgyzstan: caught between elite capture and control of political and business interests', The Foreign Policy Centre (1 March 2021), available at: https://fpc.org.uk/media-landscape-in-kyrgyzstan-caught-between-elite-capture-and-control-of-political-and-business-interests

511 More information about the Facebook official fact-checkers is available at: https://about.fb.com/news/2018/06/hard-questions-fact-checking

512 Pavlina Pavlova, 'Human rights-based approach to cybersecurity: addressing the security risks of targeted groups', Peace Human Rights Governance 4 (3) (2020), available at: https://doi.10.14658/pupj-phrg-2020-3-4

The international community also has a role in supporting quality journalism and impartial information provided by civil society watchdogs. Likewise, relevant actors need to highlight the importance of international assistance in promoting accountability, capacity building and media literacy, and support governments in their capacity to promote digital security training for journalists and HRDs. Concurrently, the efforts of the international community should be aimed at supporting and incentivising Central Asian governments to adhere to their human rights commitments. As the countries foster a multi-vector policy and aim to decrease their reliance on Russia, democracies have a window of opportunity to coordinate their foreign policy agendas to increase their soft power in the region against the model of platform governance that is coming from Russia, China and their allies. Effective leverage can be introduced through multistakeholder cooperation that builds on transparency and accountability principles. While the coordination required for secure online communications necessitates global engagement by nation states, the multifaceted nature of this issue requires the involvement of a variety of stakeholders to provide an understanding of the gaps in current practices and exchange knowledge of their differentiated impacts on diverse communities. Outreach to journalists, HRDs and civil society to contribute to multistakeholder processes must not be only a formal exercise; their participation should be enabled and encouraged. There is a need for a clear understanding of what the outcome of their engagement is, how their participation is structured, how stakeholders are selected and what kind of support can be provided to facilitate their involvement.

While multi-stakeholder platforms such as the Internet Society and the regional and global Internet Governance Forum (IGF) provide valuable fora for discussions and raising awareness about critical issues, organisations with unique access to state actors can move the agenda forward on a political level. Regional bodies such as the OSCE representative on the freedom of the media (RFoM) and the OSCE Office for Democratic Institutions and Human Rights (ODIHR), which actively engage in the region, can help to establish meaningful communication bridges between stakeholders.[513] Their intergovernmental structure and track record of work with civil society to increase their participation in public policy issues, increase oversight capacities and facilitate issue-driven coalition building position institutions favourably to assist in establishing ties between states and civil society.

513 Human Rights Watch (see note 23 above).

Donor countries could also support multi-stakeholder initiatives through their contributions in terms of extrabudgetary funding, and thus tap into the OSCE's potential as an accessible platform for confidence-building and multi-stakeholder efforts. Similar points apply to other regional bodies, and importantly the EU, with the potential to bring diverse actors to the table and contribute to building mutual confidence and trust. All such initiatives depend on political will, which the regional elites have been lacking so far—but coupled with incentives they can be of particular importance in the current framework when the regional ties and dependencies on Russia are being re-evaluated.

Conclusion

Digital authoritarianism adopted across Central Asia created an uneasy space for journalists, HRDs and other civil society actors to conduct investigations, engage with their audiences, mobilise and disseminate information online. These trends have intensified with the war in Ukraine, as political elites in the region attempt to demonstrate pragmatic impartiality or even distance their countries from Russia.[514] Parallel to the political line, states are tilting towards censorship to placate both Moscow and domestic audiences—setting internet freedom in the region on a course to deteriorate further. Since the start of Russia's invasion of Ukraine, social media platforms have taken important but ad hoc steps to counter disinformation, limit state-sponsored and state-affiliated media with ties to Russia and introduce safety measures that reinforce the protection of their users. However, progress has not occurred in a systematic way that could contribute to broader transparency and accountability measures or meaningfully retract the past concessions that private actors made to authoritarian governments. Journalists, HRDs and other targeted groups find themselves on the sharp end of platform governance without the necessary means to have a voice in forming policies and practices. Building communication bridges between states and non-state actors and including their views through a multi-stakeholder approach

514 'President Tokayev urges Russia and Ukraine to reach agreement through negotiations, says Kazakhstan ready to provide mediation, if needed', Ministry of Foreign Affairs of the Republic of Kazakhstan (1 March 2022), available at: https://www.gov.kz/memleket/entities/mfa/press/news/details/334985?lang=en; Wilder Alejandro Sánchez, 'Kazakhstan continues to break ranks with Russia,' The Diplomat (23 September 2022), available at: https://thediplomat.com/2022/09/kazakhstan-continues-to-break-ranks-with-russia; Stefan Hedlund, 'Uzbekistan's bumpy ride out of Russia's orbit', GIS Reports (24 August 2022), available at: https://www.gisreportsonline.com/r/uzbekistan-russia-relations

can facilitate important contributions in support of accessibility, integrity and confidentiality of information online, and thus create leverage against digital authoritarianism in the region and beyond.

INTERNATIONAL LAW AND HUMAN RIGHTS PERSPECTIVES

CHAPTER 11

Pulling the strings in cyberspace

Legal attribution of cyber operations based on state control

EVGENI MOYAKINE

'States and individuals look to the power of international law to regulate cyberspace, deter and suppress unwanted or injurious cyber activities and hold those responsible to account. The institution of responsibility is at the heart of international law and is part of the constitution of the international community. This being said, cyberspace poses numerous challenges to international law's central objective of ensuring responsibility.'[515]

515 Russell Buchan and Nicholas Tsagourias, 'Special Issue: Non-State Actors and Responsibility in Cyberspace: State Responsibility, Individual Criminal Responsibility and Issues of Evidence', Journal of Conflict & Security Law 21 (3) (2016), 371–381: 377.

Introduction

Nowadays, both state and non-state actors have various technical tools at their disposal that can be used in theory, and are used in practice, to conduct operations in cyberspace with the aim of harming interests of other states and inter alia causing damage to their infrastructures.[516] These so-called 'cyber operations' (COs) are a notion that has a broader scope than the terms 'cyber warfare'[517] or 'cyber war' often used in the literature. This notion is used throughout the Tallinn Manual 2.0 that has been drafted by a group of highly qualified legal experts and clarifies how international law applies to cyberspace.[518] COs are carried out not only in the context of armed conflicts but also in times of peace, and fall under different legal regimes, including humanitarian and human rights law.[519] They are defined as operations that involve the employment of capabilities aimed at achieving certain objectives in or through cyberspace.[520] 'Cyberspace' can be understood as both the physical and non-physical domain consisting in part or in whole of some essential components such as computer systems, communications networks, software, digital information including content and traffic data, and persons and entities using these data.[521]

A series of distributed denial-of-service attacks were carried out in January 2022 by hackers allegedly affiliated with the Russian government against Ukraine.[522] In April 2022—15 years after the digital attacks[523] on Estonia that,

516 Samuli Haataja, Cyber Attacks and International Law on the Use of Force: The Turn to Information Ethics. Emerging Technologies, Ethics and International Affairs (Abingdon: Routledge, 2019), 2, 51.

517 Michael Gervais, 'Cyber attacks and the laws of war', Journal of Law & Cyber Warfare 1 (1) (2012), 17–24.

518 Michael N. Schmitt (ed.), Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations (Cambridge: Cambridge University Press, 2017).

519 Paul Ducheine, Joop Voetelink, Jan Stinissen and Terry Gill, 'Towards a legal framework for military cyber operations', in Paul Ducheine, Frans Osinga and Joseph Soeters (eds), Cyber Warfare: Critical Perspectives (The Hague: T.M.C. Asser Press, 2012), 111–113.

520 See the definition in Rule 20 in Yoram Dinstein and Arne Willy Dahl, Oslo Manual on Select Topics of the Law of Armed Conflict: Rules and Commentary (Cham: Springer, 2020), 19; US Chairman of the Joint Chiefs of Staff, Joint Publication 3-0: Joint Operations, 17 January 2017 (Incorporating Change 1, 22 October 2018), GL-8, available at: https://irp.fas.org/doddir/dod/jp3_0.pdf

521 Government of Israel, Advancing National Cyberspace Capabilities: Resolution No. 3611 of the Government of August 7, 2011 (unofficial translation), available at: https://nsarchive2.gwu.edu/dc.html?doc=3346587-Document-05-Government-of-Israel-Resolution-No

522 Luke Harding, 'Ukraine hit by "massive" cyber-attack on government websites', The Guardian (14 January 2022), available at: https://www.theguardian.com/world/2022/jan/14/ukraine-massive-cyber-attack-government-websites-suspected-russian-hackers

523 James Pamment et al., 2007 Cyber Attacks on Estonia (Riga: NATO Strategic Communications Centre of Excellence, 2019), available at: https://stratcomcoe.org/cuploads/pfiles/cyber_attacks_estonia.pdf

among other things, led to the adoption of the abovementioned manual—it was revealed that the hackers' group Sandstorm (often linked to the Russian military intelligence agency) launched a cyber-attack on Ukrainian energy facilities but had been prevented from causing damage.[524] This and similar COs are in certain instances perceived as the use of force, which is prohibited under international law, while in other scenarios they fall below this threshold and land in a grey, unexplored zone between war and peace. Importantly, the practice of launching cyber-attacks reveals that states operating in the international arena may use proxies and violate international law without facing responsibility for malicious cyber actors' activities, while at the same time achieving their strategic objectives and advancing their interests.[525] In other words, states are capable of acting as 'puppeteers', hiding behind their 'puppets' in the dark of cyberspace and escaping any form of responsibility.[526]

The current contribution investigates the issue of state responsibility for the involvement of states in COs by approaching this subject from the angle of legal attribution. More particularly, it explores the level of state control that is required for attributing operations of private actors posing cyber threats,[527] such as individual hackers and hackers' groups, to states exercising control over them. This study proposes to take into account customary international law that does not prescribe a specific test for examining the degree of state control in the physical world or in cyberspace in every single instance of determining whether an internationally wrongful act has been committed. It argues that there is a need to reconsider the application of the existing control theories in the light of the current realities in order to assess the possibility of legal attribution and assigning responsibility to states exercising varying degrees of control. In addition, it uses the proposed approach for determining the level of state control by applying it to the facts of the case study on the infamous Stuxnet attack involving state actors, which could be seen as a violation of international law.

The next section of this chapter gives a brief overview of the doctrine of state responsibility in relation to COs and touches on a number of attribution modes.

524 Ryan Gallagher, 'Russian hackers tried damaging power equipment, Ukraine says', Bloomberg (12 April 2022), available at: https://www.bloomberg.com/news/articles/2022-04-12/russian-hackers-tried-damaging-power-equipment-ukraine-says?srnd=technology-vp

525 Tim Maurer, Cyber Mercenaries: The State, Hackers, and Power (Cambridge: Cambridge University Press, 2018), 22–25.

526 Rebecca Crootof, 'International cybertorts: expanding state accountability in cyberspace', Cornell Law Review 103 (3) (2018), 565, 569.

527 Ducheine et al. (see note 5 above), 106.

The section following it deals with the control theories that have been articulated by two international judicial bodies: the International Court of Justice (ICJ) and the International Criminal Tribunal for the former Yugoslavia (ICTY). In addition, it explains how the question of control should be approached by using the two abovementioned control tests. Then, the findings of the previous sections are used to shed light on the issue of attribution in the case study conducted on Stuxnet. The final section provides a conclusion.

State responsibility

First and foremost, cyberspace does not exist in a legal vacuum and international law regulates activities taking place in this 'fifth domain' of warfare existing alongside the domains of sea, land, air and space.[528] This means that secondary rules of international law—the law of state responsibility—find application in operations originating in cyberspace, in addition to primary rules of numerous treaties and customary international law.[529] International responsibility rules and principles had been laid down in the (Draft) Articles on the Responsibility of States for Internationally Wrongful Acts (ARSIWA) by the International Law Commission, and the UN General Assembly took note of this important document in its resolution.[530] The commission of such acts by states triggers their international responsibility, as indicated by the basic principle of Article 1 ARSIWA, which has been widely applied by international judicial bodies. As a result, new legal relationships between states arise under international law and bring about certain legal consequences, such as an obligation of making reparations for the damage incurred. An internationally wrongful act of a state has been committed if conduct is attributable to the state in question under international law and constitutes a breach of an international obligation of that state, according to

528 Schmitt (see note 4 above), 3, 12; United Nations, Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, UN Doc. A/70/174 (New York: UN Headquarters, 2015), 12–13; Peter Margulies, 'Sovereignty and cyber attacks: technology's challenge to the law of state responsibility', Melbourne Journal of International Law 14 (2) (2013), 496–521: 505; Ducheine et al. (see note 5 above), 104.

529 Nicholas Tsagourias, 'Non-state actors, ungoverned spaces and international responsibility for cyber acts', Journal of Conflict and Security Law 21 (3) (2016), 455–474: 461.

530 United Nations General Assembly (UNGA), Report of the International Law Commission – Fifty-Third Session, Fifty-Sixth Session, Supplement No. 10, UN Doc. A/56/10; UNGA, Resolution Adopted by the General Assembly on January 28, 2002 – Responsibility of States for Internationally Wrongful Acts, Fifty-Sixth Session, UN Doc. A/56/83.

Article 2 ARSIWA. On the one hand, attribution—the subjective element—means that an action or omission of an individual or a group of individuals through whom states act must be attached to a state as a subject of international law, so that it is perceived as a state act.[531] On the other hand, a breach as the objective element entails a violation of international law consisting of not only treaty but also non-treaty state obligations in which a state engages, such as using force contrary to Article 2(4) UN Charter.[532]

Attribution is a multi-layered process including different dimensions.[533] Firstly, there is technical attribution that is far from unproblematic due to inter alia the use of botnets and IP spoofing and is concerned with identifying the actual perpetrators of COs by using a variety of technical tools and techniques in the context of forensic investigations.[534] Secondly, political attribution deals with connecting harmful activities in the cyber domain to states or entities associated with them. It is performed at the political level, where political consequences of COs and possible retaliation play the core role.[535] Political attribution does not by definition carry the intention of holding other states responsible.[536] This is what legal attribution is used for: assigning international responsibility to states that are to a certain degree involved in COs.[537] It is obvious that the process of legally attributing malicious cyber activities to states is dependent on the output of technical attribution in the form of forensic evidence needed for establishing who the involved states and non-state actors are and what their specific relationships are. Legal attribution is also closely linked to public attribution and both may form part of states' attribution efforts, having not only technical and operational but also strategic levels and resulting in communication to the public, politicians and others.[538] In general, states cannot be held responsible for the conduct of

531 UNGA, Report of the International Law Commission – Fifty-Third Session, 35, paras 4–6.

532 Ibid., paras 7–8.

533 Nicholas Tsagourias, 'Cyber attacks, self-defence and the problem of attribution', Journal of Conflict and Security Law 17 (2) (2012), 229–244: 233.

534 Heather Harrison Dinniss, Cyber Warfare and the Laws of War (Cambridge: Cambridge University Press, 2012), 99–102; Nicholas Tsagourias and Michael Farrell, 'Cyber attribution: technical and legal approaches and challenges', European Journal of International Law 31 (3) (2020), 941–976: 942; Marco Roscini, Cyber Operations and the Use of Force in International Law (Oxford: Oxford University Press, 2014), 33; Jack Goldsmith, 'How cyber changes the laws of war', European Journal of International Law 24 (1) (2013), 129–138: 135; Tsagourias (see note 19 above), 233.

535 Jon R. Lindsay, 'Stuxnet and the limits of cyber warfare', Security Studies 22 (3) (2013), 365–404: 398.

536 Tsagourias and Farrell (see note 20 above), 946.

537 Margulies (see note 14 above), 504; Tsagourias and Farrell (see note 20 above), 946.

538 Thomas Rid and Ben Buchanan, 'Attributing cyber attacks', Journal of Strategic Studies 38 (1–2) (2015), 4–37: 8–11, 14–30, 34.

non-state actors, and only the conduct that is attributable to them triggers their international responsibility for the international law breaches committed. Legal attribution or imputation in the law of state responsibility is in essence 'the operation of attaching a given action or omission to a State'.[539]

There are four modes of legal attribution derived from customary international law that are of utmost importance to the engagement of states in COs. The first attribution mode, found in Article 4 ARSIWA, concerns the conduct of state organs, such as the armed forces, constituting state acts on the basis of international law.[540] In this regard, functions exercised by state organs are irrelevant, similarly to their positions within the state machinery.[541] What is of relevance is that the status of a state organ must be conferred on a person or an entity by the national law of that state. The second attribution modality is Article 5 ARSIWA, dealing with the conduct of persons or entities that are empowered by the domestic law of a state to exercise elements of the governmental authority and actually act in that capacity. If semi-public entities, private companies and others are involved in COs, are tasked by the internal law of a state with carrying out governmental functions and act in breach of international obligations of that state, the latter can be held responsible.[542] The third applicable type of attribution is Article 11 ARSIWA, dealing with situations when non-state actors' actions or omissions are acknowledged and adopted by a state as its own.[543] This should not be a sole support of such conduct but a state is required to consider it as its own and, for instance, take steps to ensure that the conduct does not stop.[544] Clearly, in practice states would rarely go as far as to use their official organs and individuals or entities exercising governmental authority to directly conduct COs violating international law or to acknowledge and adopt such operations as their own. If they do, however, the attribution grounds presented in Articles 4, 5 and 11 ARSIWA will be expected to be applicable and the application of these

539 UNGA (see note 17 above), 36, para. 12.

540 Schmitt (see note 4 above), 87–9, paras 1–7; Michael N. Schmitt and Sean Watts, 'Beyond state-centrism: international law and non-state actors in cyberspace', Journal of Conflict and Security Law 21 (3) (2016), 595–611: 603; Roscini (see note 20 above), 34; François Delerue, Cyber Operations and International Law, Cambridge Studies in International and Comparative Law, 146 (Cambridge: Cambridge University Press, 2020), 115.

541 James Crawford, State Responsibility: The General Part, Cambridge Studies in International and Comparative Law, 100 (Cambridge: Cambridge University Press, 2013), 118–124.

542 UNGA (see note 17 above), 43, paras 2–4; Crawford (see note 27 above), 126–132; Schmitt (see note 4 above), 89–90, paras 8–11; Schmitt and Watts (see note 26 above), 603; Roscini (see note 20 above), 35; Delerue (see note 26 above), 124.

543 Crawford (see note 27 above), 181–188; Schmitt (see note 4 above), paras 15–18; Roscini (see note 20 above), 39–40.

544 Schmitt and Watts (see note 26 above), 605.

provisions in cyberspace will be less problematic than the final and most significant mode of attribution laid down in Article 8 ARSIWA.[545] It is about the conduct of non-state actors operating on instructions of states or under their direction and control.[546] In the case of instructions, intentions of a state must be manifestly indicated with respect to international law violations that are authorised by it.[547] With respect to direction, the commission of those violations must be directed by a state providing instructions to non-state actors in a constant manner.[548] It is difficult to prove that instructions have been given or that direction has taken place.[549] State control appears to be a more practically relevant manner of legally attributing private conduct under Article 8 ARSIWA, which at the same time also poses a number of legal problems, to be tackled below.

Control theories in the age of cyber

Delving into the question of state control exercised over COs and actors participating in them as the most valuable type of legal attribution outlined in Article 8 ARSIWA requires elaboration of two control theories: the effective control test put forward by the ICJ and the ICTY's overall control test. Both theories are discussed in the context of state responsibility in the original Tallinn Manual, but the 2.0 version mentions only the ICJ's perspective and its approach as seemingly the leading theory for attribution on the basis of control without much further explanation.[550]

In its judgment in the Nicaragua case, the ICJ held that the conduct of contras would trigger international responsibility of the US if it was proved that the state had effectively controlled the military or paramilitary operations during which

545 Kubo Mačák, '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 21 (3) (2016), 405–428: 426; Delerue (see note 26 above), 150; Schmitt (see note 4 above), 95–99, paras 1–14.

546 Crawford (see note 27 above), 144.

547 United States Diplomatic and Consular Staff in Tehran (United States of America v. Iran), ICJ Reports 1980 (ICJ, 24 May 1980) (Tehran Hostages judgment), 30–31, para. 59; Schmitt and Watts (see note 26 above), 604.

548 Mačák (see note 31 above), 418.

549 Tsagourias (see note 15 above), 472.

550 Michael N. Schmitt (ed.), Tallinn Manual on the International Law Applicable to Cyber Warfare (Cambridge: Cambridge University Press, 2013), 32–33, para. 10; Schmitt (see note 4 above), 96, para. 6.

international law violations had taken place.[551] Involvement of the state in the form of financing, equipping or providing training to the contras, participation in the operational planning and general state control would not be sufficient to conclude that the perpetration of acts contrary to international law was directed or enforced by that state. The 'effective control' test has a significantly high threshold[552] and is hard to satisfy in practice: it requires a state to effectively determine how COs breaching its international law obligations are to be conducted and to monitor their execution on a continuous basis.[553] In the judgment in the Bosnian Genocide case, the ICJ stated that the effective control theory is the one to be followed for the purposes of Article 8 ARSIWA rather than the easier to meet 'overall control' test enunciated by the ICTY in the Tadić decision.[554] Although the discussion might seem to have ended there, much can be said about the necessity for applying a control theory with a lower threshold that would not allow states to escape international responsibility by using non-state actors controlled by them.

It follows from the reasoning of the ICTY's Appeals Chamber that the required level of state control over individuals and groups of individuals is different in nature: this idea was supported by the vice-president of the ICJ, Awn Shawkat Al-Khasawneh, in the Bosnian Genocide case, who drew attention to the fact that 'the test of control is a variable one'.[555] In this regard, two factual scenarios should be distinguished in cyberspace: (1) private individuals or unorganised groups of individuals conducting COs and engaging in violations of international law on behalf of states;[556] (2) organised and hierarchically structured groups of individuals committing those violations.[557] The first situation would require the degree of state control over COs to be effective in order to make legal attribution possible. For the second scenario, state control of an overall character over

551 Case Concerning Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America), ICJ Reports 1986 (ICJ, 27 June 1986) (Nicaragua judgment), 54–55, para. 115.

552 Delerue (see note 26 above), 130, 134.

553 Scott J. Shackelford and Richard B. Andres, 'State responsibility for cyber attacks: competing standards for a growing problem', Georgetown Journal of International Law 42 (4) (2011), 971–1017: 987–988.

554 Case Concerning the Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro), ICJ Reports 2007 (ICJ, 26 February 2007) (Bosnian Genocide judgment), 168–9, para. 398; Prosecutor v. Tadić, ICTY-IT-94-1-A (ICTY, 15 July 1999) (Tadić judgment); Delerue (see note 26 above), 141.

555 Tadić judgment, 47–8, para. 117; Dissenting Opinion of the Vice-President of the ICJ A. S. Al-Khasawneh in Bosnian Genocide judgment (Dissenting opinion of Al-Khasawneh, Bosnian Genocide judgment), 216–217, para. 37.

556 Tadić judgment, 48, para. 118.

557 Ibid., 49, para. 120.

cyber groups would be sufficient, meaning that states are required not only to have equipped and financed them, but also to have coordinated their activities or helped in their general planning.[558] According to the judges of the ICTY, control exercised by a state over non-state actors, such as militias or paramilitary groups, can have an overall nature, which does not require those actors to operate under specific orders or direction of a state and, more explicitly, means that a state must have played 'a role in organising, coordinating or planning the military actions of the military group, in addition to financing, training and equipping or providing operational support to that group'.[559]

It is worth noting that attribution rules of the ARSIWA concerning state control do not prescribe one specific control theory to be used in different contexts.[560] Relying on the hard-to-meet criteria of the effective control test having a high threshold would not be in conformity with the basic premise of the state responsibility doctrine, allowing states to escape international responsibility.[561] Moreover, it would be rather unrealistic to attempt to hold states to account if one merely applied the effective control requirements, especially in the current day and age. An argument can be made that synchronously with numerous technological developments taking place nowadays, the law of international responsibility, including attribution rules, which is based in customary international law, is constantly evolving and should not be considered static. The adoption of more flexible attribution standards can, for instance, be observed in the area of anti-terrorism.[562]

By acknowledging the widespread reliance of states on ICT and rethinking the existing international law rules and principles, the focus must be put back on the overall control theory. This control test is supported by state and judicial practice,[563] has a lower threshold[564] than the effective control theory and can be useful in establishing the actual close relationship between states and organised and hierarchically structured groups acting in cyberspace. Such groups must,

558 Ibid., 56, para. 131; Roscini (see note 20 above), 37.

559 Tadić judgment, 58–59, para. 137.

560 Crawford (see note 27 above), 147.

561 William Banks, 'State responsibility and attribution of cyber intrusions after Tallinn 2.0', Texas Law Review 95 (7) (2017), 1511; Margulies (see note 14 above), 500; Antonio Cassese, 'The Nicaragua and Tadić tests revisited in light of the ICJ judgment on genocide in Bosnia', European Journal of International Law 18 (4) (2007), 649–668: 654; Tadić judgment, 49–50, para. 121.

562 Derek Jinks, 'State responsibility for the acts of private armed groups', Chicago Journal of International Law 4 (1) (2003), 88–9.

563 Tadić judgment, 51–62, paras 124–145.

564 Delerue (see note 26 above), 134.

however, have a structure, a chain of command, a set of rules of operation and certain outward symbols of authority, as stated by the ICTY.[565] It should not be forgotten that the Tadić decision was rendered more than two decades ago, when cyber-attacks were not a widespread threat,[566] which is a lifetime in the realm of technology: therefore, it could be argued that the elements of the overall control test should be re-evaluated. Currently, not only typical armed groups but also hackers' collectives and other groups involved in COs can be highly organised and have hierarchical structures.

A structure that a group has does not need to be of military nature. Arguably, a degree of organisation making group members understand their position and function within a group would be sufficient. In addition, a well-structured group can only be functional if there is a chain of command, meaning that individual members should not be operating on their own but must be subjected to the authority of their superiors, who are expected to steer the group's activities.[567] A set of rules according to which a group functions implies not only written but also unwritten rules that must be observed by its members. These three criteria could possibly be met by the groups of individuals acting on behalf of states in cyberspace, but it is also the case that hackers and others using advanced cyber tools to cause harm and destruction cannot easily be distinguished from civilians, given that this distinction is certainly more blurred in the context of COs. The fourth requirement, of outward symbols of authority, is more relevant to the engagement of military and paramilitary units in armed conflicts who wear special uniforms, carry weapons and are distinguishable from civilians, and should not be regarded as the core one. This follows from the reasoning of the ICTY referring to the Stephens case, in which the Mexico–United States General Claims Commission used the overall control test despite the fact that the irregular armed group had not had uniforms and insignia.[568] Logically, it is to be observed that groups carrying out COs on behalf of states can meet the identified criteria.

In conclusion, both the effective and overall control theories are useful instruments in determining the extent of control exercised by states in relation to cyber non-state actors and their operations and in attributing breaches of international

565 Tadić judgment, 49, para. 120.

566 Collin Allan, 'Attribution issues in cyberspace', Chicago-Kent Journal of International and Comparative Law 13 (2) (2013), 60.

567 Tadić judgment, 49, para. 120.

568 United States v. Mexico, Reports of International Arbitral Awards (vol. IV) (Mexico-United States General Claims Commission, 15 July 1927) (Stephens case), 266–267, paras 4–7; Tadić judgment, 51, para. 125.

law to those states on the basis of Article 8 ARSIWA. In the light of recent technological advances, they are not solely applicable to military operations and should be resorted to in the cyber context depending on the specific circumstances of a case in which individuals and groups might be involved.

The Stuxnet incident

In relations between themselves, states have yet to invoke international responsibility for malicious cyber activities. This could be explained by many uncertainties regarding not only their general unwillingness to prevent further conflicts from escalating but also the application of international law to cyberspace and legal attribution.[569] It is, however, to be expected that a cyber dispute will be adjudicated in the (near) future on the basis of international law[570] and therefore it is reasonable to apply the above considerations to a well-documented real-life scenario and to determine whether and to what extent legal attribution of COs constituting violations of state obligations could be established.

The deployment of the malicious Stuxnet worm against Iran has been frequently referred to as the use of the world's first cyber weapon.[571] It has led to the realisation that COs amounting to cyber warfare are a major threat to any state around the globe and can have significant consequences: it is true that many states develop and use similar offensive cyber capabilities.[572] According to the statements by anonymous US, European and Israeli officials, the US and Israel

569 Tsagourias and Farrell (see note 20 above), 946; Banks (see note 47 above), 1510; Jody M. Prescott, 'The law of armed conflict and the responsible cyber commander', Vermont Law Review 38 (1) (2013), 103–140: 104.

570 Possibly before the ICJ: see Delerue (note 26 above), 146.

571 Brad Jones, 'How Stuxnet, the first weapons-grade malware, kicked off a cyber arms race,' Digital Trends (3 February 2016), available at: https://www.digitaltrends.com/computing/the-legacy-of-stuxnet/; Kim Zetter, Countdown to zero day: Stuxnet and the launch of the world's first digital weapon (New York: Crown Publishers, 2014); Goldsmith (see note 20 above), 138; Ralph Langner, 'Stuxnet: dissecting a cyberwarfare weapon', IEEE Security and Privacy 9 (3) (2011), 49–51; Holger Stark, 'Mossad's miracle weapon: Stuxnet virus opens new era of cyber war', Spiegel Online (8 August 2011), available at: https://www.spiegel.de/international/world/mossads-miracle-weapon-stuxnet-virus-opens-new-era-of-cyber-war-a-778912.html; Harrison Dinniss (see note 20 above), 2.

572 Brandon Valeriano and Benjamin Jensen, 'The Myth of the Cyber Offense: The Case for Restraint,' CATO Institute Policy Analysis no. 862 (15 January 2019), available at: https://www.cato.org/policy-analysis/myth-cyber-offense-case-restraint

were jointly involved in the Stuxnet operation.[573] With an aim of sabotaging the Iranian uranium enrichment programme, a covert project code-named 'Olympic Games' was authorised by the Bush administration and further advanced during the Obama administration.[574] Israel also made preparations for this covert action against Iran and, reportedly, carried out a number of tests before the actual deployment of Stuxnet.[575]

In June 2009, the computer worm of around 500 KB (packed with Ultimate Packer for eXecutables (UPX)) and approximately 1.2 MB (unpacked))[576] had infiltrated the computer systems located at the Natanz nuclear enrichment plant via a USB stick and spread via local area networks.[577] Described as 'an unprecedently masterful and malicious piece of code', Stuxnet operated in three phases: (1) attacking machines running Microsoft Windows and replicating itself through computer networks; (2) looking for Siemens Step7 software required for programming industrial control systems; and (3) opening access to the programmable logic controllers, allowing the attackers to damage a significant part of the industrial machinery by altering the rotation speed of the centrifuges.[578] As a computer worm of high sophistication, Stuxnet had exploited four Windows zero-day vulnerabilities in a fascinating manner and was able to effectively intervene in the functioning of the centrifuges without revealing itself until it was discovered in June 2010 by a small IT security company from Belarus.[579]

It appears that the development and deployment of the Stuxnet worm were certainly not an amateur work of 'a ragtag group of black-hat hackers'.[580] It was an effort possibly undertaken by a large group of skilled professionals who had a

573 Ellen Nakashima and Joby Warrick, 'Stuxnet was work of U.S. and Israeli experts, officials say', Washington Post (2 June 2012), available at: https://www.washingtonpost.com/world/national-security/stuxnet-was-work-of-us-and-israeli-experts-officials-say/2012/06/01/gJQAlnEy6U_story.html; David E. Sanger, 'Obama order sped up wave of cyberattacks against Iran,' New York Times (1 June 2012), available at: https://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html; Delerue (see note 26 above), 154.

574 Nakashima and Warrick (see note 59 above); William J. Broad, John Markoff and David E. Sanger, 'Israeli test on worm called crucial in Iran nuclear delay', New York Times (15 January 2011), available at: https://nyti.ms/2jR5DSA; David E. Sanger, 'U.S. rejected aid for Israeli raid on Iranian nuclear site', New York Times (10 January 2009), available at: https://nyti.ms/2qnVWfy

575 Broad et al. (see note 60 above).

576 Marco De Falco, Stuxnet Facts Report: A Technical and Strategic Analysis, Report (CCDCOE, 2012), p. 2, available at: https://ccdcoe.org/library/publications/stuxnet-facts-report-a-technical-and-strategic-analysis-2/

577 David Kushner, 'The real story of Stuxnet', IEEE Spectrum 50 (3) (2013), 49–50, available at: https://doi.org/10.1109/MSPEC.2013.6471059

578 Ibid., 50.

579 Ibid., 51; Kim Zetter, 'How digital detectives deciphered Stuxnet, the most menacing malware in history', Wired (7 November 2011), available at: https://www.wired.com/2011/07/how-digital-detectives-deciphered-stuxnet/

580 Kushner (see note 63 above), 51.

specific aim of targeting the nuclear facilities of Iran: 10 people would need approximately two or three years to create it.[581] Due to the complex architecture of the worm, its clear destructive intent directed towards the Iranian nuclear plants and its development costs, many experts tend to believe that it could not have been done without involvement of a government.[582] While it is not known who exactly initiated the Stuxnet infection, the US and Israeli governments are often claimed to be behind this attack.[583] The question is, however: what was the degree of their involvement?

Before looking into the question of attribution, one should establish whether an internationally wrongful act has been committed in the Stuxnet incident. For this particular exercise, the order of elements of an internationally wrongful act—attribution and a breach of international law—can be reversed for the sake of clarity, as was done in the ICJ's Bosnian Genocide judgment.[584] It could be argued that the Stuxnet operation is a cyber-attack[585] constituting the use of force contrary to Article 2(4) UN Charter. This provision, having a customary international law status, prohibits threats or uses of force against territorial integrity or political independence of states or threats or uses of force inconsistent with the purposes of the United Nations.[586] In this respect, it is not important what type of 'cyber weapon' has been used. The ICJ has underlined in its advisory opinion that 'any use of force, regardless of the weapons employed' is covered by this treaty prohibition.[587] The scale and effects of the Stuxnet operation are comparable to the scale and effects of the actual use of armed force that could severely damage or destroy the uranium enrichment facility and render its centrifuges

581 Ibid.

582 Samuli Haataja and Afshin Akhtar-Khavari, 'Stuxnet and international law on the use of force: an informational approach', Cambridge International Law Journal 7 (1) (2018), 99–121: 101–2; Elies van Sliedregt, 'Command responsibility and cyberattacks', Journal of Conflict and Security Law 21 (3) (2016), 505–521: 507; 'Stuxnet: rumours increase, infections spread', Network Security 2010 (10) (2010), 1–2; Bruce Schneier, 'The story behind the Stuxnet virus', Forbes (7 October 2010), available at: https://www.forbes.com/2010/10/06/iran-nuclear-computer-technology-security-stuxnet-worm.html; 'Stuxnet may be the work of state-backed hackers', Network Security 2010 (9) (2010), 1–2.

583 Joby Warrick, 'Iran's Natanz nuclear facility recovered quickly from Stuxnet cyberattack', Washington Post (15 February 2011), available at: https://www.washingtonpost.com/world/irans-natanz-nuclear-facility-recovered-quickly-from-stuxnet-cyber-attack/2011/02/15/ABUIkoQ_story.html; Thomas M. Chen and Saeed Abu-Nimeh, 'Lessons from Stuxnet', Computer 44 (4) (2011), 93.

584 The ICJ determines first whether certain private acts are attributable to a state and only then whether they are incompatible with international obligations of that state, as, for instance, it has done in the Tehran Hostages judgment: see Tehran Hostages judgment, 29–30, para. 56; this order of the elements has been changed in the Bosnian Genocide judgment for the sake of documenting the atrocities that took place in the former Yugoslavia.

585 Van Sliedregt (see note 68 above), 508.

586 Nicaragua judgment, 89–91, paras 188–190; Schmitt (see note 4 above), 342, para. 10.

587 Legality of the Threat or Use of Nuclear Weapons, ICJ Reports 1996 (ICJ, July 8, 1996), 22, para. 39.

inoperative.[588] It could even be considered an armed attack justifying the exercise of the inherent right of states to self-defence in accordance with Article 51 UN Charter, as some Tallinn Manual 2.0 experts have also indicated.[589] If the Stuxnet attack was carried out by private actors and is attributed to certain states, these states would be considered to have acted in violation of the use of force prohibition. So far, however, states have not recognised this or other similar cyber incidents as a situation of war.[590]

The examination of the possibility of attribution in this case is a rather complicated matter that is used to illustrate the scope of the problem associated with the importance of collecting sufficient evidence to prove states' involvement in COs. Although there is not enough evidence to suggest that state organs of the US and Israel were directly involved in the development and deployment of the Stuxnet worm, there are some claims that the US National Security Agency (NSA) cooperated with Israel on the Stuxnet operation.[591] There are also many markers in the code of the worm indicating the Israeli involvement,[592] and a reference to Stuxnet was reportedly made in a video about military successes of the head of the Israel Defence Forces pointing at the participation of Israeli security forces in the operation.[593] If this is indeed the case and there is more sufficient evidence backing up these claims and going beyond the available circumstantial evidence, the main conclusion will be that this CO carried out by the US and Israeli agencies falling under the definition of state organs could be attributed to the respective states under Article 4 ARSIWA. This is, however, hard to prove.

Given that there is also no information on possible cooperation between persons and entities exercising elements of the governmental authority of the

588 Schmitt (see note 4 above), 330–331, para. 1; in addition, the non-material harm and 'informational violence' can be said to be the result of this CO, necessitating the adoption of a broader notion of the use of force under Article 2(4) UN Charter: Haataja and Akhtar-Khavari (see note 68 above), 115–121; Haataja (see note 2 above), 136–166.

589 Schmitt (see note 4 above), 341, para. 6.

590 Gary P. Corn, 'Cyber national security: navigating gray-zone challenges in and through cyberspace', in Winston S. Williams and Christopher M. Ford (eds), Complex Battlespaces: The Law of Armed Conflict and the Dynamics of Modern Warfare (Oxford: Oxford University Press, 2019), 405.

591 Nakashima, and Warrick (see note 59 above); David E. Sanger, 'U.S. blames China's military directly for cyberattacks,' New York Times, 6 May 2013, available at: https://www.nytimes.com/2013/05/07/world/asia/us-accuses-chinas-military-in-cyberattacks.html; Thomas Brewster, 'NSA hacked? "Shadow Brokers" crew claims compromise of surveillance op', Forbes, 15 August 2016, available at: https://www.forbes.com/sites/thomasbrewster/2016/08/15/nsa-hacked-shadow-brokers-equation-group-leak; 'Edward Snowden interview: the NSA and its willing helpers', Spiegel Online, 8 July 2013, available at: https://www.spiegel.de/international/world/interview-with-whistleblower-edward-snowden-on-global-spying-a-910006.html

592 Schneier (see note 68 above).

593 Christopher Williams, 'Israel video shows Stuxnet as one of its successes,' Telegraph (15 February 2011), available at: https://www.telegraph.co.uk/news/worldnews/middleeast/israel/8326387/Israel-video-shows-Stuxnet-as-one-of-its-successes.html

US and Israel in the Stuxnet operation and the participation in it has not been acknowledged or specifically approved by any state, as required by Articles 5 and 11 ARSIWA respectively, particular attention should be devoted to Article 8 ARSIWA and it is to be established whether certain individuals or groups were acting on instructions, under direction or control of these states. There are no indications or clear evidence of any instructions or direction provided by the allegedly involved states. Nevertheless, the possibility exists that third parties were engaged in the process of developing and deploying the Stuxnet worm and that they were operating under state control. According to researchers, it does not seem probable that such complex project could have been the result of the efforts of a 'lone wolf' hacker.[594] Given the characteristics of the Stuxnet infection, it is also difficult to imagine that an unorganised group of individuals could have successfully carried out this attack, because the development of such malware requires a high level of coordination and cooperation. As elaborated on in the previous section, legal attribution in these scenarios would be possible if the involved states effectively controlled the operation in question, which is a high threshold to meet.

Considering Article 8 ARSIWA and bearing in mind the technically complex and innovative nature of Stuxnet, it is reasonable to assume that—rather than isolated hackers or loose bands of hackers—an organised and hierarchically structured group or various groups of this nature might have acted on behalf of the respective governments: at this point, it is certainly a speculation, albeit one with roots in the facts of other cyber-incidents and theories formulated by experts, linking the development of this malware to cybercrime/cyberattack groups.[595] Under these circumstances, overall control would be the appropriate test to determine whether the two states played a role in organising, coordinating or planning of the CO in addition to their involvement in financing, training and equipping or providing operational support to the group(s). While sufficient evidence to make sound conclusions on this matter is not at hand, a claim

594 Patrick Fitzgerald, 'The hackers behind Stuxnet', Symantec Official Blog (21 July 2010), available at: https://www.symantec.com/connect/blogs/hackers-behind-stuxnet

595 James P. Farwell and Rafal Rohozinski, 'Stuxnet and the future of cyber war', Survival 53 (1) (2011), 26–27, available at: https://doi.org/10.1080/00396338.2011.555586; Marie Baezner and Patrice Robin, Hotspot Analysis: Stuxnet (Zürich: Center for Security Studies (CSS), 2017), 8–9, available at: https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Cyber-Reports-2017-04.pdf; Kaspersky Lab, Equation Group: Questions and Answers (Moscow: Kaspersky Lab HQ, 2015), available at: https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/08064459/Equation_group_questions_and_answers.pdf; Charlie Osborne, 'Beyond Stuxnet and Flame: Equation "most advanced" cybercriminal gang recorded', ZDNet (16 February 2015), available at: https://www.zdnet.com/article/beyond-stuxnet-and-flame-equation-group-most-advanced-cybercriminal-gang-recorded/

could be advanced that the US and Israel organised, coordinated and planned the Stuxnet worm attack and also financed and possibly equipped the hired hackers. A demonstration of the 'unity of goals, unity of ethnicity and a common ideology'[596] could be an indication of this. For instance, it has been claimed that the Equation Group—a sophisticated threat actor— worked closely with the NSA as a team of developers behind the Stuxnet malware or assisted the Stuxnet developers in their efforts.[597] The NSA probably not only organised, coordinated and planned the actions of this organised and highly structured group, but also paid them for the delivered work and even supplied them with certain tools, such as the advanced keylogger 'Grok' developed by the NSA.[598] There are also reasons to believe that operational support had to be provided to that and other hackers' collectives given the high level of sophistication of the execution of this reportedly joint operation. If this was indeed the case and the US and Israel had overall control in the CO, while Articles 4, 5 and 11 ARSIWA were not applicable, attribution per Article 8 ARSIWA would be possible, leading to the international responsibility of both states.[599] Moreover, according to Articles 16 and 17 ARSIWA, other states aiding or assisting the US and Israel in the attack or directing and controlling them would also face responsibility if they were aware of the circumstances of the internationally wrongful act, and the attack could be qualified as such act if conducted by them.[600]

Similarly to the imputation on the basis of overall control exercised by the involved states in the Stuxnet case, the attribution of COs that could be categorised as violations of international law in other scenarios—with regard to the prohibition of not only the use of force but also intervention and the principle of sovereignty—will be possible if there is no indication of effective state control. For instance, the global WannaCry 2.0 malware attack from May 2017 was publicly attributed to North Korea by a number of states, including the US and the UK,

596 Dissenting opinion of Al-Khasawneh, Bosnian Genocide judgment, 216, para. 36.

597 Thomas Brewster, 'Equation = NSA? Researchers uncloak huge "American cyber arsenal"', Forbes (16 February 2015), available at: https://www.forbes.com/sites/thomasbrewster/2015/02/16/nsa-equation-cyber-tool-treasure-chest; Brandon Valeriano, Benjamin M. Jensen and Ryan C. Maness, Cyber Strategy: The Evolving Character of Power and Coercion (New York: Oxford University Press, 2018), 193–194; Kaspersky Lab (see note 81 above).

598 Dan Goodin, 'How "omnipotent" hackers tied to NSA hid for 14 years – and were found at last', Ars Technica (16 February 2015), available at: https://arstechnica.com/information-technology/2015/02/how-omnipotent-hackers-tied-to-the-nsa-hid-for-14-years-and-were-found-at-last

599 Tsagourias and Farrell (see note 20 above), 963.

600 Schmitt (see note 4 above), 100–103, paras 1–8.

several months after its discovery.[601] It has been claimed that the Lazarus Group as a hacking team allegedly sponsored by the North Korean government conducted this attack targeting public and private organisations worldwide.[602] There is no evidence of official state organs of North Korea or persons and entities of that state exercising governmental authority having accomplished this CO unaccompanied by others, and it is obvious that Pyongyang would not acknowledge and adopt this malware operation as its own. Legally proving the exercise of effective control in this particular scenario would be exceedingly difficult, if not impossible: it cannot be proved that the state in question effectively controlled the CO and was able to continuously decide how and when the related activities of the hackers should be carried out, closely monitor the execution of the operation and order its cessation at any given point in time. Although an extensive investigation into the factual circumstances of WannaCry 2.0 by the competent authorities and experts and a high level of transparency on the matter are certainly needed, it can be maintained that the abovementioned hackers' collective receives financial and other support from the North Korean government, which also organises, coordinates or plans COs together with the members of this group, which are carried out on its behalf.

Likewise, the SolarWinds hack that had targeted various governmental institutions and private organisations in the US and was revealed in December 2020 cannot be legally imputed to the Russian Federation under Articles 4, 5 and 11 ARSIWA, despite the fact that the attack has been attributed to the Russian foreign intelligence service at the political level.[603] This is the case due to the lack of credible evidence indicating that this particular CO was directly conducted by the actors specified in Articles 4 and 5 ARSIWA and the unwillingness of Moscow to consider this unlawful conduct its own. On the one hand, this CO, time and again, illustrates the complexity of imputation based on the effective control test of Article 8 ARSIWA, even if it can be established that the Russian hackers' group Cozy Bear or Nobelium affiliated with the abovementioned foreign intelligence

601 Paul Sandle, 'Britain joins U.S. in blaming North Korea for "WannaCry" attack', Reuters (19 December 2017), available at: https://www.reuters.com/article/us-usa-cyber-northkorea-britain-idUSKBN1ED1SK

602 Olivia Solon, 'WannaCry ransomware has links to North Korea, cybersecurity experts say', The Guardian (15 May 2017), available at: https://www.theguardian.com/technology/2017/may/15/wannacry-ransomware-north-korea-lazarus-group

603 Danny Palmer, 'SolarWinds: US and UK blame Russian intelligence service hackers for major cyberattack', ZDNet (15 April 2021), available at: https://www.zdnet.com/article/solarwinds-us-and-uk-blame-russian-intelligence-service-hackers-for-major-cyber-attack/

service was reportedly behind the operation.[604] On the other hand, the overall control theory offers a helpful instrument for linking this and similar operations conducted by organised and well-structured groups of individuals in the digital domain to states, if those powerful 'puppeteers' play a role in organising, coordinating or planning the activities of their skilful 'puppets', while also providing financial and other forms of support to them.

Conclusion

As explained above, there are still various pertinent issues that can be identified in the process of legally attributing COs to states under the existing control theories. The analysis of the case study of Stuxnet indicates that there is an obvious problem of evidence or—to be more precise—the lack thereof that needs to be solved in the process of technical attribution. Only then would legal experts be able to assign international responsibility to states for malicious cyber-activities. The examination has revealed, in a rather speculative manner, that the Stuxnet operation could be attributed to the US and Israel on the basis of Article 4 ARSIWA if their state organs had conducted it or on the basis of Article 8 ARSIWA and the overall control test if organised and hierarchically structured (hackers') groups had been involved in the development and deployment of the worm.

A few policy recommendations can be made following the analysis performed. The law of international responsibility should not fade into oblivion, but it is essential to develop it further in conformity with legal and technological developments shaping our physical and digital worlds. Given the rapid proliferation of cyber warfare capabilities, the reliance of states on various non-state groups—even those specialising in cyber terrorism[605]—for carrying out COs and issues associated with establishing a high degree of state control in cyberspace, the overall control theory must receive more judicial and political attention. Its requirements should be updated in accordance with the current practice of state involvement in COs, the criterion of outward authority symbols can be abolished

604 Alyza Sebenius, 'U.S., U.K. reveal code flaws abused by SolarWinds hackers', Bloomberg (7 May 2021), available at: https://www.bloomberg.com/news/articles/2021-05-07/u-s-and-u-k-release-details-on-russia-s-solarwinds-hackers

605 Thomas M. Chen, Cyberterrorism After Stuxnet (Carlisle: Strategic Studies Institute and U.S. Army War College Press, 2014); Scott J. Shackelford, Managing Cyber Attacks in International Law, Business, and Relations: In Search of Cyber Peace (Cambridge: Cambridge University Press, 2013), 17–18.

and the alignment between motivations, goals and general backgrounds of involved states and cyber actors is to be taken into account. As argued by some authors, there could be merit in developing and introducing new theories, such as 'soft control'[606] or 'working in tandem'[607] tests, although the overall control concept grounded in customary international law seems to be the more appropriate choice. Furthermore, evidence-related questions must be answered and there is perhaps a need for creating international/regional attribution mechanisms[608] based on cooperation between states and other organisations and persons with relevant expertise, including those from academia,[609] that are capable of striking a delicate balance between the secretive nature of attribution and its transparency. More transparency about national attribution efforts is in any case a welcome development that could contribute to further crystallisation of customary international law in this field.[610] Finally, other possibilities for holding states responsible, such as cyber due diligence,[611] should be explored in order to prevent international law from being violated through the omnipresent yet distant fifth domain.

606 Tsagourias and Farrell (see note 21 above), 965.

607 Allan (see note 52 above), 81.

608 Tsagourias and Farrell (see note 21 above), 959–961.

609 Florian J. Egloff and Myriam Dunn Cavelty, 'Attribution and knowledge creation assemblages in cybersecurity politics', Journal of Cybersecurity 7 (1) (2021), 10.

610 Banks (see note 47 above), 1512.

611 Eric Talbot Jensen and Sean Watts, 'A cyber duty of due diligence: gentle civilizer or crude destabilizer?', Texas Law Review 95 (7) (2017), 1565–1567; Tsagourias (see note 15 above), 466.

CHAPTER 12

Is cybersecurity the sole responsibility of states?

The concept of 'active defence' and the role of non-state actors in responsible state behaviour in cyberspace

JAIME BELLO

What is 'active defence' and why do some private actors promote its use?

It goes without saying that companies and other private organisations suffer most of the cyber-attacks around the world on a daily basis,[612] and that the traditional response to such attacks is simply to increase the means of defence, building up layers and layers of security (usually from different vendors) that, unfortunately, fail to prevent attackers from continuing to achieve their goals, which cause huge damage and loss of information and resources.

612 Ryan Manship, 'The top 6 industries at risk for cyber attacks', Red Team Secure, available at: https://www.redteamsecure.com/blog/the-top-6-industries-at-risk-for-cyber-attacks

It has been estimated that global spending on cybersecurity in 2020 amounted to more than US$130 billion,[613] with sustained annual growth over the past few years and prospects for even stronger rises. However, it seems that the total impact of cyber-attacks in 2020 in the context of the pandemic may have reached a trillion dollars,[614] including the direct damage due to business interruption (like that suffered by a global shipping company for two weeks in 2017, which cost it US$300 million[615]), loss of data or theft of intellectual property, in what has been described as 'the greatest unwilling transfer of wealth in history'.[616]

The attacked organisation may decide to report the attack to its national law enforcement authorities (or not). In most cases, however, this does not result in the arrest of the perpetrators, who continue to act with total impunity against other victims.

In cyberspace, traditional investigative action (entrusted to the public authorities) has been accused of sometimes being slow and ineffective. Despite increased training and high levels of specialisation of police forces and intelligence agencies—especially in Western countries—law enforcement is facing several problems that make it extremely difficult to crack down on cybercrime.

Firstly, the attack surface itself, i.e. the number of devices, networks and systems potentially vulnerable to a cyber-attack, has been growing exponentially in recent years. The proliferation of cloud computing, the digitisation of processes and the Internet of Things (IoT) make it impossible to chase every cyber-attack suffered by organisations. In 2016, the Cloud Evidence Group of the Council of Europe sadly concluded, among other things, that:

> "cybercrime, the number of devices, services and users (including of mobile devices and services) and with these the number of victims have reached proportions so that only a minuscule share of cybercrime or other offences involving electronic evidence will ever be

613 Gartner, 'Gartner forecasts worldwide security and risk management spending to exceed $150 billion in 2021' (17 May 2021), available at: https://www.gartner.com/en/newsroom/press-releases/2021-05-17-gartner-forecasts-worldwide-security-and-risk-managem

614 Tonya Riley, 'The Cybersecurity 202: global losses from cybercrime skyrocketed to nearly $1 trillion in 2020, new report finds', Washington Post, 7 December 2020), available at: https://www.washingtonpost.com/politics/2020/12/07/cybersecurity-202-global-losses-cybercrime-skyrocketed-nearly-1-trillion-2020/

615 Jill Leovy, 'Cyberattack cost Maersk as much as $300 million and disrupted operations for 2 weeks', Los Angeles Times (17 August 2017), available at: https://www.latimes.com/business/la-fi-maersk-cyberattack-20170817-story.html

616 General Keith Alexander (Commander US Cyber Command), 'Statement before the Senate Committee on Armed Services' (12 March 2013), available at: https://careersdocbox.com/US_Military/68063297-Statement-of-general-keith-balexander-commander-united-states-cyber-command-before-the-senate-committee-on-armedservices.html

> recorded and investigated. The vast majority of victims of cybercrime cannot expect that justice will be served."[617]

Secondly, the shortage of personnel and resources in law enforcement units fighting cybercrime means that they are forced—logically—to prioritise the security of public administrations, critical infrastructures and other essential elements in our democratic systems (such as the protection of elections against vote counting and disinformation attacks).

And thirdly, most of the attackers are located in states that do not collaborate as they should in the effective repression of such activities, that protect the attacking groups or even, in the most extreme cases, that promote and finance their actions by using them as proxies for the fulfilment of their own geopolitical objectives.

In view of these obstacles, in a context where a minimal percentage of cyber-attacks are being pursued, resulting in more and more losses, what should companies do? Should they limit themselves to withstanding attacks by maintaining a passive–defensive posture, i.e. strictly within their security perimeters?

For some, the answer to this question is 'no', and they advocate allowing organisations that are victims of cyber-attacks to take action through so-called 'active defence' measures. Although there is no consensus on what this concept should comprise, for our purposes we will consider the whole 'spectrum of proactive cybersecurity measures ranging from traditional passive defense to offensive action',[618] thus including a range from the less aggressive actions (threat intel gathering through the use of honeypots or other measures of deception) to those that are slightly more serious and outside the organisation's perimeter (botnet patching, dye-packets, beaconing), to the most damaging, considered as hacking back or cyber-retaliation, whereby the victim breaks into the attacker's system to retrieve information or cause damage.

617 Cloud Evidence Group – Council of Europe, 'Criminal justice access to electronic evidence in the cloud: recommendations for consideration by the T-CY Final report of the T-CY Cloud Evidence Group' (16 September 2016), available at: https://rm.coe.int/16806a495e

618 Center for Cyber & Homeland Security – The George Washington University, 'Into the gray zone – the private sector and active defense against cyber threats' (29 November 2016), available at: https://www.businessofgovernment.org/blog/gray-zone-private-sector-and-active-defense-against-cyber-threats

How does the active defence of private organisations affect the responsible behaviour of states in the cyber world?

The adoption of active defence measures, especially in their most aggressive modalities, is likely to engage the responsibility of the states in which the private organisations are located, not only towards their citizens (internal responsibility) but also in their relations with other sovereign states (external responsibility). The solution to the former would be articulated through criminal law and the repression of this kind of conduct in the cyber world, while the latter would involve aspects of international law that are still the subject of much discussion.

Internal responsibility: the criminal law response

Budapest Convention

There is as yet no rule in European legislation expressly authorising the use by private organisations of measures for the active defence of their computer systems. On the contrary, the criminalisation of cyber-offences (understood as those specifically linked to the cyber world or having a computer component) throughout the EU takes as its starting point the Budapest Convention.[619] Since its adoption in 2001 under the initiative of the Council of Europe, it has become the first experience of an international treaty on the subject, with the accession of numerous non-European countries including the USA, Japan, Canada, Nigeria and Morocco.

619 Council of Europe, Convention on Cybercrime (ETS No. 185), open for signature on 23 November 2001 by the member states and the non-member states that have participated in its elaboration and for accession by other non-member states, ratified by 66 countries, available at: https://rm.coe.int/1680081561

Through a harmonised description of cybercrimes, boosted by an effective mechanism of police and judicial cooperation, the Convention makes the prosecution of attackers, when they are located in the territory of one of its member states, extraordinarily effective except in very few cases.

Precisely because of this effectiveness, organisations must be very careful not to commit offences arising from the Budapest Convention (as implemented into national law) when undertaking active defence measures. Thus, the organisation that adopts these types of measures must be aware that, for example, by implementing beaconing measures, activating dye packets or even collecting advanced intelligence in the attacker's own systems, it could be committing a crime of illegal access or illegal interception, or that by adopting very aggressive measures such as sending logic bombs, patching third-party equipment or, directly, introducing malware, ransomware or any other means of destroying systems or data, it may be committing a crime of data or system interference.

In this regard, it is worth considering whether, in a case of active defence, the commission of these crimes could fall into one of the justifications provided for in the various criminal laws, such as self-defence and plea of necessity. In this sense, although its application remains doubtful and lacks precedents, it seems that in certain cases and under certain circumstances, courts may eventually accept such reasoning to exempt from liability the organisation that implements those measures.

Not surprisingly, although the text of the Convention does not expressly refer to such a possibility, its Explanatory Report[620]—approved on the same date—does seem to envisage the possible 'justified' commission of cybercrimes. Indeed, the Report not only admits the possibility of decriminalising such conduct when it can be subsumed within the traditional concepts of self-defence or plea of necessity, but also opens the door to finding 'other principles and interests' that may exclude criminal liability. Consequently, the 'legitimate' commission of such crimes would be unpunishable provided that it was covered by justifications, excuses or legal defences or other relevant principles provided for in domestic law. However, so far none of the countries that have signed and ratified the Convention seem to have invoked this possibility in order to specifically cover the use of active defence measures by the victims of a cyber-attack.

620 Council of Europe, Explanatory Report to the Convention on Cybercrime (23 November 2001), available at: https://rm.coe.int/16800cce5b

United States

In the USA, the voices calling for the legalisation of active defence measures by the private sector are increasing. Numerous authors and think-tanks are in favour of allowing organisations to defend themselves, especially to prevent the massive theft of information and intellectual property by foreign actors.

As recent experience has shown, no proceedings have been initiated so far as a result of the adoption of hacking-back measures. However, some of them could have been considered an offence under the Computer Fraud and Abuse Act (CFAA).[621] Indeed, this legislation would in fact prohibit many of the measures defined as active defence, insofar as it prevents organisations from accessing without authorisation and/or damaging systems or devices of third parties, even when these are located abroad.

In 2017, Georgia Congressman Tom Graves (Rep.) introduced a bill named 'Active Cyber Defense Certainty Act' (ACDC)[622] (H.R. 3270), which sought to amend the CFAA to decriminalise (while maintaining the civil remedies) the use of limited defence measures that can go beyond the boundaries of the systems themselves in order to monitor, identify and disrupt attackers. In this way, the ACDC sought to exclude from criminal liability private organisations leaving their networks in order to (1) establish attribution of an attack, (2) disrupt the cyber-attack without damaging third-party equipment, (3) recover or destroy information belonging to the victim, (4) monitor the attacker's behaviour, and (5) use beaconing technologies.

At the same time, the ACDC bill required the National Cyber Investigative Joint Task led by the FBI to be notified before any active defence measure was taken,[623] which would allow prior control to ensure responsible use. As a third pillar, the bill included a voluntary certification mechanism for active defence techniques and measures,[624] which would make it possible to increase the legal certainty deriving from their use, their legal compliance and the improvement of their operation.

Despite the enthusiasm with which the bill was received by some, and its public support by several bipartisan congressmen, it did not receive the necessary backing due to its evident risks, which finally led to its stalling.

---

621 Computer Fraud and Abuse Act, available at: https://www.congress.gov/bill/99th-congress/house-bill/4718/text

622 Active Cyber Defense Certainty Act, available at: https://www.congress.gov/bill/116th-congress/house-bill/3270

623 Section 5, ACDC.

624 Section 6, ACDC.

Singapore

The Republic of Singapore deserves special mention because it has the most advanced legislation in force in this area. Despite its small size, this state is one of the major financial hubs in Southeast Asia, which makes it a prime target for cybercriminals. Not surprisingly, in recent years it has suffered several major attacks, which seems to have prompted the adoption in 2018 of the 'Cybersecurity Act'.[625]

While it does not constitute a full legalisation of private active defence, the Act adopts an intermediate solution by creating a mechanism of active defence by order of the state to ensure the protection of its national critical infrastructures (especially financial ones). This means that the Minister for Home Affairs can grant to any person or organisation powers to 'access, inspect and check' any system within or outside Singapore that relates to an attack (whether the attacker's or a third party's) or even decrypt any information existing therein, whenever the minister deems it necessary for the purpose of 'preventing, detecting or countering any serious and imminent threat to (a) the provision of any essential service; or (b) the national security, defence, foreign relations, economy, public health, public safety or public order of Singapore'.[626] Persons subject to such an order would be granted immunity for such acts,[627] while failure to comply with it may be punishable by imprisonment for up to 10 years and fines of up to 50,000 Singapore dollars.[628]

For the first time, therefore, private organisations can be allowed to adopt active defence measures which, in the absence of greater precision, may include the penetration of third-party systems, even on a preventive basis (if they constitute a serious and imminent threat). However, by requiring prior instructions from the minister, the state can maintain control and limit how far private actors can go.

625 Cybersecurity Act 2018, No. 9 of 2018, Singapore Statutes Online, available at: https://sso.agc.gov.sg/Acts-Supp/9-2018

626 Section 23(1) and 23(2).

627 Section 23(3).

628 Section 23(4).

External responsibility: the rules of international law

The adoption of active defence measures (especially those closer to hack-back) by private organisations could also jeopardise or engage the external responsibility of states under international law, insofar as such actions are likely to cause damage on the territory of other states. Indeed, in 2013 a group of experts from 15 countries—within the framework of the United Nations—reached a consensus report[629] (known as 'GGE 2013') in which they proclaimed that international law (including the rules for the Responsibility of States for International Wrongful Acts[630] adopted by the International Law Commission in 2001) and the UN Charter were also applicable in cyberspace.

Even though to date neither such texts nor the subsequent interpretative efforts (such as the Tallinn Manual[631] developed by an international group of experts within the framework of NATO's CCD-COE) have expressly pronounced on the adoption of active defence measures by non-state actors, they provide us with a reasonable basis for the regulation of private active defence and allow us to point out possible answers to some of the questions analysed here.

Could the action of a non-state attacker entail the responsibility of the host state?

Certainly: although under international law the operations of private persons (non-state actors) are not in principle attributable to states, states could indeed be made responsible under some circumstances.

Firstly, this can apply when the state breaches its international obligations, either by action or by omission.[632] We must refer here to the breach of the 'due diligence principle', according to which states have the duty to ensure that their

629 United Nations, Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security: note by the Secretary-General, 24 June 2013, available at: https://digitallibrary.un.org/record/753055

630 Available at: https://legal.un.org/ilc/texts/instruments/english/draft_articles/9_6_2001.pdf

631 M. Schmitt (general editor), Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations, 2nd ed. (New York: Cambridge University Press, 2017).

632 Article 2, Rules for the Responsibility of States for International Wrongful Acts.

territory is not used to harm other states,[633] which means that failure to do so may entail the non-diligent state's responsibility, not for the attack itself but for its failure to prevent it.

Secondly, the state's responsibility can be entailed where the cyber operations are attributable to the state because carried out on its instructions or under its direction or control.[634] The concept and degree of such 'control' is crucial in this regard, since mere general support by the state to a group responsible for the action has been deemed not sufficient for such attribution, even if there is a 'preponderant or decisive [participation] in the financing, organising, training, supplying and equipping'[635] of the attackers. Finally, also attributable to the state would be cyber operations of non-state actors that are acknowledged and adopted 'as its own'.[636] In this case, the mere tacit approval of such action would not be sufficient; express endorsement by the host state would be required.

However, it is worth noting that the control and the acknowledgment of the cyber-attacks are elements that will not always be clearly proven, given the obvious tendency of states using this type of proxy to deny any connection with such groups.

Could active defence measures fall under the state's self-defence exception?

Under international law, states may exercise their inherent right to self-defence when they are the object of an armed attack (article 51 of the UN Charter).[637] Such a right is easily recognised when the attack comes from another state, but is less easy to consider and assess when it comes from a non-state actor. In any case, the exercise of states' self-defence would also require the concurrence of

633 See among others, Corfu Channel ICJ Judgment (1949), para. 22, or Tehran hostages ICJ Judgment (1980), paras 67–68.

634 Article 8, Rules for the Responsibility of States for International Wrongful Acts.

635 Nicaragua ICJ Judgment (1986), para. 115.

636 Article 11, Rules for the Responsibility of States for International Wrongful Acts.

637 Available at: https://legal.un.org/repertory/art51.shtml

the traditional elements of necessity, proportionality[638] and immediacy, which are not always easy to prove in the cyber realm.

Even if these requirements are met, ICJ case law and most scholars consider that for the retaliation to be covered under the victim state's self-defence right, the cyber-attacks of non-state actors also need to be attributable to the host state,[639] either because (i) they are carried out on the instructions of the state or under its effective direction or control, or (ii) they have been expressly recognised or adopted by the state.

Indeed, although Messerschmidt[640] has pointed out that such self-defence may even be invoked against operations carried out by non-state actors in scenarios other than (i) and (ii) above when their host states do not exercise due diligence, setting such a connection would be extremely rare and unlikely to be upheld by the international courts and fora,[641] even though the debate is currently open.[642]

638 Kimberley N. Trapp, 'Back to basics: necessity, proportionality, and the right of self-defence against non-state terrorist actors', International and Comparative Law Quarterly, 56(1), 141-156 (2007), available at: https://www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/abs/back-to-basics-necessity-proportionality-and-the-right-of-selfdefence-against-nonstate-terrorist-actors/F9260D68A2005754BA0FCCAE24E62090

639 DRC v Uganda ICJ Judgment (2005), para. 147

640 Jan E. Messerschmidt, 'Hackback: permitting retaliatory hacking by non-state actors as proportionate countermeasures to transboundary cyberharm', Columbia Journal of Transnational Law 52 (275) (2013), available at: https://ssrn.com/abstract=2309518. He argued that 'upon a state's breach of this obligation, affected states may be entitled to reciprocate by neglecting their own due diligence obligation, and allowing their victimized nationals to hackback'.

641 As was the case in Israel's rescue operation in Entebbe airport (1976) or bombings against the PLO's headquarters in Tunisia (1985), in which Israel invoked those states' failure to prevent their territory from being used as a base for terrorist operations. Such explanation was rejected by the international community in those cases but seemed to be accepted in the context of the 2001 US-led campaign in Afghanistan (against both Al Qaeda and the state apparatus, even though the former's actions were not necessarily attributable to the latter), as explained in Trapp; see note 27 above.

642 Especially after US air strikes in Syria against ISIS, as explained by Paola D. Reyes Parra, 'Self-defence against non-state actors: possibility or reality?', Revista Facultad de Jurisprudencia 9 (2021), 151–176, available at: https://www.redalyc.org/journal/6002/600266295004/html/

Balance of interests: prioritising the defence of national interests vs responsible behaviour in cyberspace

In view of the significant legal complications raised by private active defence, the question of whether to prioritise the defence of national interests—even if these are of a private nature—by allowing organisations to repress attacks by themselves where the state fails to reach, or to give preeminence to states' responsibility, strictly retaining the monopoly on the use of force in order to ensure responsible behaviour in cyberspace, remains unclear.

For centuries, the major maritime powers allowed retaliation by private individuals and organisations—privateers—against foreign targets, either other private actors or the fleets of other states. Thus, the action of privateers was legitimised by commissions or the so-called 'letter of marque', by means of which states delegated part of their traditional monopoly of force to these private fleets to repel pirates or state enemies on the open sea, allowing them to keep the booty obtained. Although this practice was abolished internationally by the Declaration of Paris[643] of 1856, some today advocate the need to apply similar principles to cyber-attacks by adopting a set of rules at an international level that would allow the implementation of active defence measures by authorised private organisations.

For the time being, however, most voices are against the recognition of a private right to adopt active defence measures, especially those that have been described as hack-back, because of the many disadvantages that would be caused by leaving the fight against cybercrime in private hands. Countries such as France have already declared themselves openly opposed to legalising these practices. Thus, in its 2018 Strategic Cyber Defense Review,[644] the General Secretariat for Defense and National Security included among its priority recommendations the 'prohibition of hack-back by private sector actors in cyberspace'.

643 Declaration Respecting Maritime Law, Paris, 16 April 1856, available at: https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Article.xsp?action=openDocument&documentId=473FCB0F41DCC63BC12563CD0051492D

644 SGDSN, 'Revue stratégique de cyberdéfense' (12 February 2018), available at: http://www.sgdsn.gouv.fr/uploads/2018/02/20180206-np-revue-cyber-public-v3.3-publication.pdf

Many voices in the US are also speaking out against hack-back. Indeed, the large technology industry, grouped around the Cybersecurity Tech Accord,[645] calls for the implementation of such measures to be forbidden and urges the cooperation of the entire sector to implement best practices in this area. In fact, this forum is behind Principle #8 of the Paris Call for Trust and Security in Cyberspace,[646] an initiative launched in 2018 which, based on nine guiding principles, 'calls for binding rules and standards to build trust in cybersecurity and further advance digitalization'. Principle #8 (No private hack-back), calls for 'steps to prevent non-State actors, including the private sector, from hacking-back, for their own purposes or those of other non-State actors'. Like most detractors of hack-back, they not only advise against it because of the already explained risk of committing similar cybercrimes to those committed by the attacker, but also express their concerns about the fact that the proliferation of this type of behaviour entails inherent problems and risks that are very difficult to circumvent, even if in some cases private organisations have greater resources than the states themselves.

Attribution

The first and foremost of the problems with hack-back practices is the difficulty of attributing an attack with certainty. Obviously, attribution techniques and the measures adopted by attackers to avoid them are based on the very functioning of the internet, whose initial design did not show much concern for security or for the ability to attribute without doubt the origin of a communication. Consequently, although investigation methods are becoming increasingly sophisticated, the use by attackers of equally sophisticated means to mask their actions (such as botnets, identity cloaks, proxies, dynamic IP assignment or onion routing) makes it extremely difficult in certain cases to attribute with a reasonable certainty the origin of an attack.

645 The Cybersecurity Tech Accord is a coalition of over 140 global technology firms committed to advancing trust and security in cyberspace. https://cybertechaccord.org/

646 French Ministry for Europe and Foreign Affairs, available at: https://pariscall.international/en/

Retaliation and escalation

Another problem often invoked is that on many occasions active defence or hack-back measures, far from ending the problem, may encourage the attacker (by considering the target to be worthwhile) or provoke an escalation of hostilities.[647] This is particularly dangerous in cases where the attackers are an integral part of a hostile state (intelligence or defence units) or are supported by it, since, as described above, an overly aggressive response can engage the responsibility of the victim state and trigger serious diplomatic and even warlike repercussions.

It could also be pointed out, in this regard, that the existence and development of a private sector dedicated to hack-back could in turn encourage cybercriminals to act on both sides of the fence, motivated by the enormous economic gains that such a market could involve.

Collateral damage

Finally, detractors draw attention to the risk of causing collateral damage[648] to third parties who have nothing to do with the attack. Attackers can often use third-party systems to camouflage their actions, use botnets of hundreds of computers or devices, or even carry out 'false flag attacks' by pretending to be another state or group. Hack-back measures in such cases, without being able to determine whether the attribution of the attack is reliable or not, can have unexpected and extremely damaging consequences on innocent victims, such as the loss of service of an entire network of cameras, the interruption of the activity of a hospital or the disruption of the transport network or the power grid of a third state.

647 Jan Ellis, 'Hack back is still wack', Rapid 7 (last updated 18 November 2021), available at: https://www.rapid7.com/blog/post/2021/08/10/hack-back-is-still-wack/

648 Martin Giles, 'Hacking back makes a comeback—but it's still a really bad idea', MIT Technology Review (1 December 2017), available at: https://www.technologyreview.com/2017/12/01/147357/hacking-back-makes-a-comeback-but-its-still-a-really-bad-idea/

A third way? Outsourcing the cyber response

Given the majority position against active defence measures (at least against those consisting in hacking back) on the one hand, and the persistence of the problem of cybercrime and its foreseeable exponential increase in the coming years on the other, we propose—on the basis of recent experiences and studies—to explore a third option in which the state, as holder of the monopoly of force, would 'subcontract' the necessary response to attacks to private organisations, through a system of authorisations that should ensure the responsible behaviour of the state in cyberspace under the standards of international law.

To this end, with the ultimate aim of guaranteeing the rule of law, as well as avoiding or minimising the pernicious effects that have been pointed out—that is, the escalation of hostilities, the attribution problems and the damage to innocent third parties—the implementation of such a third way should be strictly subject to the following requirements.

Public enabling scheme and enhanced fluidity of public–private cooperation

In this scheme, the adoption of active defence measures, especially the most aggressive ones (excluding destructive counter-hacking), would be based on an authorisation granted by the state, following verification that the private organisation has the necessary capabilities to carry out such actions (either by itself or through equally authorised providers), as well as internal mechanisms for auditing and ensuring regulatory compliance that allow for *ex ante* and *ex post* review of any response action.

Then, the authorised organisations would immediately notify the administration of any serious attack received (or imminently expected) and the measures to be taken, and the state would be able in each case to authorise the response or, if deemed appropriate, to force the organisation to suspend it or even transfer it to law enforcement or military units.

The authorising commission would be composed of both government representatives (home affairs, defence and foreign affairs) and members of the judiciary, so that the interests of the state and the necessary balance of fundamental rights could be taken into account.

In addition, the collaboration and exchange of information between the state and the authorised private organisations should be extremely close, through the establishment of public–private partnerships, reinforced obligations of transparency and information to the authorising commission (with periodic controls), as well as a regime of heavy penalties in case of non-compliance.

Some recent experiences have demonstrated the success of a public–private partnership. For example, in 2020, Microsoft coordinated various actors in up to 35 countries to successfully take down the 'Necurs' botnet,[649] one of the most extended at the time (up to 9 million computers affected), which was behind millions of fraudulent emails and indiscriminate attacks, including the dissemination of malware. This action was carried out under the authorisation of a US court and involved the coordination of multiple public and private partners, including computer emergency response teams (CERTs) and police units from various countries.

Involvement of an auditor of the system: the insurance industry

Hoffman and Levite[650] have explored the idea that, in order to provide even greater guarantees to the system, the authorised organisations, besides strict prior controls and periodic checks, would be subject to the supervision of the insurance industry. This would take on an essential role, taking charge of establishing the principles and practices, standards and 'rules of engagement' for an adequate response to attacks, thus preventing the measures adopted in each case from exceeding reasonable limits. In this regard, the authors pointed out that the insurance industry has the potential capacity to influence all the actors involved (by setting premiums and available coverage) and gather the necessary intelligence from them, allowing it to became 'the most adroit incentivizer and steward of effective risk reduction'.[651]

This scheme has proved successful in the physical world in other contexts of private response to threats from hostile groups, and could be appropriate for

649 Microsoft, 'New action to disrupt world's largest online criminal network' (10 March 2020), available at: https://blogs.microsoft.com/on-the-issues/2020/03/10/necurs-botnet-cyber-crime-disrupt/

650 Wyatt Hoffman and Ariel Levite, 'Private Sector Cyber Defense', Carnegie Endowment for International Peace, 2017), available at: https://carnegieendowment.org/2017/06/14/private-sector-cyber-defense-can-active-measures-help-stabilize-cyberspace-pub-71236

651 Idem, p. 26.

relieving the public administration of an excessive supervision burden. Thus, for example, in the framework of the armed response to maritime piracy in the Horn of Africa, experience shows that the onboarding of private maritime security contractors (heavily armed) radically reduced the threat,[652] and that the insurance industry quickly backed the practice (long before the states). During the worst days of the piracy wave in the region, one leading insurance broker successfully offered significant discounts on insurance for ships hiring a private security contractor.[653] Additionally, in order to avoid the 'wild west' situation of a proliferation of unlicensed security contractors (sometimes without sufficient training or resources), the insurance industry worked on the development of a 'bespoke comprehensive insurance package designed specifically for the private maritime security industry'[654] with the Security Association of the Maritime Industry (SAMI), an organisation gathering 180 of these contractors which was involved in the development of such industry standards as ISO 28007 and the 100 Series Rules for the Use of Force. It is an important precedent showing the willingness of the insurance sector to work with companies with sufficient maturity and ensuring the respect of certain rules and standards.

How could this audit mechanism work in practice? According to our proposal, in order to obtain the necessary administrative authorisation for the exercise of active defence measures, an organisation would have to prove beforehand its subscription—with a reputable company—of a mandatory insurance policy covering any damage caused to innocent third parties. This would impose a strict liability regime similar to that which covers traffic accidents under compulsory automobile insurance. The policy would modulate premiums not only according to the company's cyber capabilities but also based on the catalogue of active defence measures that it could eventually implement, from mere intelligence-gathering measures (deception, honeypots, beacons) to the most aggressive activity outside its own systems and measures that could result in damage to third parties, such as system patching, botnet takedowns, distributed denial-of-service (DDoS) counter-attacks or sinkholing. Of course, measures of voluntary destruction or disruption of the hacker's systems would remain out of coverage and, in our proposal, would also be excluded from any administrative authorisation.

652 Idem.

653 Insurance Journal, 'Marsh teams up with marine security firm REDfour to combat piracy' (13 October 2009), available at: https://www.insurancejournal.com/news/international/2009/10/13/104479.htm

654 Maritime Cyprus, 'The Security Association for the Maritime Industry (SAMI) announces voluntary liquidation' (19 April 2016), available at: https://maritimecyprus.com/2016/04/19/the-security-association-for-the-maritime-industry-sami-announces-voluntary-liquidation/

A mandatory insurance policy would also enable faster and more transparent responses to cyber-attacks. Not surprisingly, one of the most common causes of delays in the adoption of responsive measures (and eventually of retaliatory or active defence measures) is the indecision that affects companies' reaction protocols. Thus, fearful of the possible sanction resulting from the recognition of a breach, organisations are forced to obtain legal, financial and reputational expert advice in a decision-making process that ends up being lengthy and ineffective. This process could be simplified and lightened if the organisation is aware from the outset that the measure taken and any damage involuntarily caused to innocent third parties will be covered by the active defence insurance policy.

Consider as an example one of the most famous cases of hack-back recognition by a non-state actor: the so-called Operation Aurora, described by Egloff.[655] When Google detected the security breach in mid-December 2009, it took several days before it adopted countermeasures (tracing back the attack to a server in Hong Kong), a few more before it decided to inform the State Department (on 11 January 2010) and several more before it started collaborating with the NSA in the investigation of the incident. In our opinion, under the mechanism proposed here, not only would Google have reacted with greater legal certainty by responding hand in hand with the federal government (under the prior administrative authorisation regime), but the decision and communication times would have been considerably shortened as the decision-making executives would have had, from the very discovery of the breach, a route marked both by the authorisation regime and by the insurance company.

Another example of the advantages of such compulsory insurance policies is when damage to an undetermined number of third parties is more likely, as for example in botnet-takedown operations. Indeed, although several botnets have been successfully disabled in recent years by the tech industry, it has also been reported that some of these operations 'took up to five million unrelated websites offline as a collateral damage'.[656] Thus, on the one hand, the initiator of this type of measure can be exposed to huge claims if sensitive or critical infrastructures are affected (think of a hospital or an electricity grid), so the subscription of the policy would allow this risk to be covered. But, on the other hand, what for the innocent victims would normally consist of an enormously complicated claim in

655 Florian Egloff, Cybersecurity and Non-State Actors: a Historical Analogy with Mercantile Companies, Privateers, and Pirates (PhD thesis, University of Oxford, 2018), available at: https://ora.ox.ac.uk/objects/uuid:77eb9bad-ca00-48b3-abcf-d284c6d27571

656 Dennis Broeders, 'Private active cyber defense and (international) cyber security—pushing the line?', Journal of Cybersecurity 7 (1) (2021), available at: https://doi.org/10.1093/cybsec/tyab010

multiple jurisdictional forums could be greatly facilitated if the mandatory policy were generalised as a corollary of the administrative authorisation mechanism of active defence.

In sum, competition between insurance and reinsurance companies themselves, as well as their intention not to assume disproportionate compensations derived from excessive responses (or against state-backed cyberattacks[657]), could make it possible to achieve a necessary balance and standardise the active defence measures to be implemented in each circumstance.

Legal certainty: making the concept of self-defence more flexible and assuming states' responsibility in the event of escalation or retaliation

Finally, in order to cover cases in which, as a consequence of a defective attribution, harm is caused to innocent third parties within the framework of the active defence response, a flexible application of the legal justifications (self-defence or plea of necessity) could be considered, so that those authorised organisations would be exempted from criminal liability, assuming, however, the payment of the corresponding civil remedies or compensations for damages caused to third parties (assumed by the insurance industry, as explained).

As described above, this possibility is raised by the Explanatory Report of the Budapest Convention itself, which, in our opinion, Member States should consider in order to provide greater legal certainty to the active defence system.

Furthermore, to minimise the risk of escalation of hostilities, in the most serious cases, active defence actions could be assumed by the state under its own self-defence exception, and afterwards resolved in the corresponding international jurisdictions.

657 James Rundle, 'Lloyd's to exclude catastrophic nation-backed cyberattacks from insurance coverage', Wall Street Journal (18 August 2022), available at: https://www.wsj.com/articles/lloyds-to-exclude-catastrophic-nation-backed-cyberattacks-from-insurance-coverage-11660861586

Conclusion

At present, there are still many voices opposed to active defence measures, especially in their most aggressive forms, not only because of the various practical problems they raise but also because, as this study has shown, many of them are likely to engage the responsibility of states, both towards their own citizens and towards other sovereign states.

For this reason, the opportunity of a third way has been analysed, based on a scheme of public authorisation in which the state would retain the monopoly of the power of repression, but would subcontract—when deemed appropriate—the adoption of the appropriate response to private organisations (which have greater means at their disposal), all under the supervision of the insurance industry as overseer of the system and providing the criminal-legal system with a flexible interpretation of self-defence, while the damage caused to innocent third parties would not be left without adequate coverage.

If it were implemented, the success of this third way would however ultimately depend on how key issues were resolved: among others (i) the speed at which the administration would grant authorisation in each case; (ii) what level of aggressivity would be authorised; (iii) how problems of attribution or the possible escalation of hostilities would be dealt with despite being backed by the state; (iv) what standards or requirements would be needed for authorisation; or (v) whether the authorised organisation could in turn resort to specialised providers (perhaps located in other states) for the response.

CHAPTER 13

Humanitarian organisations under cyber-attack

Emerging threats and humanitarian actors' responsibilities under international human rights law

FRANCESCA ROMANA PARTIPILO AND MARTA STROPPA[658]

Introduction

On 18 January 2022, the International Committee of the Red Cross (ICRC) detected a cyber-attack against its servers, resulting in the hacking of information connected to the Red Cross and Red Crescent Movement's Restoring Family Links services. To this day, the ICRC does not know who was

658 Although the authors have equally contributed to this chapter and share the responsibility for the entire work, just for evaluation purposes, paragraphs 1, 2, 3 and 8 should be attributed to Francesca Romana Partipilo, while paragraphs 4, 5, 6 and 7 should be attributed to Marta Stroppa.

responsible for the attack or whether stolen information has been made available to others; nonetheless, there is a considerable risk that the data have been sold to third actors who might use them to identify and target vulnerable people.[659] Robert Mardini, ICRC's director-general, strongly condemned the attack, declaring that it was 'an affront to humanity, endangering those already suffering the effects of war or disaster'.[660]

The cyber-attack against the ICRC is an example of cyber threats that can undermine the work of humanitarian actors, with detrimental consequences for vulnerable communities. Information is indeed a basic asset in humanitarian response: humanitarian organisations handle sensitive information in most of their daily activities. Whenever this information falls into the wrong hands, it might be used to attack both the organisations and the vulnerable communities to which humanitarian aid is provided.

Acknowledging that cyber-attacks are generally treated as a security threat rather than as a human rights issue, this chapter intends to challenge the traditional framework by introducing a human-rights-based approach to address emerging threats in cyberspace against humanitarian actors. Thus, it will explore the human-rights responsibilities of humanitarian organisations collecting and processing personal data.

First, the chapter will consider the evolution and digitalisation of humanitarian information systems. Secondly, it will analyse humanitarian information systems' main vulnerabilities to cyber threats, while taking account of their impact on vulnerable people's rights. Thirdly, it will focus on the existing standards on data protection applicable to humanitarian organisations. It will be argued that humanitarian organisations' obligations on data protection are still not clear and have often been overlooked in discussions on the matter. Thus, it will suggest the adoption of a human-rights-based approach to data protection in the humanitarian field, by arguing that humanitarian organisations are increasingly bound to promote and ensure the respect of human rights, including when violations might result from cyber-attacks on their own servers. Finally, this chapter will suggest some recommendations on what humanitarian actors should do to mitigate and respond to cyber threats while enhancing the protection of vulnerable individuals' data.

659 ICRC, 'Cyber-attack on ICRC: what we know' (16 February 2022), available at: https://www.icrc.org/en/document/cyber-attack-icrc-what-we-know

660 Statement by Robert Mardini, director-general, ICRC, 'Hacking the data of the world's most vulnerable is an outrage' (29 January 2022), available at: https://www.icrc.org/en/document/hacking-data-outrage

Significantly, the chapter will refer to humanitarian organisations with an international scope, operating across several countries. Due to the nature and the objectives of the work, aimed at enhancing the protection of the beneficiaries of humanitarian organisations, it will address both intergovernmental and non-governmental organisations in the hope of providing the whole humanitarian community with a global framework for data protection. Particular attention will be paid to the humanitarian organisations that do not enjoy special privileges and immunities and that are subject to the national laws of the states in which they are located and operate, in order to provide them with a consistent framework of minimum standard protection that is applicable wherever they conduct their activities. It is worth noting, however, that the same minimum standards of protection might be well extended also to international humanitarian organisations that enjoy special privileges and immunities and are not subject to national laws: including those on data protection.

The 'humanitarian cyberspace'

Today, digitalisation and the recourse to new technologies have a significant impact on the activities carried out by humanitarian organisations worldwide.[661] The use of mobile phones, social media platforms and geospatial technologies has fundamentally altered the environment in which humanitarian organisations operate as well as their modus operandi.[662] Not only are humanitarian organisations reliant on information communication technologies (ICTs) services for their daily functioning, they also exercise an active role within cyberspace, offering their services digitally: they provide assistance to populations in need through mobile cash transfers; have recourse to cloud services to store vital information for their activities; gather and process personal data on a large scale; employ biometric identification technologies as tools for emergency support and

661 Massimo Marelli, '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 102 (913) (2020), 367–387: 367.

662 Kristin Bergtora Sandvik, Maria Gabrielsen Jumbert, John Karlsrud and Mareile Kaufmann, 'Humanitarian technology: a critical research agenda', International Review of the Red Cross 96 (893) (2014), 219–242: 2.

refugee management; and so on.[663] With a view to capturing the expanding presence of humanitarian actors in cyberspace, the expression 'humanitarian cyberspace' was coined.[664] The 'humanitarian cyberspace' resembles the traditional humanitarian space in which humanitarians interact with other actors—such as people of concern, host states, private actors, non-state actors—and try to assist those in need. The difference is that in the humanitarian cyberspace, these undertakings occur through, or are enabled by, ICTs.[665]

Cyber-threats against humanitarian organisations and their detrimental impact on vulnerable people's rights

Regrettably, while the possibility for humanitarian organisations to offer their services digitally has the potential to contribute substantially to the effectiveness of their activities, it might also pose some dangers. Due to their expanding presence in cyberspace, indeed, humanitarian organisations are increasingly faced with the 'techno-violence' carried out therein.[666]

The notion of cyber-threat identifies a wide array of online menaces, ranging from cybercrimes to cyber-attacks, or even cyber-warfare.[667] This chapter primarily focuses on cyber-attacks and cybercrimes, as humanitarian organisations are more easily exposed to these cyber-threats. While cyber-attacks may be described as 'efforts to alter, disrupt, or destroy computer systems or networks

663 Ibid. As for biometric technologies, UNHCR's first 'trialling' of biometric refugee registration was in 2002, when the technology was introduced as a mandatory part of a repatriation programme along the Afghan–Pakistani border. In this regard, see Peter Kessler, 'Afghan "recyclers" under scrutiny of new technology', UNHCR News (3 October 2002).

664 Kristin Bergtora Sandvik, 'The humanitarian cyberspace: shrinking space or an expanding frontier?', Third World Quarterly 37 (1) (2016), 17–32.

665 Ibid., 18.

666 The concept of techno-violence was introduced in 2011 by Lorenzo Magnani, 'Structural and technology-mediated violence: profiling and the urgent need of new tutelary technoknowledge', International Journal of Technoethics 2 (4) (2011), 1–19.

667 Hemen Philip Faga, 'The implications of transnational cyber threats in international humanitarian law: analysing the distinction between cybercrime, cyber-attack, and cyber warfare in the 21st century', Baltic Journal of Law and Politics 10 (1) (2017), 1–34: 4.

or the information or programs on them',[668] the notion of cybercrime refers to 'any crime that is facilitated or committed using a computer network or hardware device'.[669]

Cyber-threats may come from a wide range of actors. States might target a humanitarian organisation to gain access to the data of people resident on their territory who are seeking humanitarian assistance, in order to persecute them and their families. Similarly, private individuals and non-state actors may pose significant security challenges to humanitarian organisations relying on ICTs. Criminals might steal the data stored in humanitarian organisations' servers in order to sell them for profit.[670] Each of these cyber-threats may compromise the work of humanitarian actors by altering the perception of their neutrality, impartiality and independence or even resulting in restrictions on access to beneficiaries or to territories for the organisation.[671] Furthermore, the hacking of sensitive information might endanger some of the fundamental rights enjoyed by the beneficiaries of humanitarian organisations' activities. For instance, the breach of one's right to privacy might be accompanied by huge risks for the life and physical integrity of refugees, asylum-seekers, political dissidents and so on.

By way of example, in the 2022 cyber-attack against the ICRC, there is a considerable risk that the stolen data have been sold to third parties. Such parties might use the data to identify and target vulnerable people who are receiving assistance from the Red Cross and Red Crescent Movement. The danger posed by such an attack has been explicitly recognised by the ICRC director-general, Robert Mardini, in his 'Statement on existing and potential threats in the sphere of information security'.[672] Mardini underlined that while cyber-attacks often mean lost profit or exposed credit card details, in this case the data obtained could potentially be used to cause harm to extremely vulnerable people, including unaccompanied or separated children, detainees, migrants and missing

668 Matthew C. Waxman, 'Cyber-attacks and the use of force: back to the future of Article 2(4)', Yale Journal of International Law 36 (2011), 421–459: 422.

669 Sarah Gordon and Richard Ford, 'On the definition and classification of cybercrime', Journal of Computer Virology 2 (2006), 13–20: 14. Other definitions of cybercrime are broad enough to include not only all crimes committed by means of a computer, but also any crime involving a computer as means or target. See in this respect Debra Little, John Shinder and Ed Tittel, Scene of the Cybercrime: Computer Forensics Handbook (Rockland, MA: Syngress Publishing, 2002), 17.

670 Jack M. Beard, 'Legal phantoms in cyberspace: the problematic status of information as a weapon and a target under international humanitarian law', Vanderbilt Journal of Transnational Law 47 (1) (2014), 67–144: 113.

671 Massimo Marelli and Tilman Rodenhäuser, 'Cyber Disruption of Humanitarian Assistance', NATOCCDCOE (2021), available at: https://cyberlaw.ccdcoe.org/wiki/Scenario_25:_Cyber_disruption_of_humanitarian_assistance

672 ICRC, 'Statement on existing potential threats in the sphere of information security' (31 March 2022), available at: https://www.icrc.org/en/document/humanitarian-data-infrastructures-must-be-protected

people's families.[673] He added that the attack harmed the ICRC's global network's ability to locate missing people and reconnect families. For instance, in the aftermath of the tsunami-induced flooding in Tonga, the ICRC's ability to provide assistance to families and missing people has been hampered. In addition, the tracing work of the ICRC has been endangered in conflict areas, such as for Afghans fleeing violence.[674]

The raising of new cyber-threats against humanitarian actors raises two important issues. On the one hand, it is crucial to understand to what extent humanitarian organisations are protected from cyber-threats under international law, and states' related obligations. On the other hand, it is important to understand the responsibilities that humanitarian organisations have in the field of data protection. In the following sections, both questions will be explored.

Humanitarian organisations' protection under international law

Humanitarian organisations are protected from attack under international humanitarian law (IHL), the branch of international law applying in the context of armed conflicts. Humanitarian organisations not only enjoy the same protection granted to civilian objects and civilians under Articles 51 and 52 of Additional Protocol I to the Geneva Conventions;[675] they also enjoy ad hoc protection under IHL due to the exceptional risks they face while carrying out relief operations. Article 70 of Additional Protocol I to the Geneva Conventions states that once impartial humanitarian operations have been agreed to by the parties to a conflict, they shall be allowed and facilitated.[676] Moreover, according to Article 71 of Additional Protocol I to the Geneva Conventions, the humanitarian staff shall be respected and protected against any harm that may be caused by the

673 ICRC, 'Hacking the data of the world's most vulnerable is an outrage' (28 January 2022), available at: https://www.icrc.org/en/document/hacking-data-outrage

674 Ibid.

675 See Articles 51 and 52 of Additional Protocol I to the Geneva Conventions, which set out the principle of distinction, one of the cardinal pillars of IHL (see also Rules 1 and 7 of the ICRC Customary IHL Study).

676 See Article 70 of Additional Protocol I to the Geneva Conventions. This provision reflects international customary law: see Rules 32 and 55 of the ICRC's Customary IHL Study.

belligerents.[677] Targeting a humanitarian organisation in violation of one of these provisions may amount to a war crime.[678]

Whereas IHL surely provides for a solid protection framework for impartial humanitarian organisations in times of armed conflict, clarification is still needed as to how it applies in the digital sphere.[679] In fact, none of the rules of IHL explicitly deal with cyber operations.[680] There is no doubt that a cyber operation leading to the death or injury of humanitarian personnel, or to the physical damage of objects used in humanitarian operations, amounts to a prohibited attack, in violation of Articles 51 and 52 of Additional Protocol I.[681] When it comes to cyber operations that do not cause physical damage to civilian objects or death or injury to civilians, but rather a disruption of digital infrastructures, however, IHL rules are less clear and a variety of opinions exist in this regard.[682] Nonetheless, even if one assumes that cyber operations not causing physical damage will not be considered an 'attack' under IHL, thus not representing a breach of Articles 51 and 52 of Additional Protocol I, they might still be prohibited under Articles 70 and 71 of Additional Protocol I. Disruption of the digital infrastructure, as well as hacking, stealing or manipulation of humanitarian data, would indeed unduly interfere with humanitarian operations,[683] jeopardise the humanitarian

677 See Articles 71 of Additional Protocol I to the Geneva Conventions. This provision reflects international customary law: see Rule 31 of the ICRC's Customary IHL Study.

678 See Article 8, paragraphs (b) (iii) and (e) (iii), of the Statute of the International Criminal Court.

679 Tilman Rodenhäuser, Balthasar Staehelin and Massimo Marelli, 'Safeguarding humanitarian organizations from digital threats', ICRC Blog (13 October 2022), available at: https://blogs.icrc.org/law-and-policy/2022/10/13/safeguarding-humanitarian-organizations-from-digital-threats/

680 Dan-Iulian Voitasec, 'Applying international humanitarian law to cyber-attacks', Lex et Scientia International Journal XXII (1) (2015), 124–131: 128.

681 Rodenhäuser et al. (see note 21 above).

682 See Michael N. Schmitt, Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations (Newport, RI: United States Naval War College, 2017), commentary on Rules 92 and 100; for the ICRC position, see ICRC, 'International humanitarian law and cyber operations during armed conflicts' (November 2019), available at: https://www.icrc.org/en/document/international-humanitarian-law-and-cyber-operations-during-armed-conflicts; for national positions, see the NATO CCDCOE Cyber Law Toolkit, https://cyberlaw.ccdcoe.org/wiki/Attack_(international_humanitarian_law)

683 See the Tallinn Manual 2.0, Rule 145, which states that 'Cyber operations shall not be designed or conducted to interfere unduly with impartial efforts to provide humanitarian assistance.'

personnel's safety and security, and undermine the perception of humanitarian organisations' impartiality.[684]'

States should further clarify IHL rules in the light of the new realities spurred by the digital revolution, in order to enhance humanitarian organisations' protection in cyberspace.[685] Discussions on humanitarian organisations' protection under international law should take into account innovative solutions that are currently studied by the ICRC and the civil society, such as the adoption of a digital emblem—'a distinctive emblem, signal or other digital means to identify the data and digital infrastructure of organisations and entities entitled to display the distinctive emblems recognised under international humanitarian law and to indicate, where applicable, their legal protection'[686]—or the development of a 'sovereign humanitarian cloud' to protect humanitarian data.[687]

Furthermore, states should reach an agreement over the protection of humanitarian organisations outside the context of armed conflicts, based on the long-standing international consensus on the importance of impartial humanitarian activities.[688] As underlined by the ICRC in its resolution 'Safeguarding Humanitarian Data', it is crucial that impartial humanitarian organisations be

684 See Rodenhäuser et al. (note 21 above); Tilman Rodenhäuser, 'Hacking humanitarians? IHL and the protection of humanitarian organizations against cyber operations', EJIL:Talk! (16 March 2020), available at: https://www.ejiltalk.org/hacking-humanitarians-ihl-and-the-protection-of-humanitarian-organizations-against-cyber-operations/; Laurent Gisel, Tilman Rodenhäuser and Knut Dörmann, '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 102 (913) (2020), 287–334: 329. The fact that cyber operations may erode trust in impartial organisations, jeopardising their ability to operate as well as the safety of their staff and of the people in need, has been underlined by the ICRC in 'Safeguarding humanitarian data', Resolution CD/22/R12 (June 2022), para. 3, available at: https://rcrcconference.org/app/uploads/2022/06/CD22-R12-Safeguarding-Humanitarian-Data_23-June-2022_FINAL_EN.pdf

685 In this respect, states may also rely on the Martens Clause, according to which 'in cases not covered by specific international agreements, civilians and combatants remain under the protection and authorities of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience'. As pointed out by the International Court of Justice in its 1996 Advisory Opinion on the Legality of the Threat and Use of Nuclear Weapons, the Martens Clause is 'an effective means of addressing the rapid evolution of military technology' (para. 78). Furthermore, the Martens Clause allows both legal and moral arguments to be taken into account. Thus, while considering whether harming impartial humanitarian organisations in cyberspace is against customary law, even if the attack does not amount to an 'armed attack' under IHL, states should also consider whether it is against the principles of humanity and the dictates of public conscience.

686 ICRC (see note 26 above), para. 1. See also Tilman Rodenhäuser, Larry Maybee, Fabrice Lauper, Laurent Gisel and Hollie Johnston, 'Signaling legal protection in a digitalizing world: a new wea for the distinctive emblems?', ICRC Blog (16 September 2021), available at: https://blogs.icrc.org/law-and-policy/2021/09/16/legal-protection-digital-emblem/; Felix E. Linker and David Basin, 'Signaling legal protection during cyber warfare: an authenticated digital emblem', ICRC Blog (21 September 2021), available at: https://blogs.icrc.org/law-and-policy/2021/09/21/legal-protection-cyber-warfare-digital-emblem/; and Antonio De Simone, Erin Hahn and Brian Haberman, 'Identifying protected missions in the digital domain', ICRC Blog (23 September 2021) available at: https://blogs.icrc.org/law-and-policy/2021/09/23/protected-missions-digital-domain/

687 Massimo Marelli, 'The SolarWinds hack: lessons for international humanitarian organizations', International Review of the Red Cross 104 (919) (2020), 1267–1284.

688 ICRC (see note 26 above), preamble.

respected and protected offline and online not only during warfare, but also during natural disasters and other emergencies (which are outside the scope of IHL).[689] Currently, however, the only obligations states have with respect to humanitarian organisations may be found in IHL, which only applies in times of armed conflicts. For this reason, it is crucial that humanitarian organisations adopt and implement cybersecurity measures in order to protect themselves and the data they have collected and processed from possible incoming attacks.

Humanitarian organisations' responsibilities in the field of data protection

Data protection legislation has evolved rapidly in recent years, to the extent that around 120 countries now have national laws or statutory requirements concerning data protection in their domestic legislation.[690] The rapid evolution of national data protection legislation, however, might raise some challenges for humanitarian organisations operating in several countries. They deal with a patchwork of national provisions that may differ from one another, creating confusion as to humanitarian organisations' responsibilities and gaps in the protection of data.[691] Furthermore, in some states where humanitarian organisations operate, the personal data protection legislation might be embryonic, non-existent or not entirely enforceable (given the extraordinary circumstances characterising humanitarian emergencies).

In addition, not all humanitarian organisations are subject to national laws. In fact, international organisations providing humanitarian assistance, such as UN agencies or the ICRC, enjoy special privileges and immunities from domestic legislation. In order to fill the protection gap, some organisations have adopted

689 Ibid., paras 6 and 12.

690 UN Conference on Trade and Development (UNCTAD), 'Data protection regulations and international data flows: Implications for trade and development' (2016), available at: http://unctad.org/en/pages/PublicationWebflyer.aspx?publicationid=1468

691 UNOCHA, 'Humanitarianism in the age of cyber-warfare: Towards the principled and secure use of information in humanitarian emergencies' (October 2014), p. 8, available at: https://www.unocha.org/sites/unocha/files/Humanitarianism%20in%20the%20Cyberwarfare%20Age%20-%20OCHA%20Policy%20Paper%2011.pdf

their own internal policies and strategies for data responsibility in humanitarian action. At the UN level, in 1990 the General Assembly adopted a first set of Guidelines for the Regulation of Computerized Personal Data Files, which called inter alia for a careful application of data protection in humanitarian emergencies.[692] In 2010, the International Organization for Migration (IOM) adopted the Data Protection Manual,[693] which includes the IOM's data protection principles as informed by relevant international standards, as well as generic templates and checklists to be followed when collecting and processing personal data. Similarly, in 2015 the UN High Commissioner for Refugees (UNHCR) adopted its own Policy on the Protection of Personal Data of Persons of Concern to the UNHCR,[694] followed in 2018 by a related guidance.[695] In 2021, the UN Office for the Coordination of Humanitarian Affairs (OCHA) adopted its Data Responsibility Guidelines,[696] a set of principles, processes and tools that support data responsibility in OCHA's work. Outside the UN, in 2020 the ICRC established its own data protection framework, intended to ensure that its humanitarian activities and operations are carried out in a manner consistent with internationally recognised standards for protecting personal data.[697] The ICRC's data protection framework includes, inter alia, the ICRC Rules on Personal Data Protection,[698] adopted in 2015 and amended in 2019, as well as the establishment of the ICRC Data Protection Office and Data Protection Commission—two supervisory bodies responsible inter alia for monitoring the application of the ICRC Rules.

Whereas these guidelines certainly constitute considerable progress in the protection of personal data in humanitarian emergencies, they only apply to the adopting agency. Still, they have been crucial in raising awareness on the need to adopt a universal framework applying to the overall humanitarian community. Such urgency has been underlined also in international fora such as the 33rd

692 UNGA, Guidelines for the Regulation of Computerized Personal Data Files (Resolution 45/95, 14 December 1990).

693 IOM, Data Protection Manual (2010), available at: https://publications.iom.int/books/iom-data-protection-manual

694 UNHCR, Policy on the Protection of Personal Data of Persons of Concern to the UNHCR (2015), available at: https://data2.unhcr.org/en/documents/details/44570

695 UNHCR, Guidance on the Protection of Personal Data of Persons of Concern to UNHCR (2018), available at: https://www.refworld.org/docid/5b360f4d4.html

696 UNOCHA, Data Responsibility Guidelines (2021), available at: https://centre.humdata.org/the-ocha-data-responsibility-guidelines/

697 ICRC, Data Protection Framework (2020), available at: https://www.icrc.org/en/document/icrc-data-protection-framework

698 ICRC, Rules on Personal Data (2015, amended in 2019), available at: https://www.icrc.org/en/publication/4261-icrc-rules-on-personal-data-protection

Annual Conference of Data Protection and Privacy Commissioners, which took place in 2011 in Mexico City and resulted in the Resolution on Data Protection and Major Natural Disasters,[699] and the 37th International Conference of Data Protection and Privacy Commissioners, which was convened in Amsterdam in 2015 and produced the Resolution on Privacy and International Humanitarian Action.[700] Both resolutions underlined the risks deriving from the use of ICTs in humanitarian emergencies and called for states and international organisations to take them into account in their response efforts. Furthermore, the Amsterdam Resolution echoed the UNOCHA's Report on Humanitarianism in a Networked Age[701] and the IFRC's World Disaster Report[702] call for the adoption of 'clear guidelines and standards for how and by whom the information they collect will be processed, used and stored'.[703] Another important tool reflecting the urgency to adopt a framework to grant humanitarian data protection is the abovementioned ICRC Resolution 'Safeguarding Humanitarian Data', adopted in June 2022 in the wake of the cyber-attack against the ICRC's servers in January 2022.[704]

To date, however, the sole attempt to adopt some practical guidelines on data protection in the humanitarian field may be found in the 2018 Handbook on Data Protection in Humanitarian Action, elaborated by the ICRC together with the Brussels Privacy Hub.[705] The handbook, now in its second edition, was inspired by existing guidelines, working procedure and practices that have been established in the humanitarian domain, as well as by a wide variety of international data protection instruments and standards.[706] Its main objective is that of 'providing specific guidance on the interpretation of data protection principles in the context of Humanitarian Action, particularly when new technologies are employed'.[707] While it certainly constitutes valuable guidance on data protection

699 ICDPPC, Resolution on Data Protection and Major Natural Disasters (2011), available at: http://globalprivacyassembly.org/wp-content/uploads/2015/02/Resolution-on-Data-Protection-and-Major-Natural-Disasters.pdf?mc_phishing_protection_id=28047-br1tehqdu81eaoar3q10

700 ICDPPC, Resolution on Privacy and International Humanitarian Action (2015), available at: https://edps.europa.eu/sites/default/files/publication/15-10-27_resolution_privacy_humanitarian_action_en.pdf

701 UNOCHA, Humanitarianism in the Network Age (2012), available at: https://www.unocha.org/sites/unocha/files/HINA_0.pdf

702 IFRC, World Disaster Report (2013), available at: https://reliefweb.int/attachments/2be4c3dc-b6dd-32de-83d2-47acb6c9df82/World%20Disasters%20Report.pdf

703 Ibid., p. 145.

704 ICRC (see note 26 above).

705 Cristopher Kuner and Massimo Marelli (eds), Handbook on Data Protection in Humanitarian Action, 2nd ed. (Geneva: ICRC, 2020), available at: https://www.icrc.org/en/data-protection-humanitarian-action-handbook

706 Ibid., 20–23.

707 Ibid., 21.

in the humanitarian field, however, it fails to address a crucial issue, namely to what extent humanitarian actors are required to adopt a data protection framework and on which basis.[708] In the next section, a human-rights-based approach to cybersecurity in the context of humanitarian emergencies will be suggested.

A human-rights-based approach to cybersecurity in humanitarian emergencies

The cyber threats to which humanitarian organisations are exposed have magnified the urge to adopt cybersecurity measures in the humanitarian field. In the absence of an international instrument on the matter, this chapter suggests that humanitarian organisations shall adopt cybersecurity measures as part of their efforts to protect vulnerable people's human rights, and that such a framework should be derived from extant international human rights provisions.

Traditionally, cybersecurity has largely been perceived as a security issue. In our view, however, it should also be considered a human rights issue. As previously demonstrated, indeed, stolen data may be easily used by other actors to identify and attack vulnerable people, putting at stake their fundamental rights, including the right to life and to personal integrity. Thus, any cybersecurity framework adopted by humanitarian actors should encompass a human-rights-based approach.

While one could easily argue that human rights protection should be granted by states and not by humanitarian actors, since international human rights law generally places primary obligations on states, this view is increasingly and constantly challenged. As argued by Clapham, indeed, human rights are often spelled out in norms and provisions that are not written 'with regard to a specific duty-holder'.[709] Therefore, it has 'increasingly been accepted, in the second half of

708 See also Asaf Lubin, 'Data protection as an international legal obligation for international organizations: the ICRC as a case study', in Russel Buchan and Asaf Lubin (eds), The Rights to Privacy and Data Protection in Times of Armed Conflict (Tallinn: NATO CCDCOE Publications, 2022), 256.

709 Andrew Clapham, Human Rights Obligations of Non-State Actors (Oxford: Oxford University Press, 2006), 34.

the 20th century, that non-state actors are, or should be, the bearers of international legal obligations'.[710]

This is even more evident in the context of international disaster law: as underlined by Natoli, 'In disaster-related issues, the agency of non-state entities is likely to have relevant consequences on proper systemic dynamics as law-making and liability, also in the light of their capacity to carry out public-like functions in emergency situations in which the host State could temporally fail.'[711] For this reason, humanitarian actors are increasingly recognised as playing a crucial role in the implementation of human rights obligations.

Human rights protection is indeed at the core of humanitarian action, which follows the principles of impartiality, neutrality, independence and humanity. Significantly, humanitarian actors' protection efforts should be aimed at preventing the negative effects not only of the crisis but also deriving from the humanitarian response to the emergency. According to the well-established 'do no harm' principle, indeed, while carrying out their activities, humanitarian organisations should 'avoid exposing people to further harm as a result of your actions'.[712]

While it is true that the right to data protection is still not settled in customary law, it is strongly interconnected with other fundamental rights, such as the right to life and personal integrity and the right to privacy. Furthermore, as more and more nations are adopting data protection as a mandatory legal framework, it might well be that a similar right will crystallise in international law, as argued by other commentators.[713] Therefore, human rights protection efforts should be accordingly extended also to the digital sphere, to prevent any cyber-threat directed against servers from resulting in a violation of human rights.

710 Jean d'Aspremont, André Nollkaemper, Ilias Plakokefalos and Cedric Ryngaert, 'Sharing responsibility between non-state actors and states in international law: introduction', Netherlands International Law Review 62 (2015), 49–67: 50.

711 Tommaso Natoli, 'Non-state humanitarian actors and human rights in disaster scenarios: normative role, standard setting and accountability', in Flavia Zorzi Gustiniani, Emanuele Sommario, Federico Casolari and Giulio Bartolini (eds), Routledge Handbook of Human Rights and Disasters (New York: Routledge, 2018), 149–164: 149.

712 Sphere Project, 'Humanitarian Charter and Minimum Standards in Humanitarian Response' (2011), p. 33, available at: https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/2011SPHEREHandbookHC-PP-Annex1-Annex2.pdf

713 Lubin (see note 50 above), 256–257. See also Christopher Kuner, 'The internet and the global reach of EU law', in Marise Cremona and Joanne Scott (eds), EU Law Beyond EU Borders: The Extraterritorial Reach of EU Law (Oxford: Oxford University Press, 2019), 112–145: 131.

Some recommendations for a 'cyber-secure' humanitarian action

Regrettably, several humanitarian organisations are still characterised by limited cybersecurity preparedness, organisational readiness and digital literacy. The wide variety of national and regional data protection rules, the lack of an international treaty on data protection in the humanitarian field and the lack of consensus on the recognition of a customary right to data protection are indeed hindering the emergence of a global framework for data protection in the humanitarian domain.

On the one hand, states should take account of the overlaps between data protection and humanitarian assistance, in order to elaborate a global framework aimed at granting a minimum level of data protection in humanitarian response all over the world. The debate on humanitarian data protection under IHL might be a valid starting point for further reflecting on the protection of data, including outside armed conflicts.[714] At the same time, however, humanitarian organisations should adopt their own data protection frameworks, aimed at offering the maximum protection to the data they collect and use while carrying out their activities.[715]

The following non-binding recommendations are specifically intended to assist humanitarian organisations in the fulfilment of their protection efforts in the digital sphere. Acknowledging the wide spectrum of national and regional legislations on data protection that applies to humanitarian actors all over the world, these recommendations are not intended to come into conflict with the applicable laws, but rather to provide minimum standards for the protection of data in the humanitarian context. It follows that humanitarian actors may be able to provide for stricter critria whenever they deem it necessary or it is required by domestic or regional provisions.

714 For a comprehensive analysis of the debate over data as 'object' under international humanitarian law, see Kubo Mačák, 'Military objectives 2.0: the case for interpreting computer data as objects under international humanitarian law', Israel Law Review 48 (1) (2015), 55–80.

715 As underlined by the ICRC in its resolution 'Safeguarding Humanitarian Data', 'the digital transformation of impartial humanitarian organizations' structure and activities, including the processing of humanitarian data, entails and important responsibility for these organizations to adopt and implement cyber security measures and data protection practices'. See ICRC (note 26 above), para. 1.

Recommendation 1: Humanitarian organisations shall define the 'cyber-perimeter' of their operations and the potential threats and harms inherent in their action.

As a first step, any humanitarian organisation relying on ICTs should define what has been called the 'cyber-perimeter' of its operations. As suggested by Massimo Marelli, head of the Data Protection Office of the ICRC:

> Clearly defining the digital boundaries within which they carry out operations lays the groundwork for humanitarian organizations to develop a strategy to support and protect humanitarian action in a digital environment, channel available resources to where they are most needed, remain effective in their relationship with host countries and other stakeholders in cyber geopolitics, and understand the areas in which their operational dialogue and working modalities need to be adapted for cyberspace.[716]

Thus, humanitarian organisations should first and foremost define which activities will be carried out throughout cyberspace, as well as how they will be conducted and for what purpose. In this respect, humanitarian organisations should also conduct a rigorous questioning as to whether their reliance on data (especially very sensitive data, such as facial recognition, fingerprinting or other biometric data) is necessary in the first place for carrying out certain activities.[717]

716 Massimo Marelli, 'Hacking humanitarians: moving towards a humanitarian cybersecurity strategy', ICRC Blog (16 January 2020), available at: https://blogs.icrc.org/law-and-policy/2020/01/16/hacking-humanitarians-cybersecurity-strategy/ Massimo Marelli and Adrian Perrig, 'Hacking humanitarians: mapping the cyber environment and threat landscape', ICRC Blog (7 May 2020), available at: https://blogs.icrc.org/law-and-policy/2020/05/07/hacking-humanitarians-mapping-cyber-environment/

717 As underlined by Ella Jakubowska, reliance on new technologies (such as biometric systems) in humanitarian assistance is often not necessary—and it might be very dangerous if those data are stolen or fall into the wrong hands. As such, it is crucial that humanitarian organisations carefully consider whether the collection of sensitive data is actually necessary for the conduct of their activities or whether there might be other possible paths to follow. Furthermore, humanitarian organisations should consider whether the assisted populations are comfortable with the use of said technologies. As reported by Petra Molnar, indeed, 'under the justification of efficiency, refugees in Jordan have their irises scanned in order to receive their weekly rations. Some refugees in the Azraq camp have reported feeling like they did not have the option to refuse to have their irises scanned, because if they did not participate, they would not get food. This is not free and informed consent.' See, respectively, Ella Jakubowska, 'Do no harm? How the case of Afghanistan sheds light on the dark practice of biometric intervention', European Digital Rights (EDRi) (17 November 2021), available at: https://edri.org/our-work/do-no-harm-how-the-case-of-afghanistan-sheds-light-on-the-dark-practice-of-biometric-intervention/; Petra Molnar, 'The human rights impacts of migration control technologies', European Digital Rights (EDRi) (12 February 2020), available at: https://edri.org/our-work/the-human-rights-impacts-of-migration-control-technologies/.

Furthermore, humanitarian organisations should analyse the possible risks they might face while carrying out their activities in the digital sphere, with particular attention to threats that might jeopardise the human rights of particularly vulnerable persons they are assisting. Indeed, only by defining the activities that will be carried out in cyberspace, the modus operandi and the potential challenges that might arise will it be possible to adopt ad hoc measures aimed at protecting the personal data of the most vulnerable persons.

Significantly, the cyber-perimeter of a specific organisation should also take account of activities conducted by the third parties with whom humanitarian organisations are working, such as technology service providers. Whenever there might be a 'conflict of perimeter' between two or more humanitarian organisations or between a humanitarian organisation and its partners, the actors should operate according to measures that provide maximum protection of data for the whole duration of the collaboration.

Recommendation 2: Humanitarian organisations shall adopt data protection policies.

Humanitarian organisations should also adopt internal data protection policies, which should offer at least the same protection as recognised in existing international principles and guidelines on data protection and should be tailored to the realities of humanitarian response. Furthermore, they should be implemented at the practical level, into all stages of activities (from project design to data collection, storing, analysis and treatment).

Recommendation 3: Humanitarian organisations shall develop a cybersecurity strategy that provides for concrete steps to ensure personal data protection against digital threats.

While implementing their internal policies on data protection, humanitarian organisations should also adopt a solid cybersecurity strategy, which should provide for concrete steps to ensure data protection against possible digital threats. Significantly, a cybersecurity strategy should be intended as more than merely protecting the humanitarian organisations' cyber networks and tools: it should also be considered as a way of granting the protection of human rights in a networked age. For this reason, the cybersecurity strategy should be comprehensive

and be aimed at addressing all the threats that a humanitarian organisation may encounter in the digital environment, so as to grant maximum protection to the vulnerable people whose data are collected and treated in the conduct of humanitarian action.

In the delineation of a cybersecurity strategy, humanitarian organisations should first define which is the applicable law and which are their duties and obligations with respect to data protection. Humanitarian organisations should consider the legal framework applicable both where they have their headquarters and in the countries in which they are operating. Furthermore, humanitarian organisations should take account of the law applicable to the partners they are working with. If a humanitarian organisation is an international organisation enjoying privileges and immunities from national law, it should still clarify the application of those privileges and immunities to the data it stores and processes, directly or through a third-party service provider and other partners.[718]

A solid cybersecurity strategy should provide for ad hoc technical measures aimed at granting special protection to the personal data collected and treated, which might be stricter than the ones required by the applicable law. As mentioned above, humanitarian organisations often work with personal data of people affected by armed conflicts or other situations of violence. For this reason, it is of particular importance they adopt ad hoc measures aimed at addressing the main threats they may face in the digital sphere, with a view to protecting the vulnerable persons they are assisting. These technical measures should be tailored to the specific activities carried out by humanitarian organisations: they should take account of the context in which the organisations are operating, the actors involved, the nature of the personal data they are collecting and for which purpose, as well as the vulnerabilities of the persons to which they provide digital services. Furthermore, these technical measures should be based on the 'confidentiality, integrity and availability' principles: (1) access to data shall be granted by humanitarian organisations and their partners to intended users only;[719] (2) data shall be collected and treated accordingly to the principles of accuracy and integrity, by avoiding any form of intentional or non-intentional manipulation; (3) humanitarian organisations relying on ICTs for their activities

718 Marelli (see note 3 above).

719 Confidentiality is particularly crucial, since humanitarian organisations may come under pressure to provide humanitarian data to national authorities wishing to use such data for other purposes, which might result in a violation of vulnerable people's fundamental rights. For this reason, whenever humanitarian organisations process humanitarian data, they should do solely for purposes that are compatible with their exclusively humanitarian mandate. See ICRC (note 26 above), preamble and para. 2.

shall also adopt any feasible measure to grant constant and safe access to their digital services.[720]

Of course, in order to grant legal and technical protection to personal data, humanitarian organisations should cooperate with several actors. For this reason, it is of particular importance that they also adopt specific procedures to ensure that their dialogue with any relevant stakeholder is always confidential, neutral and impartial. Massimo Marelli has identified three categories of relevant stakeholders, namely the 'cyber-host state' (that hosts the necessary infrastructure for humanitarian organisations' digital services), the states where the humanitarian organisation intends to offer its digital services, and other state and non-state actors that are operating in the territory of the state where the humanitarian organisation is operating.[721] All these actors play a central role in the success of the humanitarian organisations' digital services and respect of data protection standards.

Recommendation 4: Humanitarian organisations shall undertake a continuous due diligence process on data protection during all the phases of their digital activities.

Cyber-threats are constantly changing, and every day hackers find new vulnerabilities to exploit for launching their cyber-attacks. For this reason, it is fundamental for humanitarian organisations to constantly carry out a due diligence process on data protection in order to (1) identify and assess new potential cyber threats; (2) integrate the findings of such assessment in their internal policies, strategies and procedures in order to prevent or mitigate adverse impact on data protection and vulnerable people's human rights; (3) monitor the effectiveness of the adopted measures and (4) provide an account of the means by which they have addressed such digital threats and the ultimate outcome of the adopted measures.

720 Marelli and Perrig (see note 58 above).

721 Massimo Marelli and Martin Schüepp, 'Hacking humanitarians: operational dialogue and cyberspace', ICRC Blog (4 June 2020), available at: https://blogs.icrc.org/law-and-policy/2020/06/04/hacking-humanitarians-dialogue-cyberspace/

Recommendation 5: Humanitarian organisations shall establish independent supervisory and monitoring mechanisms.

In order to supervise the actual implementation of internal policies and of the due diligence process, humanitarian organisations should establish, whenever possible, an ad hoc independent body.[722] This body should be in charge of monitoring the correct application and implementation of internal policies and procedures, including the due diligence process, at the organisational level. It should also be in charge of reviewing and updating internal policies and procedures in the light of regulatory developments and changes in the humanitarian organisation's activities, as well as of advising on data protection matters. The same body should ensure remedy to subjects who want to file complaints against the organisation for a violation of their rights. Finally, in the case of a data breach, this body should coordinate actions aimed at mitigating its impact on data protection and human rights, as well as on the humanitarian organisation's activities.

Recommendation 6: Humanitarian organisations shall adopt ad hoc procedures to follow in case of a data breach.

In the event that the humanitarian organisation is targeted by a malevolent cyber operation resulting in a data breach, it should have in place ad hoc procedures that allow a prompt response so as to minimise and mitigate the impact of the breach on both vulnerable people's human rights and the humanitarian organisation's personnel and activities. Such procedures should include, inter alia, the immediate mitigation of the risks deriving from the data breach; notification to the affected persons whose data were stolen and the adoption of any feasible measure to grant them protection; collaboration with the humanitarian organisation's partners and with the relevant stakeholders; a security enhancement of the ICT systems; and the restoration of the service.

722 By way of example, see the ICRC Data Protection Office, established in 2020 by the ICRC as part of its Data Protection Framework: https://www.icrc.org/en/document/icrc-data-protection-framework

Conclusion

Acknowledging that cyber-attacks are commonly treated predominantly as a security issue rather than a human rights one, this chapter challenged traditional assumptions by introducing a human-rights-based approach to emerging cyber-threats against humanitarian actors. It first outlined the widening role of humanitarian actors in cyberspace, pointing to the potential gains and dangers connected to a so-called 'humanitarian cyberspace'. Then it analysed states' obligations to respect and protect humanitarian organisations, as well as humanitarian organisations' responsibility to protect themselves and the data they collect and process from incoming attack.

The work underlined how there is still confusion on the matter. On the one hand, states should clarify their obligations towards humanitarian organisations under IHL in the digital domain, as well as extending them outside the context of armed conflict. On the other hand, the international community should further discuss the obligations of humanitarian organisations with respect to data protection. Since states have an obligation to respect and protect humanitarian organisations only in time of armed conflict, it is crucial that humanitarian organisations adopt and implement cybersecurity measures in order to protect themselves and the data they have collected and processed from possible incoming attacks. In this respect, it was suggested that a human-rights-based approach to cybersecurity be adopted, which derives humanitarian organisations' obligations on data protection directly from international human rights law.

Finally, the chapter underlined how humanitarian organisations still have a long way to go to ensure a sufficient level of security against cyber-attacks. Therefore, the work suggested a list of actions that humanitarian actors may put in place in order to adequately protect the data of vulnerable individuals and communities, which include (1) the definition of a 'cyber perimeter' of the activities humanitarian organisations want to carry out in cyberspace; (2) the adoption of internal policies on data protection; (3) the development of a strong cybersecurity strategy; (4) the implementation of a continuous due diligence process on data protection; (5) the establishment of independent monitoring and supervisory mechanisms; and (6) the adoption of ad hoc procedures to follow in the case of a data breach. Of course, these actions alone are not sufficient: the international community should also continue discussions on the matter, in order to delineate a global framework for data protection in humanitarian emergencies that is applicable everywhere, and to all humanitarian organisations.

CHAPTER 14

A responsibility to improve

How global cybercrime cooperation frameworks must better safeguard human rights and protect the humans of cybersecurity

RAMAN JIT SINGH CHIMA

Introduction

What do you get when you mix prosecutors, police, government legal advisers, diplomats, and a small scattering of cybersecurity experts and multi-stakeholder representatives tasked with preparing a binding international treaty to combat cybercriminal activity on an ever more expansive basis, in the middle of escalating geopolitical cyber tension and ever-present worries about digital authoritarianism?

Well, in September 2023—by the 78th session of the UN General Assembly—we will find out.

A United Nations Ad Hoc Committee (UN AHC) has been constituted, on the basis of a resolution originally advanced by the Russian Federation that narrowly succeeded in the Third Committee of the General Assembly in 2019, to draft a 'comprehensive international convention to combat to misuse of ICT and cybercrime'.[723] This potential cybercrime convention is the first, negotiated global effort under the auspices of the UN on this issue, but not the first diplomatic instrument on cybercrime signed and implemented by nation states. The Council of Europe (COE) Budapest Convention has 66 signatories, with a recent second additional protocol (on cross-border investigation requests and data transfers) now open to accession.[724] Several regional cybercrime conventions exist, including the African Union Convention on Cybercrime and Personal Data (the Malabo Convention). [725] These are joined by a range of plurilateral and bilateral legal initiatives, most of which are recently focused around the issue of cybercrime legal cooperation and investigatory powers dealing with global platforms and cross-border data (e.g. the Commonwealth Computer and Computer Related Crimes Model Law, CLOUD Act agreements between the US and other states).[726]

Increased global consensus and international collaboration on countering cybercrime is arguably a net positive. For governments, consensus appears most useful around the scope of activities to criminalise across borders as cybercrime and putting in place international mechanisms for legal assistance, investigative cooperation, and related technical, capacity issues on combating cybercrime. With the even faster proliferation of digital connectivity and services into all aspects of human life during the Covid-19 pandemic, the potential for technologies to be a vector of harm is even clearer, alongside their radically liberating, rights-protecting effect. However, as the contested UN votes and tense negotiating process in its beginning phase has shown, a significant proportion of cyber-policy stakeholders aware of or involved in the UN cybercrime treaty discussions are concerned by its initiation and where it will go to. The official press

723 United Nations General Assembly, Seventy-fourth session, 'Countering the use of information and communications technologies for criminal purposes', A/RES/74/247, 26 May 2021, available at: https://undocs.org/A/Res/74/247

724 'Council of Europe Convention on Cybercrime', opened for signature 23 November 2001, European Treaty Series, no. 185, available at: https://rm.coe.int/1680081561; Cybercrime Convention Committee (T-CY), 'Second Additional Protocol to the Convention on Cybercrime on enhanced co-operation and disclosure of electronic evidence', opened for signature 17 November 2021, available at: https://search.coe.int/cm/pages/result_details.aspx?objectid=0900001680a48e4d#globalcontainer

725 'African Union Convention on Cyber Security and Personal Data Protection', adopted on 27 June 2014, available at: https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection

726 The Commonwealth, 'Model Law on Computer and Computer Related Crime', final draft made available 18 November 2002, available at: https://production-new-commonwealth-files.s3.eu-west-2.amazonaws.com/migrated/key_reform_pdfs/P15370_11_ROL_Model_Law_Computer_Related_Crime.pdf; United States Department of Justice, CLOUD Act Resources, available at: https://www.justice.gov/dag/cloudact

description from the UN itself noted the active, contested nature of the General Assembly session that saw the resolution initiating the UN AHC being adopted after extensive debate, several actively debated amendments, and a general concern at the rushed nature of the vote instead of global consensus.[727] The procedural sessions in 2021 also saw this confrontation, with several states noting that Vienna-based UN processes—particularly in the area of international criminal law matters—are normally advanced by consensus rather than a slim simple majority of UN member states. This was followed by states advancing several contested procedural motions to set in place more inclusive, transparent procedural rules for the UN AHC.[728]

A large amount of public discussion on this comes in the context of the UN cybercrime treaty process being initiated by Russia and ostensibly supported by China and other states with arguably authoritarian governments. While the UN AHC process does owe its origin to a resolution proposed and aggressively pushed forward by the Russian Federation, its initiation, discussions on its proposed substantive content, and its effect on global cybercrime and cyber-policy discussions must be examined beyond a narrow geopolitical framing.

It is a fact that all the current UN cyber processes owe their origin to the efforts initiated by Russia in the late 1990s to call for multilateral initiatives to combat what it called 'the criminal use of ICT'—by states and other actors—through the prism of safeguarding international peace and security. The Group of Governmental Experts (GGE) under the UN Security Council, the Open Ended Working Group (OEWG) under the UN General Assembly First Committee and the UN AHC deliberating the cybercrime convention have all begun under the initiative of, or specific resolutions advanced by, the Russian Federation.[729] This trend has only recently been broken, with successful votes in UN First Committee and General Assembly over November–December 2022 giving the go-ahead for the initiation of a UN Programme of Action (PoA) to advance responsible state behaviour in ICTs in the context of international security—which was originally

727 'General Assembly adopts resolution outlining terms for negotiating cybercrime treaty amid concerns over "rushed" vote at expense of further consultations', GA/12328, 26 May 2021, available at: https://press.un.org/en/2021/ga12328.doc.htm

728 Summer Walker, 'Contested domain: UN cybercrime resolution stumbles out of the gate', Global Initiative Against Transnational Organized Crime, 2 June 2021, available at: https://globalinitiative.net/analysis/un-cybercrime-resolution/

729 Elaine Korzak, 'Russia's cyber policy efforts in the United Nations', Talinn Paper no. 11, NATO CCDOE (2021), available at: https://ccdcoe.org/library/publications/russias-cyber-policy-efforts-in-the-united-nations/

initiated by France and Egypt, supported by EU member states and Canada and then approved by 150+ of the UN's member states.[730]

However, we must recognise that the initiator of the process is not the sole driver: starting a process is not the same as controlling it. Once initiated, a process is open to being steered or influenced by any state or group of states that bring their 'norm entrepreneur' energies and diplomatic channels to bear. Indeed, the UN cyber processes have borne that out. As delegates to these processes observe in discreet conversations, Russia initiated the UN GGEs but did not solely end up supporting that particular institutional process, and may therefore have ended up supporting the UN OEWGs thinking it would have more influence through a process open to all UN member states. And when conversations across different groups of countries break down, these processes can end in nearly universally recognised failure—as was the case with the UN GGE for 2015–2017.[731] While the second OEWG was quickly brought into place by Russia, demonstrating its pre-Ukraine conflict procedural prowess and UN political capital, its subsequent existence has shown that the rest of the UN's membership and competing geo-political/geographic 'blocs' do end up with a significant role in these processes. The sixth edition of the UN GGE—which did achieve consensus—was chaired by a Brazilian representative and saw active efforts by states beyond the superpowers. The UN OEWG has seen active involvement from a wide spectrum of states, with its first edition being chaired by a Swiss diplomat, its informal intercessional session led by a senior Singaporean cybersecurity official, and its deliberations seeing active involvement from states across regions and levels of development.

I would argue that this phenomenon of a Russia-triggered UN cyber process being significantly steered by other states is already holding true with the UN AHC cybercrime process—and perhaps this would have been the case even if the Ukraine war had not triggered a significant isolation of Russia within the UN and a reduction in its multilateral political capital and freedom of negotiating movement. The UN AHC's founding initial resolution may have been principally advanced by the Russian Federation, but its current rules of procedure were framed after active deliberations involving proposals and amendments coming from Brazil, Haiti (on behalf of the Caribbean Community—CARICOM) and the

730 'General Assembly adopts over 100 texts of First, Sixth Committees tackling threats from nuclear weapons, international security, global law, transitional justice', GA/12478, 7 December 2022, available at: https://press.un.org/en/2022/gadis3704.doc.htm

731 Michael Schmitt and Lois Vihul, 'International cyber law politicized: the UN GGE's failure to advance cyber norms', JustSecurity, 30 June 2017, available at: https://www.justsecurity.org/42768/international-cyber-law-politicized-gges-failure-advance-cyber-norms/

United Kingdom. The eventual final voting modalities of the UN AHC (a requirement of a two-thirds majority), regional balance in the meeting hosting (by shuttling the AHC meetings between New York, where all UN member states have representation, and Vienna, which has less representation from less prosperous states) and external participation (restrictions on the ability of states to block non-governmental participation, including civil society) were decided by these intervening proposals. It is clear that the UN AHC may have come from a Russian proposal, but its deliberations and output will be shaped by the states actively participating and canvassing support for their efforts.

What the UN AHC cybercrime discussions definitely showed, towards the conclusion of its three substantive sessions in 2022, is that there is significant interest in cross-border and potentially multilateral initiatives on cybercrime. States—and the officials who create their positions across different domestic agencies—are interested in a range of cybercrime-related problems and would like to see further innovation, international cooperation and collaboration, whether within the UN AHC processes or outside them. The negotiation of the Budapest Convention Second Additional Protocol demonstrated that, arguably, it came from clear interests from several nation states that the COE and Budapest treaty teams responded to, and likely accelerated in the end with a desire for it to be finalised prior to the UN AHC processes substantively commencing. I would submit that the complicated—and often controversial—measures on cross-border data sharing, joint investigations and the like included in the Second Additional proposal were accelerated by the COE and several of the convention's more active signatories in order to have the instrument out before the UN AHC process advanced further. In effect, the COE and several of the supporters of the Second Additional Protocol appeared to want to show it as an alternative to measures on law enforcement cooperation, data access and more to the states that were frustrated by the current status of global legal assistance on electronic evidence and cybercrime investigation issues, in order to reduce their support for a UN instrument on cybercrime advocated for by Russia and its allies.[732] Indeed, European Commission senior staff have stated as much before the European Parliament in hearings, indicating that their adamance on the EU

732 Katitza Rodriguez, 'EFF to Council of Europe: cross border police surveillance treaty must have ironclad safeguards to protect individual rights and users' data', Electronic Frontier Foundation Deeplinks, 8 September 2021, available at: https://www.eff.org/deeplinks/2021/09/eff-council-europe-cross-border-police-surveillance-treaty-must-have-ironclad

ratifying the Protocol was partly due to the UN cybercrime treaty deliberation process kicking off.[733]

We therefore must engage with international cybercrime-related legal harmonisation and cooperation discussions on a more substantive basis. And in that, I would argue that all stakeholders involved in international cybercrime-related negotiations must recognise that international legal frameworks in this area should not be based solely on responding to the 'lowest common denominator' negotiation approach to what national policymakers dealing with cybercrime wish to lock into multilateral processes. The government representatives for negotiating states—and the stakeholders engaged with advising and advocating before them—must ensure further developments in an international cybercrime legal framework helps improve the global situation, particularly around the security research community and human rights with respect to cybercrime laws and procedures.

Any international cybercrime legal harmonisation effort—especially a more high-profile one embedded within the UN system—has tremendous signalling and state practice-setting potential. Ensuring that it does not encourage the use of cybercrime laws as tools of political repression and digital authoritarianism is a global policy priority we should recognise. Of course, we must recognise the limits that any international legal instrument faces in driving domestic reforms when dealing with a particularly entrenched situation—even the most progressive international cybercrime treaty may have less effect than we would like in being a tool to reform problematic laws or instances of authoritarian state behaviour in the guise of combating cybercrime that we already see. It would, however, always play a key role in the international harmonisation of legal frameworks, at least *de jure*, and would play a key role in legitimisation of particular legal approaches and standards.

Therefore, whether for democratic or authoritarian states, I would argue there are two areas where we need more universal improvement and international refinement with regard to cybercrime legal frameworks. Stakeholders involved in drafting an international cybercrime legal instrument have a responsibility to:

733 Laura Kabekla, 'Controversy surrounds new cybercrime protocol as plenary vote still hangs in the balance', EURACTIV, 12 May 2022, available at: https://www.euractiv.com/section/data-protection/news/controversy-surrounds-new-cybercrime-protocol-as-plenary-vote-still-hangs-in-the-balance/

1. further a more robust global cybersecurity ecosystem, to ensure that legal frameworks facilitate—and do not chill—security research and are designed keeping in mind the humans who make cybersecurity possible;
2. ensure strong safeguards on new international structures on law enforcement cooperation, cross-border investigatory powers and international criminal law provisions governing access and retention of data.

International cybercrime legal harmonisation: an opportunity for reform or focus on the responsibility to prevent further harm?

International cybercrime cooperation and improved cross-border processes could represent a step up, a shift in the paradigm that enables or encourages increased reform in national legal frameworks that are yet to reach the identified ideal standards, whether in explicit mandates (states must do X reforms in domestic law and practice) or via incentives or nudges (if states make sure to do Y, they will receive benefits Z). An international cybercrime instrument could also cause harm, by allowing the spread of less well considered standards or weak institutional procedures internationally. This would require us to work to make sure that an international cybercrime instrument does not cause more harm—whether by advancing problematic standards, lowering global standards/undermining safeguards, or being used to legitimise existing problematic legal provisions and government practices by certain states.

The question addressed by many actors involved in current global cybercrime policy discussions (including technical experts, academics, civil society and government advisers) at the UN AHC is whether the focus should be on solely on preventing the legitimisation of harmful national practices through a UN instrument and reducing opportunities for inflicting more harm (i.e. preventing an even worse situation from coming about for domestic stakeholders—especially journalists, human rights defenders and vulnerable communities—because

of multilateral discussions), or whether there are positive objectives that could be achieved.

I argue that the answer is that both are required. An international cybercrime legal instrument negotiated as a UN treaty will be at least a landmark moment. If signed up to by the vast majority of UN states, it could be transformative—an international legal framework that binds together more states on this topic than before,[734] and possibly one where new structures and commitments on cybercrime legal matters are being secured. More states than ever before may agree on criminalising a range of activities online, and set in place how they will cooperate on matters pertaining to investigation, data access and sharing, technical assistance and more. All these issues would have significant effects on a variety of stakeholders in the wider cybersecurity ecosystem and on individuals across the globe—including possible new avenues for intrusion and harm by inadvertent or intentional governmental action.

The direct potential for cybercrime laws to be used as tools for content regulation, targeting of activists and persecution of those who dissent is clear. Over the past few years, there has been an ever-expanding set of examples of governments using cybercrime laws as a key part of their digital authoritarian toolkit.

A few useful—though far from exhaustive—illustrations of this are as follows:

> Bangladesh's Digital Security Act: brought forward and passed into law after an already problematic ICT Act had been regularly used to target critics and activists, the Digital Security Act created a set of new, vague cybercrime offences and gave birth to a new Digital Security Agency that from the beginning appeared not just to be focused on technical assurance in cybersecurity matters or the prosecution of cyber-dependent crime, but also had a broad remit to investigate (including via warrantless search and seizure) and punish 'digital offences'. The initial strong criticism of the law by scholars, activists, journalists and representatives of several other governments has proved well founded in its implementation: its content criminalisation provisions and wide set of powers to the Digital Security Agency to inquire into and seek punishment of individuals

734 The Council of Europe Budapest Convention currently has 67 signatories, out of the 193 current UN member states.

under the statute have been regularly used to undermine human rights and penalise critics and human rights defenders.[735]

- Kenya's Computer Misuse and Cybercrimes Act, 2018: the Kenyan law included a broad criminalisation approach, creating a range of offences focused not just on core cyber-dependent crimes but also on a wide range of cyber-enabled crimes—including provisions around false publications. It also created criminal liability for not reporting cyber-attacks or threats within a specified period, and provided for a dangerous over-broad set of government procedural and investigative powers without sufficient independent oversight.[736]
- Syria's new 2022 cybercrime law, i.e. Law 20/2022: explicitly drafted with the aim of curbing the 'misuse of technology', the law advanced by al-Assad's government created an extremely vague set of crimes—including activities relating to decency or modesty and activities 'undermining prestige' as cybercrimes. Criminal penalties for the same activities carried out through online channels were enhanced, and, in the case of the provisions relating to online slander, provided for enhanced penalties if the activities were directed at public employees—demonstrating a clear attempt to muzzle criticism. Other alarming provisions included an obligation on service providers to store and retain data, and penalties if providers failed to identify persons responsible for posting online content.[737] A similar alarming approach to criminalising political

735 Ali Riaz, 'How Bangladesh's Digital Security Act is creating a culture of fear', Carnegie Endowment, 9 December 2021, available at: https://carnegieendowment.org/2021/12/09/how-bangladesh-s-digital-security-act-is-creating-culture-of-fear-pub-85951; 'Legal Analysis – Bangladesh: Digital Security Act 2018', Article 19, November 2019, available at: https://www.article19.org/wp-content/uploads/2019/11/Bangladesh-Cyber-Security-act-2018-analysis-FINAL.pdf; Rokeya Lita, 'Bangladesh's Digital Security Act is criminalising journalism', Al Jazeera Journalism Review, 18 April 2022, available at: https://institute.aljazeera.net/en/ajr/article/1872; Access Now, 'New Digital Security Act in Bangladesh deepens threats to free expression', 21 September 2018, available at: https://www.accessnow.org/new-digital-security-act-in-bangladesh-deepens-threats-to-free-expression/; Access Now and 14 other organisations, 'Bangladesh: release Nusrat Shahrin Raka, sister of Bangladeshi journalist Kanak Sarwar', 28 January 2022, available at: https://www.accessnow.org/joint-letter-bangladesh-release-nusrat-shahrin-raka/

736 Mercy Muendo, 'Kenya's new cybercrime law opens the door to privacy violations, censorship', The Conversation, 29 May 2018, available at: https://theconversation.com/kenyas-new-cybercrime-law-opens-the-door-to-privacy-violations-censorship-97271; Article 19, 'Legal Analysis – Kenya: Cybercrime and Computer Related Crimes Bill', February 2018, available at: https://www.article19.org/wp-content/uploads/2018/02/Kenya-Cybercrime-Bill-129072014-BB.pdf

737 Marwa Fatafta, 'Syria's new "cybercrime" law adds salt to injury', Access Now, 27 May 2022, available at: https://www.accessnow.org/syria-cybercrime-law/

criticism and dissent can be seen in the UAE's Federal Decree Law No. 34 of 2021 on Combatting Rumours and Cybercrime.[738]

> The Ecuadorian Criminal Code (Article 232): the prosecution of Ola Bini, the Swedish open-source developer and activist, is an ongoing international case study on how the Ecuadorian provision criminalising 'unauthorised access to a computer system' is alarmingly over-broad and being actively misused. The language of the provision was sufficiently broad that prosecutors felt comfortable in initiating a now years-long prosecution involving an intrusive raid and detention (till the present day) of Ola Bini based on his disclosure of a Telnet request for connection to an open server. The case has been criticised by human rights groups and digital security practitioners globally, and described as a 'hacker panic' case involving fear around the perceived information and capabilities that the InfoSec community holds.[739] In February 2023, the Ecuadorian court trying Ola Bini came to a unanimous ruling upholding his innocence,[740] though it is unclear if Ecuadorian authorities will appeal the ruling.[741]

738 Access Now and 14 other organisations. 'Joint statement on the UAE's adoption of Federal Decree Law No. 34 of 2021 on Combatting Rumours and Cybercrime', 25 January 2022, available at: https://www.accessnow.org/cybercrime-law-uae/

739 Carlos E. Flores, 'Ola Bini, the cyberactivist who causes panic in Ecuador', Global Voices, 21 October 2022, available at: https://globalvoices.org/2022/10/21/ola-bini-the-cyberactivist-who-causes-panic-in-ecuador/; Danny O'Brien, 'Telnet is not a crime: unconvincing prosecution screenshot leaked in Ola Bini case', Electronic Frontier Foundation, 23 August 2019, available at: https://www.eff.org/deeplinks/2019/08/telnet-not-crime-unconvincing-prosecution-screenshot-leaked-ola-bini-case; Jason Kelley and Veridiana Alimonti, 'EFF and other civil society organizations issue report on danger to digital rights in Ola Bini trial', Electronic Frontier Foundation, 9 May 2022, available at: https://www.eff.org/deeplinks/2022/05/eff-and-other-civil-society-organizations-issue-report-danger-digital-rights-what; Gaspar Pisanu, 'Join our Statement for the Protection of Digital Rights Defenders', Access Now, 18 December 2019, available at: https://www.accessnow.org/join-our-statement-for-the-protection-of-digital-rights-defenders/

740 People's Dispatch, 'Digital Rights Activist Ola Bini Declared Innocent by Ecuadorian Court', 1 February 2023, available at: https://www.newsclick.in/digital-rights-activist-ola-bini-declared-innocent-ecuadorian-court

741 ARTICLE 19, 'Ecuador: Ola Bini innocent verdict must lead to stronger digital rights', 7 February 2023, available at: https://www.article19.org/resources/ecuador-ola-bini-innocent-verdict-must-lead-to-stronger-digital-rights/.

Excessively broad cybercrime legal provisions impact cybersecurity research and result in more instability

In the UN AHC cybercrime process, several national delegates have regularly spoken on the need to ensure that conversations on cybersecurity, national security and cybercrime are kept separate. What they are saying in effect is 'Keep counter-terrorism and ICT, cybersecurity norms and responsible state behaviour, and cybercrime law harmonisation separate from each other.'

However, advancing robust cybersecurity is a key reason for and corollary of effective legal and policy frameworks to address cybercrime. I would therefore argue that coordination on countering cybercrime and coordination on cybersecurity policy are, and always will be, intertwined. There may be a UN First Committee-anchored process in the form of the OEWG that focuses on state cyber behaviour, while the UN Third Committee-initiated AHC process focuses on international cybercrime law harmonisation. While we can define specific and narrow objectives for these different processes—recognising the challenges of the substantive topics themselves and of securing political consensus—we cannot say that international efforts to combat cybercrime should be divorced from states' efforts to secure more agreement and progress on a human-centric approach to cybersecurity. Any strict effort to separate these in the realm of policy discussions would risk creating contradictory standards or even allow for the risk of double standards in state commitments (forum shopping).

Because of multilateral politics, there is an understandable reluctance to explicitly state this in UN discussions. That said, I propose that we must recognise the following imperatives:

- We want secure, reliable networks and digital services and tools.
- We prosecute or seek to deter those who are active threats to that.
- We also need to encourage and support those who help counter bad actors and bring vulnerabilities to attention. This requires us to recognise that their efforts involve ways of working involving ethical approaches to hacking, vulnerability disclosure and related techniques, which cannot be reconciled with a blanket ban on all

forms of intrusion into devices and over-broad criminalisation of 'hacking'.

We need a systemic approach to protecting and advancing global cybersecurity instead of ad-hoc and unbalanced approaches that seek tightening on only one element of cybersecurity to the detriment of others. Conservative approaches from states at the international level may not be as well advised as they think if they fail to acknowledge the reality of how individuals play a key role in advancing cybersecurity, particularly in security research and vulnerability disclosure. And the cracks have been showing at the domestic level for some time, even on criminalising as basic a core 'cybercrime' as that of 'hacking', i.e. unauthorised access to computers and/or ICT systems.

Indeed, one could argue that cybercrime laws are the nearly universal original sin of computer abuse. From the Computer Fraud Abuse Act (CFAA) in the US onwards, domestic cybercrime/computer crime laws include core crimes around unauthorised access or hacking that are too broadly worded. This broad wording and corresponding wide, varying enforcement have had a chilling effect on cybersecurity research, which, after all, relies on hackers testing, probing, penetrating and tinkering with systems in order to see whether security vulnerabilities exist that should be reported and fixed.[742]

Requiring universal criminalisation of unauthorised access into ICT systems and other related 'core' cybercrimes without putting in place safeguards, building in lessons learned from the (mis)application of cybercrime and computer abuse laws against hackers engaged in legitimate security research, would be an extraordinary missed opportunity. We see a partial recognition of the importance of the issue in the UN AHC process, with the chair and secretariat to the process posing specific questions to UN member states on their views regarding how the criminalisation portion of the proposed treaty should be addressed.[743]

National governments have sought to address this issue in domestic policy and legal practice. The Dutch government is well known for its disclosure guidelines initiative since 2013 (revised in 2018) and its policy indicating a commitment on the part of the prosecution service not to prosecute those following

742 James Conrad, 'Seeking help: the important role of ethical hackers', *Network Security* 2012 (8) 2012, 5–8.

743 Letter from the Chair of the Ad Hoc Committee, including guiding questions, 25 May 2022 (Questions I.A.1, I.B.2), available at: https://www.unodc.org/documents/Cybercrime/AdHocCommittee/Second_session/AHC_2nd_session_-Guiding_questions_criminalization.pdf

ethical vulnerability disclosure standards.[744] The US has seen repeated calls for reform measures addressing the wide reach of the CFAA and its impact on the security research community.[745] These have included calls for amendment of the CFAA's current criminal provisions. Another proposal has been to have a standalone exemption provision that would immunise legitimate security research from prosecution or other over-broad legal claims: in effect a security researcher's 'safe harbour' provision in the US CFAA.[746]

In May 2022, the US Department of Justice announced a new policy for how it will seek to charge cases under the CFAA. It said the policy directed that, for the first time, good-faith security research should not be charged, and defined this as 'accessing a computer solely for purposes of good-faith testing, investigation, and/or correction of a security flaw or vulnerability, where such activity is carried out in a manner designed to avoid any harm to individuals or the public, and where the information derived from the activity is used primarily to promote the security or safety of the class of devices, machines, or online services to which the accessed computer belongs, or those who use such devices, machines, or online services'. [747] Notably, this reform measure was advanced as an executive policy or guidelines announcement—not the binding statutory amendments to reform the CFAA that advocates and the information security community have been asking for.[748]

However, we must engage with the fact that there appears to be a temptation for states to suggest that they will handle safeguards for security researchers at a purely domestic level, and not adopt enhanced standards for criminal intent (in the form of requirements for 'dishonest' or 'malicious' intent) or a standalone

744 'Coordinated Vulnerability Disclosure: the Guideline', 2 October 2019, available at: https://english.ncsc.nl/publications/publications/2019/juni/01/coordinated-vulnerability-disclosure-the-guideline; K. Clark, D. Stikvoort, E. Stofbergen and E. van den Heuvel, 'A Dutch approach to cybersecurity through participation', *IEEE Security & Privacy* 12 (5) (2014), 27–34.

745 Electronic Frontier Foundation, 'The Computer Fraud and Abuse Act hampers security research', available at: https://www.eff.org/document/cfaa-and-security-researchers; Riana Pfefferkorn, 'America's anti-hacking laws pose a risk to national security', Brookings TechStream, 7 September 2021, available at: https://www.brookings.edu/techstream/americas-anti-hacking-laws-pose-a-risk-to-national-security/

746 Daniel Etcovitch and Thyla van der Merwe, 'Coming in from the cold: a safe harbor from the CFAA and DMCA §1201', Berkman Klein Center, available at: https://cyber.harvard.edu/publication/2018/coming-cold-safe-harbor-cfaa-and-dmca-ss1201

747 Office of Public Affairs, United States Department of Justice, 'Department of Justice announces new policy for charging cases under the Computer Fraud and Abuse Act', 19 May 2022, available at: https://www.justice.gov/opa/pr/department-justice-announces-new-policy-charging-cases-under-computer-fraud-and-abuse-act

748 Harley Geiger, 'Proposed security researcher protection under CFAA', Rapid7, 4 June 2021, available at: https://www.rapid7.com/blog/post/2021/06/04/proposed-security-researcher-protection-under-cfaa-2/; Riana Pfefferkorn, 'The importance of protecting good-faith security research', Center for Internet and Society at Stanford Law School, 14 September 2020, available at: https://cyberlaw.stanford.edu/blog/2020/09/importance-protecting-good-faith-security-research

provision to require states to exempt the activity of legitimate security research. This appears to be the more conservative approach that several key states are taking in the initial discussions around security researchers in the UN AHC process; for example, despite its domestic policy changes, the US delegation to the AHC said that it was unsure of supporting a standalone exception for security researchers in the treaty and did not want to place a heightened intent requirement on the crime of unauthorised access in the proposed treaty.

Such an approach would be dangerous; we may end up in a situation where an international treaty on cybercrime explicitly requires the universal criminalisation of network and device intrusion by states while not placing any pressure on them to provide legal certainty to legitimate security research. This would be forcing universal criminalisation, but not learning from the mistakes and legal innovation from several UN member states.

Failing to provide legal protection for security researchers—whether in the form of heightened intent requirements for core cyber-dependent crimes (particularly unauthorised access) or a standalone legitimate security research 'safe harbour' mandate—would be a mistake that the global information security community can ill afford. Indeed, the statements by several national delegates to the UN AHC were telling—showing that suspicions of the information security community as reckless hackers continue, despite increased global governmental recognition of working with cybersecurity talent. And statements that security researchers had nothing to fear if their activities were 'authorised' by the parties beforehand betray a lack of understanding of how hackers find security vulnerabilities—especially models that rely on bug bounties or responsible vulnerability disclosure.[749]

I would argue that it is even more important that we ensure that reduction in legal uncertainty and chilling effects on legitimate security research be an international policy goal—and not merely a domestic goal left to states to manage on a best-efforts basis. Much of modern security research—especially the sector of bug bounties—is now international and conducted regularly across borders. A failure to act on this would be a significant global missed opportunity.[750]

749 Most state viewpoints on this were presented in the second session of the AHC in Vienna over 30 May–10 June 2022, clustered in the discussions on the criminalisation provisions on 30 May–2 June. States that took a view that legitimate security research would involve an element of permission or authorisation included Nigeria, Peru, Ghana and Indonesia. Some states indicated that qualifications or recognition of individuals as cybersecurity researchers could be a prerequisite for them to receive legal protections—an example was the Philippines.

750 Ryan Ellis and Yuan Stevens, 'Bounty everything: hackers and the making of the global bug marketplace', Data & Society, available at: https://datasociety.net/library/bounty-everything-hackers-and-the-making-of-the-global-bug-marketplace/

It would have been reasonable to assume that the debate on legal protections for security researchers, vulnerabilities disclosure and penetration testing would have resulted in an initial negotiating text that carried different configurations attempting to provide at least a baseline form of protection. After all, the specific discussion on this subject in the May–June second session of the AHC saw several delegations—major players and intergovernmental organisations, as well as smaller states—agree that security researchers deserved some form of legal protection in the treaty. The main disagreement appeared to be the form in which their legal position could be safeguarded—some states actively raised concerns regarding a security research exception or 'safe harbour', but indicated they were flexible on negotiating alternative legal mechanisms. Several specifically indicated that they believed that the language on intent in the provisions regarding unauthorised access and related parts of the criminalisation chapter could be drafted with higher requirements that would protect security researchers and other public-interest-related classes, such as human rights defenders and journalists.[751]

Unfortunately, the first consolidated negotiation document text prepared by the AHC chair and support team does not build strongly enough on this emerging international consensus to provide at least some form of improved legal certainty and protection for security research.[752] In the criminalisation chapter of the consolidated negotiating document, the language proposed for further requirements on intent in the proposed core cybercrimes is only recommendatory: 'A State Party may require that the offence be committed.' Draft Article 6 relating to illegal access to a computer or ICT system does not make criminal intent a mandatory element; it is prefaced with the previously mentioned 'A State Party may' language, which is also in its formulations, and does not place a strong enough emphasis on criminal intent as a prerequisite.[753] Neither does it require the establishment of harm as a condition—harm is only listed as a ground for a state party to impose an aggravation of penalty. The failure to put in clearer, heightened

751 Examples of such states in the second session included Argentina, Japan, France, Brazil, New Zealand, Australia, Slovenia, Israel, Brazil, Thailand and Oman (and the EU); there were broader statements of support from China, Jordan, Haiti, Algeria and El Salvador.

752 UN AHC, 'Consolidated negotiating document on the general provisions and the provisions on criminalization and on procedural measures and law enforcement of a comprehensive international convention on countering the use of information and communications technologies for criminal purposes', 7 November 2022, A/AC.291/16, available at: https://www.undocs.org/A/AC.291/16

753 'A State Party may require that the offence be committed by infringing security measures, with the intent of obtaining [computer data] [electronic/digital information] or other criminal intent, or in relation to [a computer system] [an information and communications technology system/device] that is connected to another [computer system] [information and communications technology system/device].'

language around intent even as a negotiating baseline option is alarming—particularly since there is no general provision providing an exception or safe harbour from criminal liability for security researchers (or for journalists and human rights defenders). This requires priority attention by delegations and stakeholders to the AHC across its 2023 sessions, and a clearer commitment by the UN AHC chair in negotiating text options with delegates.

Indeed, it is useful to adopt and adapt some of the framing that was successfully negotiated in the UN OEWG process around an international commitment to a human-centric approach to global cybersecurity. I propose that a corollary is that legal harmonization of international cybercrime must not harm the human beings who make more effective, resilient cybersecurity possible. Cybercrime harmonisation and international enforcement must not make us more cyber-insecure. This approach appears to have received at least some explicit recognition by delegations to the AHC—the Australian delegation in its 31 May 2022 remarks on the discussion around criminalisation and security research noted the importance of AHC participants' ensuring that the proposed convention did not hamstring those who are often the first line of defence against cybercrime and other online threats.

Ensuring global cyber-coordination helps to further respect for privacy and protected human rights

A key imperative for the UN AHC, as well as other international legal discussions around cybercrime collaboration, is to facilitate increased data preservation, disclosure and communications intercept sharing among law enforcement actors across different countries.

States recognise the sensitivity of this issue, but also appear to agree that it is an imperative that any future cybercrime treaty must address. Interventions from several state delegations in the second and third substantive sessions of the UN AHC in May–June 2022 in fact have established that there is support for international collaboration on law enforcement cooperation provisions to apply to not just the core cybercrimes criminalised by the proposed treaty, but a much wider set of criminal-law-related tasks.

The drafting and ratification of the Budapest Convention's Second Additional Protocol on enhancing cooperation and disclosure of electronic evidence provides a useful experience to learn from. The COE secretariat/drafting team and key states argued that they could not intervene to push for improvements in oversight, safeguards in lawful interception and data access in the disparate Budapest Convention signatories with their varying legal traditions. The argument that carried the day was that it was better for a lowest common denominator approach to be taken, with a general baseline requirement around data protection standards, rather than asking all states to reach up to a level of judicial authorisation and supervision of such powers.

But actually what was happening was a new moment, something that did not exist before. Earlier conflict of laws and the presence of tech firms outside their national borders constituted *de facto* impediments to states, now changed by *de jure* innovation. I propose that if creating something new in the space of government legal powers to make access to protected information possible, you can—and should—gatekeep it. Indeed, the rush to finalise an agreed text to the Second Additional Protocol, despite the concerns of civil society, privacy experts and others, is impacting its uptake and effectiveness as an instrument—and triggered attempts in the EU Parliament to block the EU from acceding to the treaty due to these concerns.[754] It is crucial that stakeholders to the UN AHC learn from the problems of the Second Additional Protocol and avoid rushing to agree on text regarding information sharing, data access and joint investigative efforts without putting in sufficient safeguards to satisfy human rights legal standards and to assure stakeholders that strong checks exist against misuse.

When creating something new and allowing further unprecedented global access to protected information and sensitive data, states (and their negotiating delegations) are also under an obligation to enhance global standards by ensuring that an appropriate level of safeguards is included. An international legal framework to enable such increased data sharing and cross-border investigatory powers also requires participating states to meet a higher legal standard—particularly with respect to judicial authorisation of such processes and related procedural safeguards.

The AHC Chair's consolidated negotiating document contains a basic step in this regard that requires further evolution and expansion. Besides proposing a

754 European Digital Rights, 'Civil society warn against rushed global treaty for intrusive cross-border police powers', 8 June 2021, EDRi, available at: https://edri.org/our-work/civil-society-warn-against-rushed-global-treaty-for-intrusive-cross-border-police-powers/; Kabelka (see note 11 above).

provision that would require state parties to implement their obligations under the convention in accordance with applicable human rights law, and that any person prosecuted for offences established under the convention receives human rights law provided rights and guarantees,[755] it proposes a dedicated clause on conditions and safeguards on procedural measures and law enforcement cooperation under the convention. The current text requires implementation subject to domestic law safeguards that must incorporate proportionality, necessity and legality as well as protecting privacy and personal data. Importantly, the text also mandates that such safeguards include judicial or other independent supervision, as well as requiring justification of and limitations on such powers or procedures.[756] This is a start, seemingly incorporating at least part of the approach recognised in international human rights law standards such as the Necessary and Proportionate Principles, as well as their detailed Universal Implementation Guide.[757]

An additional innovation that AHC delegations should consider is mandating a form of 'transparency reporting' on state parties' usage of such procedural measures and cross-border law enforcement under the proposed convention. This can be an adaptation of the now accepted technology sector practice of transparency reports regarding government requests around user data and content regulation. State parties should document and summarise their use of measures and cross-border cooperation, and share this on a regular basis with whichever body helps support the review and implementation of the convention—whether the United Nations Office on Drugs and Crime (UNODC), an assembly of state parties or something else. The regular reporting and publication of information regarding such measures and cross-border law enforcement cooperation in the space can help stakeholders determine trends and bring possible abuse to light—but also, importantly, demonstrate where this new international legal framework is

755 Article 5(1), Article 39(5)

756 'Article 42. Conditions and safeguards – 1. Each State Party shall ensure that the establishment, implementation and application of the powers and procedures provided for in this chapter are subject to conditions and safeguards provided for under its domestic law, which shall provide for the adequate protection of human rights and liberties, including rights and fundamental freedoms arising from its obligations under applicable international human rights law, and which shall incorporate the principles of proportionality, necessity and legality and the protection of privacy and personal data. 2. Such conditions and safeguards shall, as appropriate in view of the nature of the procedure or power concerned, inter alia, include judicial or other independent supervision, grounds justifying application, and limitation of the scope and the duration of such power or procedure.'

757 The International Principles on the Application of Human Rights to Communications Surveillance (the 'Necessary and Proportionate Principles' or '13 Principles'), May 2014, available at: https://necessaryandproportionate.org/; Access Now, 'Universal implementation guide for the International Principles on the Application of Human Rights to Communications Surveillance', July 2015, available at: https://www.accessnow.org/cms/assets/uploads/archive/docs/Implementation_guide_-_July_10_print.pdf

working and where gaps exist, helping to bring data to an otherwise often politically charged conversation around cross-border legal assistance and digital jurisdiction.

Conclusion

A new international legal arrangement is a critical moment—in its signalling across all members of the international community as well as the explicit mandates it imposes on its signing and ratifying member states and the stakeholders who reside in them. The UN AHC process must be engaged with not with a defensive strategy in mind, given its misplaced origin and framing, but in a proactive way that recognises its transformative, signalling potential. AHC participants not only must protect cybersecurity stakeholders and safeguard human rights—they have a responsibility to help improve the situation regarding the impact of cybercrime law on human rights and on security research.

Combating global cybercrime by enhancing cross-border cooperation and coordinating a more harmonised legal framework across states should not make us more cyber-insecure. AHC delegates need to do more to ensure that they proactively reduce the legal uncertainties triggered by cybercrime provisions for the human beings who make global cybersecurity possible—security researchers, digital security trainers and the wide responsible/ethical InfoSec community). Improved standards around criminal intent and harm, or other specific legal mandates around not criminalising or prosecuting legitimate security research, must be advanced—at least as a corollary to any further internationally harmonised criminalisation of unauthorised access and related core cyber-dependent crimes.

In the rush to address the political discontent around the broken mutual legal assistance system in our digital age, AHC delegates must do more to ensure safeguards and oversight on cross-border data access and sharing. Mistakes were made in the COE Second Additional Protocol that must not be replicated in a UN treaty on cybercrime—especially given the even wider membership of the latter and the increased variance in human rights standards and 'like-mindedness' among states in terms of upholding them when exercising cybercrime-related data access powers and procedural measures.

ANNEX

About the contributors

Aditi Bawa is a program coordinator in the Technology and International Affairs program at the Carnegie Endowment for International Peace. She works on projects related to cybersecurity capacity building, digital financial services, and the security of digital financial inclusion.

Luca Belli is a Professor at Fundação Getulio Vargas (FGV) Law School, Rio de Janeiro, where he heads the Center for Technology and Society (CTS-FGV) and the CyberBRICS project. He is also an Associated Researcher at the Centre de Droit Public Comparé of Paris 2 University, Member of the Board of the Alliance for Affordable Internet (A4AI), and Director of the Latin-American edition of the Computers Privacy and Data Protection conference (CPDP LatAm).

Jaime Bello is an international lawyer at CMS Albiñana & Suárez de Lezo in Madrid, Spain. He specialises in legal issues related to cyber, digital and disruptive technologies, and has advised the most innovative companies around the world for more than a decade, including reacting to cyber attacks and data breaches. He is also an expert in geopolitics and global affairs, with a special interest in the MENA region and a focus on analysing the impact of technology on public policy and international relations. He has published several articles in this field on specialised websites.

Dennis Broeders is Full Professor of Global Security and Technology at the Institute of Security and Global Affairs (ISGA) of Leiden University, the Netherlands. He is the Senior Fellow of The Hague Program on International Cyber Security and project coordinator at the EU Cyber Direct Program. His research and teaching broadly focuses on the interaction between security, technology and policy, with a specific interest in international cyber security governance. He is the author of the book *The public Core of the Internet* (2015). He served as a member of the Dutch delegation to the UN Group of Governmental Experts on international information security and the Open Ended Working Group (2019-2021) as an academic advisor. Before joining Leiden University he was Professor of Technology and Society at Erasmus University Rotterdam and senior researcher and project coordinator at the Netherlands Scientific Council for Government Policy, a think tank within the Dutch Prime Minister's office.

Enrico Calandro is the Project Leader of the EU-funded project Cyber Resilience for Development (Cyber4Dev). In that role, he leads the strategic direction of the project. Over the last fifteen years, Enrico has been working at the intersection of research, international cooperation, digital policy, and development as a research associate at Research ICT Africa. As co-founder and co-director of the Cybersecurity Capacity Centre for Southern Africa at the University of Cape Town, he has led efforts to implement the Cyber Maturity Model for Nations and raise cybersecurity awareness in the African region. He consulted on digital transformation, cybersecurity, and digital policy in Africa and more broadly in the global South for the UNDP Global Centre for Technology, Innovation, and Sustainable Development, the UNDP Office of the Human Development Report, the UK FCDO, and GIZ.

Raman Jit Singh Chima is Senior International Counsel and Asia Pacific Policy Director at the international non-profit organisation Access Now, where he also serves as Global Cybersecurity Lead. He co-founded India's SaveTheInternet.in campaign for net neutrality, and later helped co-found the Article 21 Trust and the Internet

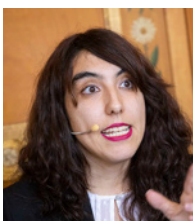
Freedom Foundation. In 2016, he was included in *Forbes Magazine*'s 30-Under-30 list of leaders in India under the Law and Policy category. He previously served as Policy Counsel and Government Affairs Manager in Google's Asia Pacific team and held senior roles in Indian industry bodies. In 2019, he was awarded a Chevening Cybersecurity Fellowship, completing a programme of study at the Defence Academy of the United Kingdom. Raman graduated from the National Law School of India University at Bangalore, where he co-founded and was later elected chief editor of the *Indian Journal of Law and Technology*.

Yasmin Curzi de Mendonça is a researcher at the Center for Technology and Society of the FGV Law School in Rio de Janeiro where she also works as Assistant Professor of Human Rights and Transnational Law. She is pursuing her PhD in Social and Political Studies at the Rio de Janeiro State University, holds a Master's degree in Social Sciences from the Pontifical Catholic University of Rio de Janeiro and Bachelor Degrees in both Law and Social Sciences from FGV Rio, with an exchange period at the Université Paris-Sorbonne. Yasmin also coordinates the Dynamic Coalition on Platform Responsibility at the UN Internet Governance Forum and is a member of the Steering Committee of the Platform Governance Research Network. Her main research interests are the impacts of technology in human rights, especially regarding gender-based violence.

Dr François Delerue is an Assistant Professor of Law at IE University. He is also an Associate Fellow of The Hague Program on International Cyber Security (Leiden University) and the GEODE Center (Paris 8 University). Additionally, he is a Co-chair of the Committee on Digital Challenges for International Law for the preparation of the 150th Anniversary Conference of the International Law Association. Previously, he was a Senior Researcher in Cybersecurity Governance at the Institute of Security and Global Affairs at Leiden University and Work Package Leader on International Law for the project EU Cyber Direct. His book *Cyber Operations and International Law* was published by Cambridge University Press in 2020 and was awarded the 2021 Book Prize of the European Society for International Law.

Mabda Haerunnisa Fajrilla Sidiq is a researcher at the ASEAN Studies Program, The Habibie Center. She is also a Cyber Diplomacy Atlas curator for Southeast Asia at EU Cyber Direct. Her research interest primarily lies at the intersection between science and technology and International Relations, particularly on global Internet governance, as well as on international politics in Southeast Asia. She previously worked as a research and teaching assistant at the Department of International Relations, Universitas Indonesia and an editorial secretary of *GLOBAL: Jurnal Politik Internasional*. Prior to working for The Habibie Center, Mabda concluded her master's degree in International Relations (Research) at the London School of Economics and Political Science and her bachelor's degree in the same subject at Universitas Indonesia.

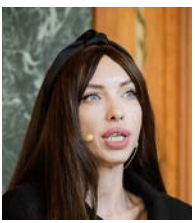
Walter B. Gaspar is a researcher at the Center for Technology and Society at Fundação Getulio Vargas' Law School in Rio de Janeiro, Brazil. His research at the Center focuses on data protection and cybersecurity matters in Brazil and other BRICS countries. He holds a Master's degree in Public Health from the Rio de Janeiro State University, focusing on the interface between Brazilian innovation and intellectual property systems. Currently, he is a PhD student in the Public Policies, Strategies and Development programme of the Economics Institute at the Federal University of Rio de Janeiro, studying innovation systems surrounding quantum technologies. Walter is the executive director of CPDP LatAm, co-authored "What is Creative Commons? New copyright models in a more creative world," was the National Coordinator of the Brazilian chapter of the NGO Universities Allied for Essential Medicines and has worked on many research projects concerning intellectual property, innovation and data protection.



Maria Pilar Llorens is a lecturer at the National University of Cordoba and a Postdoctoral Fellow of the National Council of Scientific and Technical Research (CONICET) of Argentina at the Legal and Social Research Centre of the Faculty of Law of the Universidad Nacional de Córdoba, Argentina. Her interest lies in the intersection of archival and doctrinal research from a Third World Perspective of

International Law. Her current research focuses on international law's application in/to cyberspace, with an interest on cyber-norms development and also on sovereignty. Previous research explored the influence of Argentinian scholars in the formation of international norms. She teaches in the field of international law, specifically International Public Law and Inter-American System of Human Rights. She is a member of the Latin American Cybersecurity Research Network (LA/CS Net) and of the Argentinean Association of International Law.

Moliehi Makumane is a researcher at the UN Institute for Disarmament Research (UNIDIR). She served as an expert on the UN Group of Governmental Experts on international information security (2019-2021) and as a member of the South African delegation to the Open-Ended Working Group (2019-2021). Her research is broadly on the Framework for responsible behavior on cyberspace, specifically approaches and tools for developing countries. Before joining UNIDIR, Ms. Makumane was a foreign service officer at the Department of International Relations and Cooperation, South Africa.



Andreja Mihailovic, PhD, is a lecturer, researcher, and consultant in the fields of business and criminal law, with a particular interest in cybersecurity, cybercrime, digital investigation, electronic evidence, cyber diplomacy, and innovation and technology policies. She works as a teaching assistant for the criminal law group of subjects at the Faculty of Law of the University of Montenegro, where she is involved in the development of multidisciplinary information security study programs, as well as various cyber-oriented initiatives addressing students' and the ICT market's contemporary needs. As a holder of a British Scholarship Trust scholarship, she was a PhD researcher at the University of Birmingham (UK) in 2017, where she gained advanced training in cybersecurity and digitalization. She is an authorized trainer for the Montenegrin Human Resources Administration's "cybersecurity" training programs for civil servants and national expert for evaluating digitalization skills and competencies for policymakers in government departments, agencies, and public bodies in Montenegro.

Evgeni Moyakine is an Assistant Professor IT law at the Faculty of Law of the University of Groningen, the Netherlands. He is also an Associate Member of the Security, Technology & e-Privacy (STeP) Research Group of the same university and a former Research Fellow at the Center for Cyber Law & Policy (CCLP) of the University of Haifa, Israel. His research interests include international/European law, IT law, cybersecurity, privacy and data protection. He has authored and co-authored scientific publications in the above-mentioned fields, has obtained several research grants and has been involved in various national and international research projects. At the University of Groningen, he has developed as a lead educator together with his team a massive open online course 'Understanding the GDPR' for the public at large and teaches both Bachelor's and Master's courses, including 'Telecommunications Law', 'Security & Privacy in Digital Society' and 'Accountability in the Cyberspace Era'.

Babatunde Okunoye is a Researcher affiliated with the Department of Communication and Media Studies, University of Johannesburg. He is also a Research Affiliate with the Berkman Klein Center for Internet and Society at Harvard University. His research focuses on Information Communications Technologies for Development (ICT4D), particularly in the context of the Global South.

Francesca Romana Partipilo is a PhD Candidate in Human Rights and Global Politics at the Sant'Anna School of Advanced Studies. She is a Guest Doctoral Researcher at the University of Stockholm, and previously spent six months as Visiting Researcher at Utrecht University, working with Prof. Seline Trevisanut. Her research focuses on the involvement of non-state actors in search and rescue operations at sea. She previously worked at the British Institute of International and Comparative Law and volunteered at the British Refugee Council. She holds a Law Degree, specialization in European and Transnational Law, from the University of Trento and a Master of Arts in Human Rights Law and Conflict Management from the Sant'Anna School of Advanced Studies.

Pavlina Pavlova is Public Policy Advisor at the CyberPeace Institute in Geneva, Switzerland, where she works on advancing international law and norms under the framework of responsible state behaviour in cyberspace and represents the Institute at the UN Open-Ended Working Group on ICTs and the UN Ad Hoc Committee on Cybercrime. Prior to this role, Pavlina was an official at the Organization for Security and Co-operation in Europe (OSCE), being appointed Liaison Officer of the OSCE Chairmanship. She also coordinated OSCE capacity building programmes aimed at strengthening the human dimension of security. Pavlina has been publishing and speaking on the nexus between technology, human rights, and security. Her research centres around cyber threats and cyberattacks impacting vulnerable and targeted groups and the interlink between online and offline security. She authored research papers presented at the Yale MacMillan Center for International and Area Studies, Carr Center for Human Rights Policy of the Harvard Kennedy School, and the Stanford Internet Observatory, among other fora.

Nanjira Sambuli is a researcher, policy analyst and strategist studying the unfolding, gendered impacts of digitalization/ICT adoption on governance, diplomacy, media, entrepreneurship, and culture, especially in Africa. Nanjira is a Fellow in the Technology and International Affairs Program at The Carnegie Endowment for International Peace, and a Ford Global Fellow. She is also a board member at The New Humanitarian, Development Gateway, and Digital Impact Alliance. She is a member of the Gender Advisory Board at the UN Commission on Science and Technology for Development (CSTD), and a Diplomacy Moderator at the Geneva Science and Diplomacy Anticipator (GESDA). Nanjira served as co-chair of Transform Health, as a Commissioner on the Lancet & Financial Times Governing Health Futures 2030 Commission, as a panel member on the United Nations Secretary General's High-Level Panel on Digital Cooperation, and as a deputy on the United Nations Secretary General's High-Level Panel for Women's Economic Empowerment.

Monica Nila Sari is a Ph.D. candidate at the Graduate School of Media and Governance, Keio University, Japan. Her research interests include cybersecurity cooperation in ASEAN and the effectiveness of the ASEAN regional approach in dealing with security issues. She has an international law background and obtained her Advanced Master's in Air and Space Law from Leiden University, the Netherlands. She has worked at the Ministry of Foreign Affairs of Indonesia and was assigned to the Embassy of Indonesia in the Hague, the Netherlands. During her time at the Ministry of Foreign Affairs, she received an award from the Indonesian Government as a member of Indonesia Delegate to be granted an extended Continental Shelf in the United Nations. Furthermore, she joined the ASEAN Secretariat in Jakarta, Indonesia, where she was assigned to handle ASEAN's political and security cooperation with external partners.

Sofiya Sayankina has received her PhD from Hankuk University of Foreign Studies, South Korea. She is currently doing research at the Center for International Cooperation and Strategy in Seoul. Her main research interests are cybersecurity governance and emerging technology policy.

Marta Stroppa is a PhD Candidate in Human Rights and Global Politics at the Sant'Anna School of Advanced Studies. She is currently a Visiting Researcher at the NATO Cooperative Cyber Defence Centre of Excellence and at the School of Law of the University of Westminster. She is also a Research Fellow of the Information Society Law Center of the University of Milan. Her research focuses on the legal implications of new technologies in the use of force and conduct of hostilities. She previously worked in the Legal Affairs Office of the Permanent Mission of Italy to the United Nations in New York and in the Global Maritime Crime Programme of United Nations Office on Drugs and Crime. She holds a Bachelor's and Master's Degree in International Relations from the University of Milan and a Master of Laws in International and Human Rights Law from Tilburg University.

Arun Sukumar is a post-doctoral research fellow at The Hague Program on International Cyber Security, Leiden University. He is a co-editor of *Multistakeholder Diplomacy: Building an International Cybersecurity Regime* (Edward Elgar Publishing, forthcoming) and the author of *Midnight's Machines: A Political History of Technology in India* (Penguin RandomHouse India, 2019).

This volume draws from papers presented at the conference **Closing the Gap | Responsibility in Cyberspace: Narratives and Practice**, organized in June 2022 at the Egmont Palace in Brussels, Belgium, by Leiden University, as part of the EU Cyber Direct project. The conference brought together researchers and practitioners from governments, academic institutions, technology companies, and civil society organisations around the world. Drawing on academic and policy frameworks, papers presented at the conference explored whether and how global, regional, and national narratives on responsible state behaviour in cyberspace have translated into practice. Several papers offer prescriptive solutions to bridge narrative and practice, where gaps exist.

The conference also hosted roundtables inviting experts from around the world, especially the Global South, to reflect on and improve various aspects of EU cyber diplomacy. These roundtables also saw stakeholders share ideas and experience on how to engage policymakers more effectively, and make cyber diplomacy more inclusive in the process. **Closing the Gap 2022** was conceived as a platform to facilitate exchange of perspectives from different stakeholders involved in UN cybersecurity negotiations, serving both as a neutral venue where state and non-state stakeholders could interact freely, and as a feeder process to those same UN discussions. The conference will return next year.

EU CYBER DIRECT

IMPLEMENTING ORGANISATIONS

eliss European Union Institute for Security Studies

Universiteit Leiden Institute of Security and Global Affairs

CARNEGIE EUROPE

FUNDED BY THE EUROPEAN UNION

Publications Office of the European Union