European Union
Institute for Security Studies

HACKING MINDS AND MACHINES

Foreign interference in the digital era

Edited by
Nad'a Kovalčíková

With contributions from
Rumena Filipova, Bart Hogeveen,
Ivana Karásková, Patryk Pawlak,
Andrea Salvi

The EUISS is an agency
of the European Union

CHAILLOT PAPER / 184
August 2024

European Union Institute for Security Studies (EUISS)

100, avenue de Suffren
75015 Paris

http://www.iss.europa.eu
Director: Steven Everts

The views expressed in this publication are solely those of the author(s) and do not necessarily reflect the views of the European Union.

Print		PDF	
ISBN	978-92-9462-351-5	ISBN	978-92-9462-350-8
CATALOGUE NUMBER	QN-AA-24-004-EN-C	CATALOGUE NUMBER	QN-AA-24-004-EN-N
ISSN	1017-7566	ISSN	1683-4917
DOI	10.2815/669361	DOI	10.2815/096722

Published by the EU Institute for Security Studies and printed in Belgium by Bietlot.
Luxembourg: Publications Office of the European Union, 2024.
Cover image credit: Robs/unsplash

European Union
Institute for Security Studies

HACKING MINDS AND MACHINES

Foreign interference in the digital era

Edited by
Nad’a Kovalčíková

With contributions from
Rumena Filipova, Bart Hogeveen,
Ivana Karásková, Patryk Pawlak,
Andrea Salvi

The EUISS is an agency
of the European Union

CHAILLOT PAPER / 184
August 2024

Acknowledgements

The editor would like to thank Smaranda Olariu, EUISS trainee, for her dedicated research assistance.

The editor

Nad’a Kovalčíková is a Senior Analyst in charge of the transnational security portfolio at the EUISS.

CONTENTS

INTRODUCTION

CRITICAL DOMAINS OF FOREIGN INTERFERENCE

by
NAD'A KOVALČÍKOVÁ

In today's security landscape, foreign interference[1] has become a pervasive threat. Hostile actors are infiltrating everything from social media to government websites, targeting trade secrets, and posing an increasing risk to critical infrastructure systems. This requires heightened vigilance and concerted efforts to detect, expose and counter these malign activities. The impact of intentional and harmful interference operations is amplified when wielded simultaneously across diverse societal sectors. Therefore, it is crucial to devise cross-sectoral frameworks, tools and responses and examine specific incidents of foreign interference, in order to address critical threat vectors.

In April this year, with the US presidential elections looming on the horizon and Russia's war against Ukraine having entered its third year, yet another episode of foreign interference was detected. The viral clip, containing false claims about a Kyiv troll farm attempting to interfere in the US elections[2], aimed to discredit the Ukrainian authorities. This incident was part of a larger campaign conducted by a group of disinformation experts connected to Russia's Internet Research Agency. These hostile actors are deploying increasingly sophisticated technology to disrupt Western democracies and their allies, and fabricating and spreading manipulated audio-video content online. A few months earlier, on 7 December 2023, the United Kingdom accused Russia's Security Service, the FSB, of orchestrating a 'sustained cyber-hacking campaign'[3] targeting politicians and other public representatives for a number of years, including during critical election periods. Immediately afterwards, the EU High Representative for Foreign Affairs and Security Policy (HR/VP), Josep Borrell, stated that '[a]ctivities that seek to threaten our integrity and security, democratic values and principles and the core functioning of democracies are unacceptable.' He also underlined

[1] For a definition of foreign interference, see Jones, K., 'Legal loopholes and the risk of foreign interference', European Parliament, In-Depth Analysis requested by the ING2 special committee, January 2023 (https://www.europarl.europa.eu/RegData/etudes/IDAN/2023/702575/EXPO_IDA(2023)702575_EN.pdf): '[Foreign interference] includes covert or coercive interference by a foreign power in the political or governmental system from within [...], influence of overseas regimes; influence on the political system from without, such as abuse of the lobbying system, corruption, espionage, cyber-attacks, and manipulative influence on public engagement or views, for example through online disinformation and manipulative campaigns.' The term also encompasses interference in social, economic, digital and international security domains, among others.

[2] 'Russian disinformation videos smear Biden ahead of US elections', *New York Times*, 16 May 2024 (https://www.nytimes.com/2024/05/15/us/politics/russia-disinformation-election.html).

[3] 'Russia hacking: "FSB in years-long cyberattacks against the UK, says government"', BBC News, 7 December 2023 (https://www.bbc.com/news/uk-politics-67647548).

the need to protect the European parliamentary elections 'from malign foreign actors who want Europe to fail' [4]. Only a week after the United Kingdom made similar declarations in relation to foreign (digital) interference, on 15 December 2023, the *Financial Times* published an alarming story about Chinese spies recruiting a European politician as part of an operation aimed at dividing the West [5]. These incidents, however, represent only a small part of the ongoing malicious activities being conducted by a broad range of actors. This includes, for example, the large-scale and sustained Russia-based *Doppel-Gänger* influence coperation targeting Western countries and their media outlets since 2022.

The surge in foreign interference against democracies demonstrates the rising importance and urgency of countering these hostile activities. This requires a heightened focus given their potential to critically impact national, European and transnational security. Moreover, cyberattacks and information manipulation are becoming increasingly intertwined, with growing evidence that foreign actors are generating or contributing to evolving security challenges. Their goal is to undermine their adversaries' core values and institutions and either exploit existing vulnerabilities or create new ones.

This *Chaillot Paper* examines foreign interference across a broad range of sectors [6]. It analyses how foreign information manipulation and interference (FIMI), and cyber threats are interconnected within a broader toolkit, highlighting both their points of convergence and divergence. Specific chapters dissect interference across a particular domain, exploring emerging policy approaches. Each case study follows a clear structure, identifying an *incident,* its *effects,* and the *response* measures taken, and outlining possible implications or policy recommendations to consider.

The volume explores in particular the cyber-FIMI nexus across **five key areas: the political, social, economic, digital**, and **international security** domains.

1. **The political domain**: Interference operations in this domain are among the most insidious as they are often known only to few people and are rarely exposed by those directly involved. In chapter 1 Ivana Karásková analyses various influence tactics, objectives and connections between China's information manipulation and the coercion of political representatives. A thorough understanding of these efforts is critical to prevent China's interference tactics from disrupting the political dynamics of targeted states, but also from undermining the EU's democratic processes and international standing. The author highlights the interconnectedness of the online and offline domains and draws attention to the vulnerabilities of European polities to foreign interference. She further emphasises the need for the EU and its Member States to consider developing a 'counter-coercion toolbox'.

2. **The social domain:** In chapter 2 Rumena Filipova, focusing on the case study of Bulgaria, examines how the interplay of internal and external factors can exacerbate foreign interference in a domestic societal context. She shows how, since Russia's invasion of Ukraine, a confluence of factors – increased information manipulation, the growing influence of pro-Russian media, political infiltration and cyberattacks – has made citizens more receptive to anti-Western and pro-Russian narratives aiming to fracture Bulgarian society and undermine the country's commitment to EU and NATO. The author emphasises the importance of exposing the subversive

[4] Borrell, J., 'Fighting foreign interference to protect our democracy', EEAS, 3 June 2024 (https://www.eeas.europa.eu/eeas/fighting-foreign-interference-protect-our-democracy_en).

[5] 'Chinese spies recruited European politician in operation to divide West', *Financial Times*, 15 December 2023 (https://www.ft.com/content/601df41f-8393-46ad-9f74-fe64f8ea1a3f).

[6] This *Chaillot Paper* does not claim to provide an exhaustive overview of foreign interference across every domain in which it occurs. The case studies presented explore sectors with well-documented incidents and reflect the individual expertise of the contributing authors.

activities of domestic pro-Russian actors operating within opaque networks. She stresses the need to bolster defences against foreign authoritarian influence and formulate a consistent strategy against Russian interference across various domains, including in the economic, media and political spheres.

3. **The economic domain**: In chapter 3 Bart Hogeveen analyses foreign state-sponsored cyber-enabled espionage and in particular, intellectual property theft. While individual incidents can seriously harm companies' commercial interests, the issue becomes more consequential when certain industries become targets of sophisticated and sustained digital operations that attempt to extract confidential business information. The author investigates the links between this form of cyber operations and foreign interference in the context of geostrategic and technological competition and considers the palette of responses available to states to strengthen their defences.

4. **The digital domain**: In chapter 4 Andrea Salvi delves into the increasingly sophisticated world of deepfakes, exploring their evolving features, various categories and impact. These are frequently deployed as part of targeted information manipulation efforts in electoral contexts, and beyond, feeding into broader disruptive interference operations. As artificial intelligence (AI) advances, the line between reality and fabrication is becoming increasingly blurred. After examining specific incidents, their impact and consequences, the author proposes solutions based on a collaborative approach. These solutions combine regulatory measures with initiatives to build societal resilience and public trust, empowering citizens to critically evaluate information. Such a collaborative effort is crucial to combat the misuse of deepfake technology by malicious actors who seek to shape and distort public perceptions and deepen divisions within society.

5. **The international security domain**: In chapter 5, Patryk Pawlak explores the link between FIMI and cyberattacks targeting critical infrastructure. Focusing on a set of concrete examples, he argues for a combined approach to FIMI and cybersecurity, as the information and cyber realms are inextricably interlinked. He also calls for a more rigorous approach to designating 'critical information environment infrastructure' and inclusion of the information environment as a key component in the discussion about critical infrastructure protection. The chapter examines the different categories of cyber incidents and the factors influencing attack methods. A key focus is the role of political decision-making, in particular when it comes to designating cyber incidents as foreign interference. This is especially important when state-backed perpetrators target a country's critical infrastructure with the intent to influence its foreign policy. To address risks and threats along the cyber-information continuum, critical infrastructure protection strategies should clearly define and protect their critical information environment. Since political and technical criteria for such designations can differ, the ability to distinguish between attack types is crucial for effectively assessing their impact on critical infrastructure.

This *Chaillot Paper* presents a comprehensive analysis of foreign interference tactics and their effects through the five distinct case studies outlined above. Exploring a diverse range of effects and responses, in the concluding chapter it identifies recurring patterns and exposes the interconnectedness of these interference toolboxes. The analysis not only highlights key differences as well as similarities in tactics and strategies, but also pinpoints areas where EU policies can be strengthened through integration. By offering targeted recommendations in each chapter and broader more comprehensive recommendations in the conclusion, this volume aims to equip policymakers with an effective, tailored and actionable strategy to counter these increasingly intertwined threats.

PART I: HACKING MINDS

Image: Harrison Leece/Unsplash

CHAPTER 1

CHINA'S INFORMATION MANIPULATION AS A TOOL OF POLITICAL COERCION

The case of the Czech Republic

by
IVANA KARÁSKOVÁ

INTRODUCTION

China's interference in the politics of EU Member States encompasses both online and offline attempts to manipulate their policy choices and preferences. Yet the online and offline domains are quite often treated as separate and the interconnection between them is neglected. This chapter sets out to bridge the gap and shed light on links between China's information manipulation and coercion of individual European politicians, aimed at inducing them to change their political behaviour.

Beijing's attempts to manipulate information take many forms and are driven by several objectives: acquiring 'discourse power'[1]; shaping international public opinion; and creating a favourable environment for China's rise by influencing international policies that benefit its own growth as a global power. China's strategies for information manipulation are extensive: it has invested in European media outlets, hired PR companies to facilitate inclusion of Chinese views and pro-China content in local media, and coordinated inclusion of op-eds authored by Chinese heads of mission in at least six different languages in newspapers in Central and Easten European countries. Additionally, China offers financial incentives to local media for favourable coverage or paid content supplements promoting China's agenda[2]. However some tactics involve a less transparent approach. China has been known to distribute China-prepared content through local media outlets without disclosing its origin, essentially disguising the source of the information from local

[1] Stanford University, 'Lexicon: "Discourse Power" or the "Right to Speak" (话语权, Huàyǔ Quán)', 17 March 2022 (https://digichina.stanford.edu/work/lexicon-discourse-power-or-the-right-to-speak-huayu-quan/).

[2] Karásková, I., 'How China influences media in Central and Eastern Europe', *The Diplomat*, 19 November 2019 (https://thediplomat.com/2019/11/how-china-influences-media-in-central-and-eastern-europe/).

audiences [3]. Furthermore, it has established cooperation with European news agencies, potentially influencing the flow of information on a broader scale [4].

Social media also plays a significant role in China's information manipulation efforts. State-affiliated accounts on social media platforms (such as those belonging to Chinese embassies, Xinhua, China Daily, China Radio International, etc.) are used to spread and amplify political narratives favourable to China. These accounts portray China as a responsible stakeholder in the international system, highlighting projects like the Belt and Road Initiative or its format for cooperation with Central and Eastern European countries. They also emphasise China's commitment to the principles of noninterference and respect for sovereignty.

Instances of Chinese intimidation or coercion against politicians should not be downplayed or ignored.

A significant shift in Chinese information manipulation tactics can be observed since 2019, when China perceived a greater need to step up engagement with European audiences to rewrite the narratives on the protests in Hong Kong and subsequently the origin of Covid-19. In an effort to enhance the credibility of Chinese narratives, China opted for outsourcing the production of China-related news to local partners. With the outbreak of the coronavirus pandemic Chinese messaging started to appear less in mainstream media and more in fringe media outlets in the form of anonymous articles, complicating the attribution of such outputs. Meanwhile, Chinese diplomats employed more offensive messaging, including posts and reposts of disinformation narratives on the 'real' origin of Covid-19.

After Russia's invasion of Ukraine, Chinese-state affiliated actors contrasted China's position with that of the United States, repeated Russian propaganda on the causes of the war and attempted to drive a wedge between the transatlantic partners by portraying the EU as the victim in the relationship (highlighting, for example, rising energy prices and inflation in Europe).

Chinese attempts to manipulate discourses and information in Europe evidently warrant heightened attention, especially as an ever-wider array of activities has shifted to the online sphere since the pandemic. However, this chapter focuses on the specific intersection between information manipulation and coercion of politicians which may take both online and offline forms. Instances of Chinese intimidation or coercion against individual politicians should not be downplayed or ignored. Nor should they be omitted in broader discussions on China's interference since they represent a direct attempt by China to influence democratic processes in the European Union.

The events which trigger the use of coercive measures typically revolve around China's self-defined 'core interests', such as preserving China's political regime, national security, sovereignty and territorial integrity and the stable development of China's economy and society [5]. In practical terms, harsh reactions

[3] Karásková, I.,'Analysing China Radio International's tactics: A case study of narratives disseminated in the Czech Republic,' Central European Digital Media Observatory (CEDMO), Prague, 2023 (https://cedmohub.eu/wp-content/uploads/2023/06/EN_Espresso.pdf?_gl=1*83qrpa*_up*MQ..*_ga*NTE1NjIwMTM0LjE3MDkxMTY5OTg.*_ga_44P6SY9R25*MTcwOTExNjk5Ny4xLjAuMTcwOTExNjk5Ny4wLjAuMA..).

[4] Gragnani, L., 'What's left in the West: China's short and long-term gains in Italy', China Observers in Central and Eastern Europe (CHOICE), 6 July 2021 (https://chinaobservers.eu/whats-left-in-the-west-chinas-short-and-long-term-gains-in-italy); Alliance for Securing Democracy, 'Agreements between Polish and Chinese media groups expose Polish audiences to propaganda', n.d. (https://securingdemocracy.gmfus.org/incident/agreements-between-polish-and-chinese-media-groups-expose-polish-audiences-to-propaganda/).

[5] Swaine, M.D., 'China's assertive behavior – Part One: On "core interests"', *China Leadership Monitor*, No 34, 22 February 2011 (https://carnegieendowment.org/files/CLM34MS_FINAL.pdf).

from China have followed official activities (at governmental or sub-governmental levels, or both) related to Tibet or Taiwan (e.g. meeting with the Dalai Lama, leading an official delegation to Taiwan, opening Taiwanese representative offices).

Understanding the full extent of coercion applied against political representatives by Chinese state entities is complicated by two factors. First, while instances of intimidation are by no means rare, the knowledge of these incidents is often confined to small circles of those who were directly involved – such as political representatives' staff, advisors and close contacts[6]. Occasionally, the information is shared by political representatives with the media and reported anecdotally[7]. The second challenge arises from the involvement of indirect actors. While most documented cases have involved direct confrontation between an individual politician and Chinese state representatives, a few instances have involved domestic figures acting as conduits or intermediaries for Chinese interference. These cases deserve special attention as they may complicate and obscure the evaluation of coercive tactics utilised by China. They suggest that it is not always or exclusively China's interests and actions which dictate the outcome of China's interference in European polities.

The Czech security community has been a vocal critic of China's information manipulation and cyber threats.

THE INCIDENT

The Czech Republic offers a well-documented and illustrative example of offline coercion backed by an online disinformation campaign targeting a political representative of an EU Member State. Here debates on China have long oscillated between the proponents of economic benefits allegedly deriving from closer relations with Beijing, and those who advocate a human rights-oriented standpoint, which has become deeply-rooted in the contemporary Czech foreign policy tradition.

Czech political parties quickly discovered the attractiveness of the China issue for the media and general public, and China featured prominently as a topic in political debates, including during the local parliamentary and presidential campaigns. Specifically, Tibet since the 1990s and Taiwan since 2016[8] began to be leveraged as counterweights to the pro-China policy promoted by certain politicians, including the previous president Miloš Zeman. Although Taiwan and Tibet are distant from the Czech Republic not only in geographic but also in strategic and security terms, they resonate with Czech citizens' own historical experience. Many Czechs see parallels between their own history of being surrounded and threatened by larger authoritarian regimes (Nazi Germany in 1938 and the Soviet Union in 1968) and the plight of Tibet and Taiwan.

Despite a part of the Czech political spectrum having an affinity with China, the Czech

[6] Interviews with a member of the Czech parliament and a senator, 23 July 2020, Prague. Interview with a German legislator, 13 October 2022, Berlin.

[7] See for example: Chen, Y., and Shih, H., 'Ukrainian MP under pressure from China over pro-Taiwan caucus', Focus Taiwan, 9 October 2022 (https://focustaiwan.tw/politics/202209100021).

[8] See a comparison of the frequency of these keywords in Czech media (2010-2017) in Karásková, I., Matura, T., Turcsányi, R. and Šimalčík, M., 'Central Europe for Sale: The politics of China's influence', Association for International Affairs (AMO), Prague, 2018, pp. 8 (https://www.amo.cz/wp-content/uploads/2018/04/AMO_central-europe-for-sale-the-politics-of-chinese-influence.pdf). For the symbolism of Tibet and Taiwan in Czech political debates, see an analysis of Czech MPs' standpoints in Karásková, I., Bajerová, A. and Matura, T., 'Images of China in the Czech and Hungarian Parliaments', Association for International Affairs (AMO), Prague, 2019 (https://www.amo.cz/wp-content/uploads/2019/03/AMO_Images-of-China-in-the-Czech-and-Hungarian-Parliaments.pdf).

security community has been a vocal critic of China's information manipulation and cyber threats. In late 2018 the Czech Republic's state agency, the National Authority on Cyber and Information Security (NÚKIB), issued a public warning against Chinese telecom companies ZTE and Huawei, making the country an early sceptical voice in the European debate on their possible participation in 5G networks[9].

In May 2019, the so-called Prague Proposals, an outcome of an international conference on the security of 5G networks, established guidelines for assessing risks posed by foreign vendors. A year later, the Czech Republic signed a joint declaration on 5G security with the United States.

It is in this context that the events unfolding around Jaroslav Kubera, the then President of the Czech Parliament's Senate, can be framed. In 2019 Kubera planned a visit to Taiwan as the head of a delegation. The visit was scheduled to take place in early 2020. This would have represented the highest-profile visit of a Czech politician to the self-ruling island in decades. The Chinese Embassy in Prague reacted to the announced visit by issuing a letter [10], threatening repercussions against Czech companies conducting business in China if Kubera proceeded with his plans to visit Taiwan. Interestingly, in the public domain the objections against Kubera's visit were not communicated by China's representatives in the Czech Republic or directly by Beijing, but by the Czech President Miloš Zeman, who warned that the visit would damage Czech economic interests [11]. As Kubera died suddenly in January 2020, the planned visit did not go ahead.

The letter from the Chinese Embassy was found among Kubera's belongings shortly after his death and made public. The new Senate President Miloš Vystrčil expressed his interest in going ahead with the plan to lead an official delegation to Taiwan in Kubera's stead. Vystrčil announced his intention in open defiance of the then Czech President Miloš Zeman, Prime Minister Andrej Babiš and Minister of Foreign Affairs Tomáš Petříček and despite significant pressure applied by Czech-China business associations and Czech companies fearing retribution[12].

The proposed visit sparked a diplomatic row between the Czech Republic and China. Chinese Minister of Foreign Affairs Wang Yi declared that Vystrčil would pay 'a heavy price'[13] and the Chinese state media outlet Global Times called Vystrčil a 'political hooligan'[14]. This, in turn, prompted critical reactions from abroad as an open letter[15] backing Vystrčil was signed by almost 70 members of the European Parliament and other national parliaments, including those of Australia, the United States, and Canada. Another letter in support was published jointly by the chair and co-chair

[9] NUKIB, 'Varování', 17 December 2018 (https://nukib.cz/download/uredni_deska/Varovani_NUKIB_2018-122-17.pdf).

[10] Valášek, L. and Truchlá, H., 'Za návštěvu Tchaj-wanu budete platit. Kubera si z Hradu přinesl výhrůžky od Číny' [You will pay for visiting Taiwan. Kubera brought threats from China from the Castle], Aktuálně, 19 February 2020 (https://zpravy.aktualne.cz/domaci/za-navstevu-tchaj-wanu-budete-platit-kubera-si-z-hradu-prine/r~3602b9ba51a711eaa5e40cc47ab5f122/).

[11] '"Zeman: Jestli pojede Kubera na Tchaj-wan, už není můj přítel"'[Zeman: If Kubera goes to Taiwan, he is no longer my friend], Novinky, 28 November 2019 (https://www.novinky.cz/domaci/clanek/zeman-jestli-pojede-kubera-na-tchaj-wan-uz-neni-muj-pritel-40305312).

[12] Interviews with staff at Senate President's office, 23 July 2020, Prague.

[13] 'Wang Yi: Anyone who challenges one-China principle will pay "heavy price"', *Global Times*, 31 August 2020 (https://www.globaltimes.cn/content/1199387.shtml).

[14] 'Vystrcil's Taiwan visit an opportunistic stunt: Global Times editorial', *Global Times*, 30 August 2020 (https://www.globaltimes.cn/content/1199345.shtml).

[15] 'Political leaders express solidarity with Czech Senate Presdent Miloš Vystrčil in connection to his official visit to Taiwan', website of MEP Miriam Lexmann, 25 August 2020 (https://lexmann.eu/political-leaders-express-solidarity-with-czech-senate-president-milos-vystrcil-in-connection-to-his-official-visit-to-taiwan/).

China’s sharp teeth

The letter concerning Kubera’s planned visit to Taiwan from the Chinese Embassy addressed to the Office of the Czech President

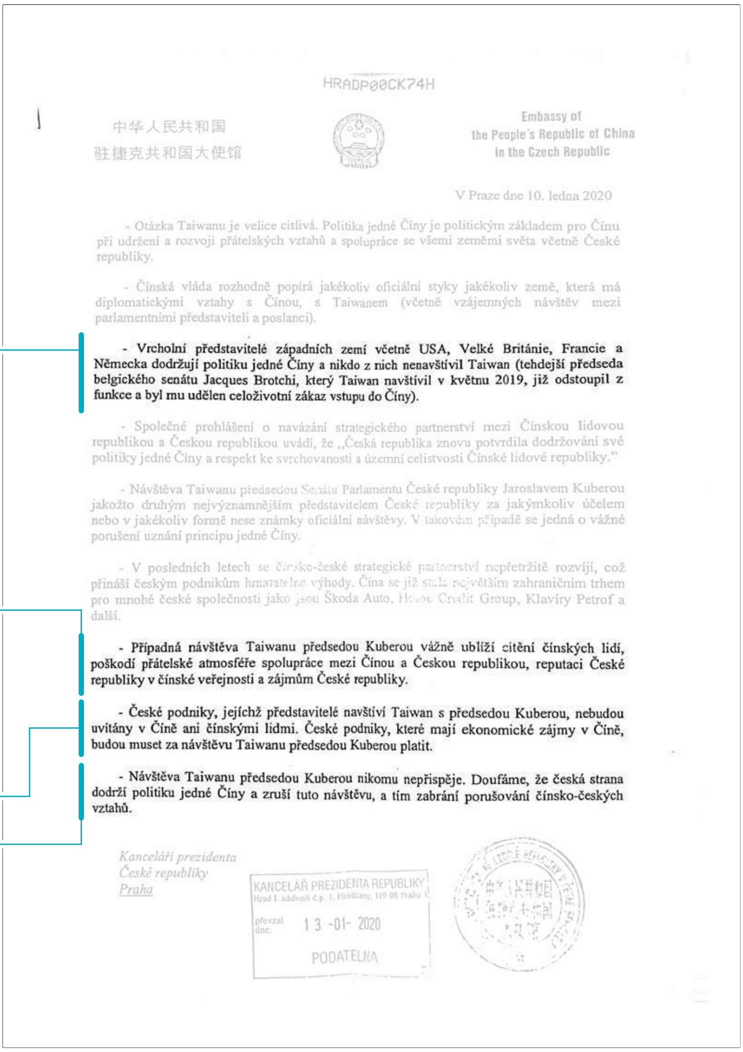
Translation of critical passages from original letter highlighting Beijing’s view and potential implications for Chinese-Czech relations

Top representatives of Western countries, including the USA, the United Kingdom, France and Germany, abide by the One-China Policy, and none of them has visited Taiwan (Jacques Brotchi, the then Chairman of the Belgian Senate, who visited Taiwan in May 2019, has already resigned from his office and received a lifetime ban from entering China).

A potential visit to Taiwan by Chairman Kubera would seriously hurt the feelings of the Chinese people, damage the friendly atmosphere of cooperation between China and the Czech Republic, the Czech Republic’s reputation among the Chinese public and the interests of the Czech Republic.

Czech enterprises whose representatives visit Taiwan with Chairman Kubera will not be welcome in China or by the Chinese people. Czech enterprises with economic interests in China will have to pay for Chairman Kubera’s visit to Taiwan.

Chairman Kubera’s visit to Taiwan will not benefit anyone. We hope that the Czech side will observe the One-China Policy and cancel this visit, thus avoiding damaging Chinese-Czech relations.



HRADP00CK74H

中华人民共和国
驻捷克共和国大使馆

Embassy of
the People´s Republic of China
in the Czech Republic

V Praze dne 10. ledna 2020

- Otázka Taiwanu je velice citlivá. Politika jedné Číny je politickým základem pro Čínu při udržení a rozvoji přátelských vztahů a spolupráce se všemi zeměmi světa včetně České republiky.

- Čínská vláda rozhodně popírá jakékoliv oficiální styky jakékoliv země, která má diplomatickými vztahy s Čínou, s Taiwanem (včetně vzájemných návštěv mezi parlamentními představiteli a poslanci).

- Vrcholní představitelé západních zemí včetně USA, Velké Británie, Francie a Německa dodržují politiku jedné Číny a nikdo z nich nenavštívil Taiwan (tehdejší předseda belgického senátu Jacques Brotchi, který Taiwan navštívil v květnu 2019, již odstoupil z funkce a byl mu udělen celoživotní zákaz vstupu do Číny).

- Společné prohlášení o navázání strategického partnerství mezi Čínskou lidovou republikou a Českou republikou uvádí, že „Česká republika znovu potvrdila dodržování své politiky jedné Číny a respekt ke svrchovanosti a územní celistvosti Čínské lidové republiky.”

- Návštěva Taiwanu předsedou Senátu Parlamentu České republiky Jaroslavem Kuberou jakožto druhým nejvýznamnějším představitelem České republiky za jakýmkoliv účelem nebo v jakékoliv formě nese známky oficiální návštěvy. V takovém případě se jedná o vážné porušení uznání principu jedné Číny.

- V posledních letech se čínsko-české strategické partnerství nepřetržitě rozvíjí, což přináší českým podnikům hmatatelné výhody. Čína se již stala největším zahraničním trhem pro mnohé české společnosti jako jsou Škoda Auto, Home Credit Group, Klavíry Petrof a další.

- Případná návštěva Taiwanu předsedou Kuberou vážně ublíží cítění čínských lidí, poškodí přátelské atmosféře spolupráce mezi Čínou a Českou republikou, reputaci České republiky v čínské veřejnosti a zájmům České republiky.

- České podniky, jejíchž představitelé navštíví Taiwan s předsedou Kuberou, nebudou uvítány v Číně ani čínskými lidmi. České podniky, které mají ekonomické zájmy v Číně, budou muset za návštěvu Taiwanu předsedou Kuberou platit.

- Návštěva Taiwanu předsedou Kuberou nikomu nepřispěje. Doufáme, že česká strana dodrží politiku jedné Číny a zruší tuto návštěvu, a tím zabrání porušování čínsko-českých vztahů.

Kanceláři prezidenta
České republiky
Praha

Source: Aktuálně.cz, February 2020

of the European Parliament Delegation to the People’s Republic of China (PRC) [16] and addressed to the Chinese ambassador in Brussels, effectively turning a bilateral issue into a matter of EU-China relations. Even Minister of Foreign Affairs Petříček who was originally sceptical about the visit summoned the Chinese ambassador to explain, while Prime

[16] European Parliament Delegation for Relations with People’s Republic of China, ‘Chair’s and Vice-Chair’s message of 20 August 2020 to Ambassador Zhang Ming’, 20 August 2020 (https://www.europarl.europa.eu/delegations/cs/d-cn/documents/communiques).

Minister Babiš called Wang's remarks incongruous[17].

China's coercive efforts did not end with issuing threats. Attempts to discredit Vystrčil have continued ever since. In November 2020, two months after his visit, an e-mail claiming Vystrčil received 4 million USD in exchange for his official visit to Taiwan started circulating in the Czech Republic[18]. The email containing the information was sent to various Czech newspapers from a Swiss-based consultancy, RefinSol Advisory Services, which claimed no personal interest in the issue yet inquired repeatedly whether the content of the mail had been published by the media.

The Kubera/ Vystrčil case shows how vulnerable European polities are to China's interference.

In the online domain, Vystrčil has been targeted by China Radio International and an interconnected ecosystem of Czech disinformation and otherwise dubious websites, known mostly for spreading pro-Russian and anti-Western narratives in the country[19].

THE EFFECTS

The Kubera/Vystrčil case demonstrates China's tactics of intimidation of and coercion against politicians that it perceives as acting in direct opposition to China's interest in preventing international recognition of and solidarity towards Taiwan. In general, it highlights the complexity and interconnectivity of information manipulation and online and offline coercion campaigns. Finally, it shows how vulnerable European polities are to China's interference.

In line with China's coercive diplomacy described in other cases, Beijing aimed at securing what it perceives as a 'core interest', i.e., preventing Taiwan from building more extensive international relations, by averting the official visit of a senior political representative from an EU Member State. It combined coercive measures targeting an individual politician (e.g. personal threats, disinformation and smear campaigns) with threats directed at the highest possible level, including warnings of economic retaliation against the state. This approach was based on the assumption that identifying the mode and scope of retaliation prior to the event taking place would lead China's challenger to back down.

Assuming the traditional secrecy of diplomatic and political practice, China was confident that its threatening letter would not be handed over to the media and publicised. When the letter was leaked to the public, it triggered a public outcry, fuelled in part by the death of Vystrčil's predecessor. This allowed Vystrčil to mobilise support from domestic and international political allies as well as from the general public. The media thus played a crucial role in empowering an individual political representative who had drawn Beijing's

[17] 'Vyjádření ministra Wanga jsou za hranou, uvedl Petříček. Ministerstvo předvolalo čínského velvyslance' [Minister Wang's statements are beyond the pale, Petříček said. The ministry summoned the Chinese ambassador], iRozhlas, 31 August 2020 (https://www.irozhlas.cz/zpravy-domov/milos-vystrcil-ministerstvo-zahranici-cesko-tchaj-wan-cina-velvyslanec_2008311533_vtk).

[18] Valášek, L. and Truchlá, H., '"Čtyři miliony dolarů pro Vystrčila". Číňané se pokoušejí očernit předsedu Senátu' ['"Four million dollars for Vystrčil", The Chinese are trying to smear the Senate President'], Aktuálně, 11 November 2020 (https://zpravy.aktualne.cz/domaci/ctyri-miliony-dolaru-pro-vystrcila-cinane-se-pokouseji-ocern/r~743c637e233511ebb408ac1f6b220ee8/).

[19] Most recently in connection with the President of Taiwan's Legislative Yuan's visit to the Czech Republic and US Congress Speaker Nancy Pelosi's visit to Taiwan, e.g., 'Předseda Strany DOMOV: Návštěva Pelosiové na Tchaj-wanu má ryze provokativní character' [Chairman of the DOMOV Party: Pelosi's visit to Taiwan has a purely provocative character], China Radio International (CRI), 30 August 2022 (https://czech.cri.cn/2022/08/30/ARTIOdymMwU800Uc55L4sZpu220830.shtml); 'Filip Andler: Vystrčile, přestaňte už škodit! [Filip Andler: Vystrčil, stop harming!]', Parlamentní listy, 27 July 2022 (https://www.parlamentnilisty.cz/arena/nazory-a-petice/Filip-Andler-Vystrcile-prestante-uz-skodit-709716).

ire and been subjected to bullying tactics. Moreover, media coverage helped to inoculate the public and other political representatives against China's future disinformation attempts by raising awareness of Chinese information manipulation and interference tactics. While the actual economic consequences of the visit were minimal, this incident heightened the Czech Republic's awareness of China's interference strategies.

THE RESPONSE

So far, the incidents of coercion against individual European politicians have not resulted in an adoption of a counter-measure toolbox. Perhaps due to China's ability to skilfully mobilise domestic political and economic interlocutors to act on its behalf, the issue may be seen more as a demonstration of a domestic political struggle rather than an act of foreign information manipulation and interference requiring special attention at the level of the EU and the Member States. The reluctance of the politicians concerned to publicise the details of China's coercion has contributed to a knowledge gap and lack of policy responses. The specific Kubera/Vystrčil case is rather unique in terms of the depth and complexity of the material which has been publicised.

The case demonstrably impacted the Czech political scene. Initially marginalised, Vystrčil's stance against Chinese pressure has since become the mainstream position with the arrival of the new government of Petr Fiala (since 2021) and its foreign policy priorities. Another delegation led by the Speaker of the Chamber of Deputies Markéta Pekarová-Adamová visited the island in March 2023[20].

Ironically, from China's point of view, the case helped Vystrčil to get re-elected to the Senate in October 2022 and to retain the position of Senate President. The disinformation campaign that China conducted against him was unsuccessful as influential Czech media refused to amplify the smears and swiftly exposed the campaign's origins. Fringe media outlets and China-affiliated state media which publicised critical content were not able to match the reach of the traditional mainstream media that rallied behind Vystrčil. Overall, in this specific case China's attempt to coerce its foreign opponents largely backfired as it led to a broader political backlash and, moreover, elevated the stature of the individual politician who challenged it.

CONCLUSION

Several broader conclusions can be drawn from Vystrčil's case which offer potential building blocks for a 'counter-coercion toolbox' for European politicians. First of all, Vystrčil adopted an approach of 'passive resistance' in which he reiterated his resolve to proceed with the visit – which eventually did indeed take place in August 2020. He gave two major reasons for his visit to Taiwan which appealed to the domestic audience: one value-based and the other economic. He claimed the visit showed the importance of values and value-based foreign policy which dictates cooperation with democracies around the world, such as that of Taiwan[21]. In this regard, Vystrčil directly linked his visit to the human rights-based approach embedded in Czech foreign policy since 1989. Moreover, he claimed the visit was morally justified as it fulfilled the wish of his deceased predecessor. The other goal he mentioned was to 'jumpstart' the Czech economy

[20] Šebok, F., 'Czech Speaker of Chamber of Deputies arrives in Taiwan on a "mission"', China Observers in Central and Eastern Europe (CHOICE), 28 March 2023 (https://chinaobservers.eu/czech-speaker-of-chamber-of-deputies-arrives-in-taiwan-on-a-mission/).

[21] 'Ukážeme, že jsme suverénní země, řekl Vystrčil k cestě na Tchaj-wan' [We will show that we are a sovereign country, said Vystrčil as he headed to Taiwan], iDNES, 27 August 2020 (https://www.idnes.cz/zpravy/domaci/vystrcil-predseda-senatu-tchaj-wan-pred-cestou.A200827_052720_domaci_kop).

in the wake of the Covid-19 epidemic, as well as to promote the Czech Republic as a 'gateway' [22] for Taiwanese investment in Europe. This argument can be regarded as a deliberate counterpoint to his opponents' criticism that the visit would lead to retaliation by China and jeopardise the Czech economy.

Second, the Czech politician managed to build a broad international coalition endorsing his visit, both in Europe and beyond (with prominent members of the Australian, Canadian and US parliaments expressing their support). The strategy shifted focus from the politician himself to the wider international backing he had secured, diverting the attention of China and its local allies.

Third, this specific incident exposed the limitations of China's leverage. The low level of economic interdependence between China and the Czech Republic meant that China simply did not have enough 'sticks' at its disposal to effectively punish the Czech Republic for its alleged 'misbehaviour' [23]. In fact, Vystrčil's visit has had minimal economic impact on trade between the two countries. Finally, he was able to secure strong backing from the mainstream Czech media which were openly critical of China and its behaviour [24]. This shows the importance of understanding that offline coercion attempts may be followed by online smear campaigns targeting an individual politician. Pre-emptive strategies to debunk disinformation should thus be prepared in advance and deployed in time to prevent Chinese online attempts to manipulate information.

To conclude, the case of Miloš Vystrčil's visit illustrates both China's practice of using coercion in international relations to advance its self-proclaimed 'core interests', and possible strategies for resisting such tactics.

[22] 'Tchaj-wan je ochoten vstoupit do Evropy přes Česko, řekl Vystrčil na závěr návštěvy' [Taiwan is willing to enter Europe via the Czech Republic, Vystrčil said at the end of the visit], ČT24, 4 September 2020 (https://ct24.ceskatelevize.cz/svet/3176244-predseda-senatu-vystrcil-konci-navstevu-tchaj-wanu-pred-odletem-se-zucastni-ekonomicke).

[23] Karásková, I., et al., 'China's sticks and carrots in Central Europe: The logic and power of Chinese influence', Association for International Affairs (AMO), Prague, 2020. (https://mapinfluence.eu/wp-content/uploads/2020/06/Chinas-Sticks-and-Carrots-in-Central-Europe__policy-paper_-1.pdf).

[24] 'Central Europe for Sale: The politics of China's Influence', op.cit.

CHAPTER 2

THE KREMLIN'S AGENDA IN BULGARIA

The role of pro-neutrality protests in disinformation campaigns

by
RUMENA FILIPOVA

INTRODUCTION

Bulgaria is subject to some of the most wide-ranging Russian influence campaigns in Central and Eastern Europe. These campaigns exploit distorted cultural proximity, historical narratives shaped by Russian propaganda, and the Kremlin's cultivation of opaque networks of patronage within Bulgaria's political, business and media spheres[1]. Domestic issues, such as vested political interests in media ownership, (self-)censorship, and low quality of content, further open the door to foreign authoritarian media influence. And as the Kremlin's war against Ukraine has unfolded, Russia has further scaled up its information manipulation efforts in Bulgaria.

At times, the Bulgarian authorities have demonstrated a robust response against Russian aggression. In January 2023, then-Prime Minister Kiril Petkov stated in an interview with the German daily *Die Welt* that 'about a third of the ammunition needed by the Ukrainian army in the early phase of the war came from Bulgaria'. Under his leadership, the country managed to provide essential aid to Kyiv, including 'weapons, ammunition and diesel fuel' and initiated the EU's first measures against Russia's full-scale invasion. In retaliation, the Kremlin was swift to respond, promptly engaging in tactics such as political infiltration, halting gas exports, and launching cyberattacks on postal services[2]. This combination of internal and external factors significantly increased the country's vulnerability to anti-Western and pro-Russian narratives.

Societal receptivity to Russia's viewpoints is shaped by Bulgaria's historical East-West ambivalence and widespread sentiment in favour of neutrality *vis-à-vis* Russia's war

[1] Filipova, R., *History Undone: Russia's historical disinformation, Bulgaria's memory politics and lessons for dealing with the past from Central and Eastern Europe*, Institute for Global Analytics, 2023 (https://globalanalytics-bg.org/2023/08/14/new-report-history-undone-russias-historical-disinformation-bulgarias-memory-politics-and-lessons-for-dealing-with-the-past-from-central-and-eastern-europe/).

[2] 'Bulgaria to the rescue: How the EU's poorest country secretly saved Ukraine', *Politico*, 18 January 2023 (https://www.politico.eu/article/bulgaria-volodymyr-zelenskyy-kiril-petkov-poorest-country-eu-ukraine/).

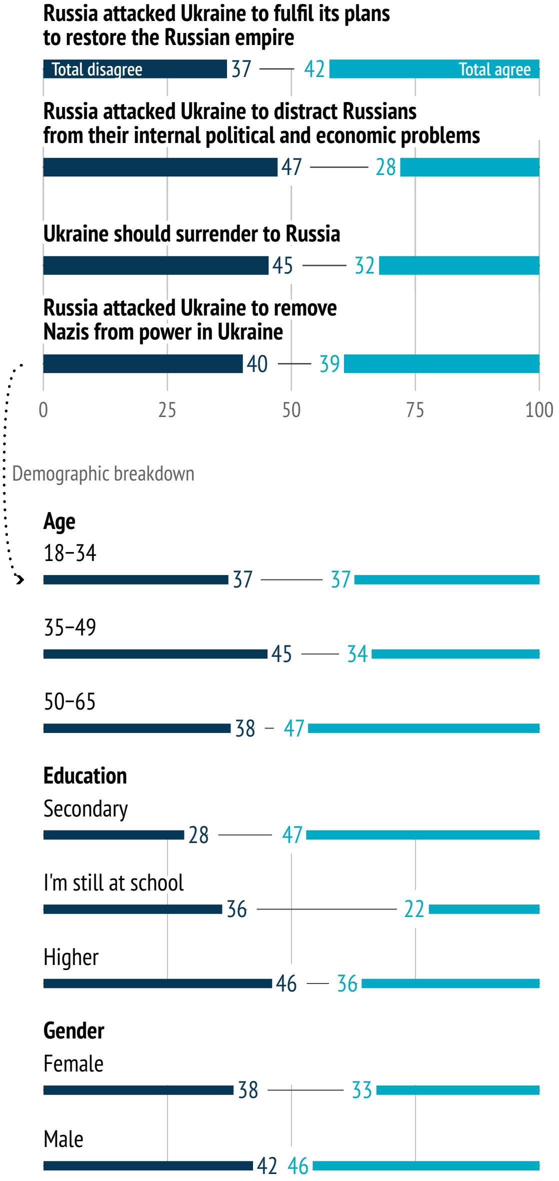
against Ukraine. Attempts to forge a neutral position are usually informed by a concealed or explicit pro-Russian bias, whereby neutrality is seen as a way to weaken Bulgaria's pro-Western orientation, rather than ensure security through military non-involvement in third-party conflicts. This pro-neutrality trend co-exists with a tug-of-war between pro-Russian and Russia-critical politicians and sections of society, further exacerbating the divide between those who support Moscow and Kyiv.

Various pro-Russian, anti-Ukrainian and anti-EU narratives have been circulating in the Bulgarian information space. These include claims that, for instance, 'Ukrainian grain is flooding the country', 'Ukrainian refugees are being hosted at luxury hotels' or 'the adoption of the euro will lead to poverty' [3]. These narratives were also amplified by the Vazrazhdane ('Revival') and Bulgarian Rise parties [4]. Furthermore, in November 2023, the Defence Minister Todor Tagarev accused Vazrazhdane and the Bulgarian Socialist Party (BSP) of spreading the pro-Kremlin narrative that 'the Bulgarian government could drag the country into a military conflict with Russia'. Tagarev voiced the belief that this might have discouraged young people from pursuing a career in the army [5].

In these circumstances characterised by strong Russia-leaning sentiment as well as political flux and instability, protests that were particularly prominent in 2022 and were spearheaded by pro-Russian actors, have served as a tool in the Kremlin's strategy of undermining Bulgaria's EU and NATO commitments, and fuelling societal polarisation. Therefore, in contrast to the dominant Western view of social movements as furthering the process of democratisation, Russian-backed social

The battle of narratives

Initial perceptions in Bulgaria of Russia's full-scale invasion of Ukraine, 2022



Data: Kantar Public & Science +, 'Bulgaria', July 2022

[3] Margova, R. and Dobreva, M., 'Disinformation landscape in Bulgaria', EU DisinfoLab, June 2023 (https://www.disinfo.eu/wp-content/uploads/2023/06/20230627_BulgariaDisinfoFS.pdf).

[4] 'The euro is poised to become the next target of disinformation in Bulgaria', *Veridica*, 30 November 2022 (https://www.veridica.ro/en/acf/the-euro-is-poised-to-become-the-next-target-of-disinformation-in-bulgaria).

[5] 'The pro-Russian disinformation has greatly affected the cadres in the army', Euractiv, 15 November 2023 (https://www.euractiv.com/section/politics/news/bulgarian-army-weakened-by-pro-russian-disinformation/).

agitation aims to undermine Europeanisation and promote authoritarian trends.

THE INCIDENT

Since the start of Russia's invasion of Ukraine, protests calling for Bulgaria to remain neutral have been instigated by domestic pro-Russian actors. Such organised public unrest exemplifies how Russia can activate its network of pro-Kremlin local proxies – encompassing political, economic and media groups and interests – to sow division within society. The predominantly informal means through which Moscow establishes its leverage over pro-Russian patronage networks in Bulgaria (and indeed the wider southeast European region) makes Russian influence even more pernicious due to its opacity and high degree of concealment[6].

The nationalist, right-wing political party Vazrazhdane is a particularly prominent pro-Russian actor in the Bulgarian public sphere, which has led protests favouring Russia. Vazrazhdane has continuously disseminated Kremlin disinformation and aligns with Moscow's policy lines in Bulgaria's domestic politics. For instance, the party refused to vote in favour of a parliamentary declaration condemning Russia's war on Ukraine and has come out against imposing sanctions on the Kremlin[7]. Notably, following the start of the Russian invasion, Vazrazhdane's leader, Kostadin Kostadinov, entered Ukraine but was expelled by the Ukrainian authorities, who banned him from entering the country for the next 10 years due to the fact that he heads a pro-Russian political party in Bulgaria and was therefore thought to represent a threat[8].

In line with Vazrazhdane's vociferous support for the Kremlin's political agenda, the party has organised protests and civil unrest calling for Bulgarian neutrality in the war. These protests condemn the dispatch of military assistance to Ukraine and advocate for closer Bulgarian-Russian political and economic cooperation. In an incident which took place on 3 March 2022 (when Bulgaria controversially celebrates[9] its national Independence Day marking liberation from the Ottoman Empire, following the Russian-Turkish War of 1877-78), Vazrazhdane attended a commemoration event, at which the party's members and supporters carried the Russian flag, chanted pro-Russian slogans and called for Bulgaria's withdrawal from NATO[10].

Shortly afterwards, at the beginning of April 2022, a pro-neutrality protest convened by Vazrazhdane was said to be motivated by the goal to get rid of 'national traitors' (i.e. those politicians and sections of society who are critical of Moscow). Moreover, it also advocated the 're-establishment of Bulgarian statehood' and an end to 'foreign occupation' (i.e. as epitomised by NATO's presence in Bulgaria) and foreign interests that are trying to drag

[6] For more on the informal character of Russian influence, see: Filipova, R. and Shopov, V., *Authoritarians on a media offensive in the midst of war: The informational influence of Russia, China, Turkey, Iran and the Gulf states in Southeast Europe*, Konrad-Adenauer-Stiftung Media Programme South East Europe, 2022 (https://globalanalytics-bg.org/2022/12/07/report-authoritarians-on-a-media-offensive-in-the-midst-of-war/).

[7] Gospodinova, V. and Dimitrov, D., 'БСП и "Възраждане" защитиха в парламента интересите на Русия' ['The BSP and Revival defended Russia's interests in Parliament'], *Capital*, 24 February 2022 (https://www.capital.bg/politika_i_ikonomika/bulgaria/2022/02/24/4316253_bsp_i_vuzrajdane_zashtitiha_v_parlamenta_interesite_na/).

[8] 'Костадин Костадинов обяви, че е изгонен от Украйна заради информация, че е проруски политик' ['Kostadin Kostadinov announced that he was expelled from Ukraine due to information that he is a pro-Russian politician'], Svobodna Evropa [Radio Free Europe/Radio Liberty], 7 March 2022 (https://www.svobodnaevropa.bg/a/31740689.html).

[9] The designation of 3 March as Bulgaria's national day is controversial. The choice stems from the Treaty of San Stefano, signed on this day in 1878 between Russia and the Ottoman Empire. The treaty is seen as a problematic foundation for marking Bulgaria's liberation as it was preliminary in character, sought to advance Russian interests, and excluded other Great Powers.

[10] 'Викове "предатели" и руски знамена посрещнаха Кирил Петков на Шипка' ['Kiril Petkov was met with "traitor" chants and Russian flags on Shipka'], Mediapool, 3 March 2022 (https://www.mediapool.bg/vikove-predateli-i-ruski-znamena-posreshtnaha-kiril-petkov-na-shipka-news332915.html).

the country into a war in which Sofia has ostensibly no stake [11]. A similar protest against Bulgaria's provision of military aid to Ukraine also took place in May 2022 [12], following Ukrainian Foreign Minister Dmytro Kuleba's visit to Sofia earlier, on 19 April. These protests spurred further initiatives by various pro-Russian groups calling for neutrality. For instance, the Public Council 'Bulgaria for Peace and Neutrality' launched a petition, arguing that pacifist goals are best served via non-interference in Russia's war in Ukraine and the demilitarisation of Bulgaria through the withdrawal of NATO's presence from the country. They claimed that Sofia's Alliance membership jeopardises Bulgaria's positive relations with Moscow, undermines sovereignty, and fosters geopolitical instability due to NATO's alleged incitement of conflicts [13].

Vazrazhdane's supporters (along with those of the Bulgarian Socialist Party) additionally protested in front of the Sofia Municipality building against the process of renaming streets adjacent to the Russian embassy 'Heroes of Ukraine' and 'Boris Nemtsov' [14]. A public petition that opposed the renaming gathered over 2600 signatures [15]. The party has also repeatedly attempted to forcefully remove the Ukrainian flag from the Municipality building.

These protest activities have been accompanied by an aggressive disinformation strategy. Vazrazhdane's posts on Facebook (and increasingly Telegram and TikTok), which promote anti-Western content, garner significant online engagement rate attracting thousands of likes and shares. Pro-neutrality positions are also amplified by mainstream print and online media. For instance, the daily newspaper *Trud* extensively publishes Vazrazhdane's statements and blames Ukraine's renunciation of a neutral position for causing the war [16]. Other newspapers like *Filter* claim that Bulgarian neutrality is the only commonsense position in the war. It draws historical analogies with World War I, where a similar lack of commonsense among Bulgarian 'hawks', who called for Bulgaria to take a certain stance, purportedly led to the country becoming party to the conflict [17].

THE EFFECTS

The protests in favour of neutrality have had a significant impact on Bulgaria's social and political scene, influencing attitudes, policy stances and even electoral outcomes.

From the outset of Russia's aggression against Ukraine, the far-right, nationalist end of the spectrum have been the most vociferous proponents of a pro-neutrality discourse, feeding into and shaping mainstream discussions on the issue. This discourse, also adopted by

[11] Mitov, B., 'Под руското знаме. "Възраждане" свика протест за "неутралитет" спрямо войната в Украйна' ['Under the Russian flag. Revival calls a protest for "neutrality" on the war in Ukraine'], Svobodna Evropa, 6 April 2022 (https://www.svobodnaevropa.bg/a/31788914.html).

[12] Darik News, 'Два протеста – „за" и „против" военна помощ за Украйна край парламента' ['Two protests – "for" and "against" military aid to Ukraine around Parliament'], 4 May 2022 (https://dariknews.bg/novini/bylgariia/dva-protesta-za-i-protiv-voenna-pomosht-za-ukrajna-kraj-parlamenta-2309360).

[13] Public Council 'Bulgaria for Peace and Neutrality', 'Позиция за мир, неутралитет и демилитаризация в България!' ['Position on peace, neutrality and demilitarization in Bulgaria!], 2022 (https://www.peticiq.com/353082).

[14] Karanyotova, K., 'Протест срещу алея "Героите на Украйна", пл. "Борис Немцов" и украинското знаме пред Столичната община' ['Protest against "Heroes of Ukraine" alley, "Boris Nemtsov" square and the Ukrainian flag in front of Sofia Municipality'], *BNT*, 21 April 2022 (https://bntnews.bg/news/protest-sreshtu-aleya-geroite-na-ukraina-pl-boris-nemcov-i-ukrainskoto-zname-pred-stolichnata-obshtina-1192402news.html).

[15] 'Петиция: Против преименуване на алеята до Руското посолство на "Героите на Украйна"' [Petition: Against the renaming of the alley next to the Russian Embassy into "Heroes of Ukraine"], Peticiq.com (https://www.peticiq.com/361121).

[16] '"Възраждане": България трябва да запази абсолютен неутралитет (ВИДЕО)' ['Vazrazhdane: Bulgaria needs to keep absolute neutrality (VIDEO]', *Trud,* 5 April 2022.

[17] Antonov, S., 'Криворазбраната диверсификация' ['Diversification wrongly understood'], *Filter*, Issue 15, 11 May – 17 May 2022, p. 19; Hristov, I., 'Русофили и русофоби да не развяват чужди знамена в България' ['Russophiles and Russophobes should not wave foreign flags in Bulgaria'], *Filter*, Issue 10, 6 April – 12 April 2022, p.21.

other pro-Russian political actors such as the Bulgarian Socialist Party, relies on propagandistic slogans. Calls for neutrality have been premised on establishing a false equivalence between weapons exports to Ukraine and the involvement of Bulgaria in a direct military confrontation with Russia (a supposed consequence said to amount to 'treason' against Bulgarian national interests). These arguments advocate Bulgaria's non-interference in the war and position Sofia as a potential mediator between Moscow and Kyiv.

In line with the prevailing public-political discourse, public opinion polls have reflected a widespread preference for neutrality. Over 60 % of polled Bulgarians expressed this view in 2022[18]. A majority (67 %) of the Bulgarian respondents did not think that a Russian victory in Ukraine would pose a threat to Bulgaria as the next target country[19]. And by the second half of March 2022, 42 % expressed the view that the EU's sanctions on Russia were too harsh[20].

Public opinion polls have reflected a widespread preference for neutrality.

At the same time, despite the presence of a neutrality-leaning majority, the pro-Russian rallies spearheaded by Vazrazhdane have sparked opposition from sections of society critical of Moscow, occasionally resulting in clashes. In April 2022, citizens gathered to protest against neutrality and expressed their support for Ukraine; in May 2022, a demonstration urging the government to send weapons to Kyiv took place alongside a protest against the provision of military aid; and later in that month a march against Russian fascism led to clashes with Kremlin supporters in front of the former monument to the Soviet army in Sofia[21]. The confrontations around the monument further reignited tensions over how to deal with the Russian and Soviet historical legacy in Bulgaria. A civic initiative known as 'Dusk for Light' called on mayors to extinguish lights illuminating Soviet statues, to indicate Bulgarians' readiness to break free from dependence on the Kremlin, support democratic values in Europe and Ukraine and ultimately remove these monuments symbolising the Soviet occupation of Bulgaria[22].

This polarisation was reflected in opinion polls conducted in 2022, 2023 and 2024. Respondents polled in 2022 were divided in their assessment as to whether Russia's attack on Ukraine could be justified or not: 41 % thought it could, whereas 44 % disagreed. In both 2023 and 2024, the same and unchanged proportion of Bulgarians surveyed, standing at 47 %, considered that the West and/or Ukraine could be blamed for the start of the war as against 44 %

[18] Kantar Public, *Uncertain times: The transmission of information and views on the war in Ukraine*, Free Press for Eastern Europe, July 2022; Hadjiski, V., '"Маркет линкс": Обществото е за неутралитет спрямо войната, подкрепящите членството в НАТО се увеличават' ['Market links: Society is in favour of neutrality regarding the war, those who support NATO membership are increasing'], *Dnevnik*, 1 April 2022 (https://www.dnevnik.bg/bulgaria/2022/04/01/4331564_market_links_obshtestvoto_e_neutralno_spriamo_voinata/?ref=home_layer2).

[19] *Uncertain times: the transmission of information and views on the war in Ukraine*, op.cit.

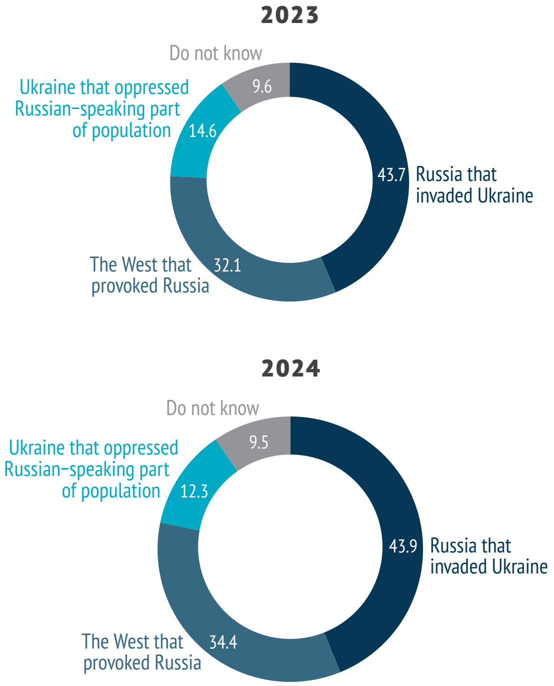
[20] Gallup International, '"Гласът на хората": Общественото мнение на европейските граждани за войната в Украйна' ['"The voice of the people": The public opinion of European citizens regarding the war in Ukraine'], 2022 (https://www.gallup-international.bg/45723/public-opinion-of-european-citizens-on-the-war-in-ukraine/).

[21] 'Два протеста – „за“ и „против“ военна помощ за Украйна край парламента' ['Two protests – "for" and "against" military aid to Ukraine around Parliament'], Darik News, 4 May 2022 (https://dariknews.bg/novini/bylgariia/dva-protesta-za-i-protiv-voenna-pomosht-za-ukrajna-kraj-parlamenta-2309360); 'Фасадата на руското посолство в София грейна с лика на Хитлер' ['The façade of the Russian Embassy in Sofia was lit up with Hitler's image'], Mediapool, 9 May 2022 (https://www.mediapool.bg/fasadata-na-ruskoto-posolstvo-v-sofiya-greina-s-lika-na-hitler-news335358.html).

[22] 'Кметове на България, изключете в Деня на независимостта прожекторите на съветските паметници!' ['Mayors in Bulgaria, turn off the projectors on the Soviet monuments on Independence Day!'], Marginalia, 2022 (https://www.marginalia.bg/aktsent/kmetove-na-balgariya-izklyuchete-v-denya-na-nezavimostta-prozhektorite-na-savetskite-pametnitsi/).

Who is to blame?

Bulgarian public opinion in 2023 and 2024 on Russia's full-scale invasion of Ukraine



Data: Adapted from Institute for Global Analytics, 'Bulgarian Public Opinion, 2024', p.9, 2024

who pinned responsibility for the aggression – in line with actual events – on the Kremlin[23]. More Bulgarians (41 % as against 34 %) considered that those parts of Ukraine which are predominantly inhabited by Russian-speaking populations should not belong to Russia. Yet a majority of Bulgarians (45 % vs 32 %) thought that Ukraine should not surrender to Russia. Bulgarian respondents were also supportive of Ukraine's Euro-Atlantic integration, with 53 % and 43 %, respectively, favouring Kyiv's EU and NATO accession[24].

The dominant preference for neutrality, coupled with opposition to Bulgaria's non-committal attitude to Russia's invasion of Ukraine on the part of a smaller but vocal part of society, may have contributed to the reticence in Bulgarian foreign policy. For example, particularly at the beginning of the war, the Bulgarian authorities' position on sending military assistance to Ukraine remained hesitant. Two months and a half after the start of the war, then Prime Minister Kiril Petkov attempted to mediate between different societal preferences and placate his coalition government partners (primarily the socialists who adamantly opposed sending weapons to Ukraine, maintaining that this would be an overt act of hostility against Moscow). This resulted in a compromise decision: the Bulgarian parliament voted to approve the provision of military-technical assistance focused on repairing Ukrainian military hardware, but not the direct shipment of weapons. This at once moved Bulgaria in the direction of contributing military aid, while stopping short of openly and publicly sending weapons exports outright as a way of appeasing pro-Russian forces[25]. Following the collapse of Petkov's government, caretaker governments appointed by President Rumen Radev stepped up their rhetoric on neutrality and alluded to a reversal of the parliamentary resolution. More weapons, they argued, would lead to an escalation of the war; peace, on the other hand, could only be achieved through negotiations[26]. However, the newly-formed government led by Prime Minister Nikolay Denkov once again demonstrated resolute support for Ukraine

[23] Filipova, R., 'Teetering on the brink of regional convergence: Bulgarians' stances on Russia's war against Ukraine, strategic orientation, democracy, media and values vis-à-vis Central and Eastern Europe', *Briefing Paper* No 3, Institute for Global Analytics, September 2023, p. 6 (https://globalanalytics-bg.org/2023/10/19/new-briefing-paper-teetering-on-the-brink-of-regional-convergence/); Filipova, R., 'Bulgarian public opinion, 2024: Increasing commitment to allies and improving perceptions of media freedom amid continuous East-West ambivalence', *Briefing Paper No 6,* Institute for Global Analytics, April 2024, p. 9 (https://globalanalytics-bg.org/2024/05/10/briefing-paper-6-bulgarian-public-opinion-2024/).

[24] *Uncertain times: The transmission of information and views on the war in Ukraine*, op.cit.

[25] Filipova, R., 'Bulgaria's balancing act', *Eurozine*, 2022 (https://www.eurozine.com/bulgarias-balancing-act/).

[26] Filipova, R., 'Bulgaria: Drifting apart from Europe', *Eurozine*, 2023 (https://www.eurozine.com/Bulgaria-drifting-apart-from-europe/).

through military aid and committed itself to combating Russian propaganda in Bulgaria.

The intensive promotion of pro-Russian and anti-NATO stances favouring neutrality, along with accompanying disinformation on social and mainstream media, has driven a significant increase in Vazrazhdane's electoral support. The party's share of the vote nearly tripled, increasing from 4.86 % in the November 2021 parliamentary elections, to 10.18 % in the October 2022 elections and to 14.16 % in April 2023[27]. In the June 2024 parliamentary elections, Vazrazhdane's support declined somewhat to 13.98 %, not least because another pro-Kremlin party, Velichie (translated as 'Greatness') made it to Parliament with 4.07 % of the vote[28]. Vazrazhdane's consistent presence, alongside that of other pro-Russian parties, in the Bulgarian political landscape means that it can continue to forcefully support the Kremlin's agenda.

> **The intensive promotion of pro-Russian and anti-NATO stances has driven a significant increase in Vazrazhdane's electoral support.**

THE RESPONSE

Since the start of Russia's war against Ukraine, parts of the Bulgarian political elite and the intelligence services have undertaken to expose the subversive activities of domestic pro-Russian actors. These actors operate within opaque and corrupt networks that Russia can manipulate to advance its own policy priorities, including through stoking division and tension in Bulgarian society.

In its annual report for 2021, Bulgaria's State Agency for National Security, which is responsible for counterintelligence, noted the Kremlin's extensive hybrid activities aimed at maintaining and expanding Russian influence in the Black Sea region. The report details the Agency's countermeasures against the activities of Russian intelligence services in Bulgaria[29]. Moreover, investigations by the Bulgarian counterintelligence agency that have not been fully disclosed publicly reportedly reveal that Bulgarian politicians, journalists and analysts receive payments from the Kremlin in order to influence public opinion in a pro-Russian direction[30]. In 2022 Bulgaria expelled 70 Russian embassy staff and more recently new cases of pro-Russian espionage have been exposed within the counterintelligence agency[31]. The decision to expel Russian diplomats was also motivated by intelligence information that they had conducted activities

[27] Central Election Commission, Избори за народни преставители [Elections for national representatives, 14 November 2021 – Results] (https://results.cik.bg/pvrns2021/tur1/rezultati/index.html); Central Election Commission, Избори за народни преставители, 2 октомври 2022 г. – резултати [Elections for national representatives, 2 October 2022 – Results] (https://results.cik.bg/ns2022/rezultati/index.html); Central Election Commission, Избори за народни преставители, 2 април 2023 г. – резултати [Elections for national representatives, 2 October 2023 – Results] (https://results.cik.bg/ns2023/rezultati/index.html).

[28] Central Election Commission, 'Избори за народни преставители, 9 юни 2024 г. – резултати' [Elections for national representatives, 9 June 2024 – Results] (https://results.cik.bg/europe2024/rezultati/index.html).

[29] '"Риск за сигурността на газовите доставки". Доклад на ДАНС говори за зависимостта от Русия' ['"A risk for the security of gas supplies". A report of the State Agency for National Security talks about the dependence on Russia'], Svobodna Evropa, 14 September 2022 (https://www.svobodnaevropa.bg/a/32033015.html).

[30] Milcheva, E., 'Кой взима хилядарките от Русия, нека разберем истината' ['Who takes money from Russia, let's learn the truth'], *Deutsche Welle*, 4 July 2022.

[31] 'ДАНС разкри руски разузнавачи, представяли се за българи' ['The State Agency for National Security exposed Russian intelligence officers, who posed as Bulgarians'], Svobodna Evropa, 26 February 2024 (https://www.svobodnaevropa.bg/a/dans-ruski-shpioni/32835953.html).

directed against the Bulgarian state, including hybrid attacks[32].

Despite such measures aimed at curtailing Russian influence in Bulgaria, however, political instability, inconclusive elections[33] and the fact that significant parts of the political spectrum demonstrate explicit or implicit pro-Russian leanings hamper the formulation of a consistent strategy against Russian interference in the spheres of the media, politics and economics. In particular, Bulgarian authorities' capacity to respond to information threats still needs to be improved. For a long period of time, the country's governance framework has been characterised by relative inaction and an insufficient recognition of the challenges posed by (pro-Kremlin) disinformation. While Prime Minister Denkov's government, which took office in June 2023, initiated a comprehensive assessment of national vulnerabilities within a risk-fraught global informational environment, more needs to be done in this vein. Moreover, frequent changes in government militate against continuity in tackling disinformation, whenever the latter is identified as a policy concern. For their part, EU-level initiatives have been slowly taken up without generating significant public debate about their merits and results. Civil society has therefore taken the lead in efforts to counter Russian media influence, as demonstrated by the development of media literacy tools and their introduction in the school curriculum. Civil society actors have also advocated for stronger journalistic practices.

Overall, therefore, the continuing hold of Russian propaganda over the Bulgarian public's mindset amid ongoing domestic social and political instability, creates a noticeable vulnerability. This situation, with the ever-present possibility of government collapse, provides a dangerous opening for the Kremlin to exploit societal divisions and pull Bulgaria towards the Russian orbit, undermining Sofia's Euro-Atlantic commitments.

[32] Hadjiski, V., 'Петков загатна за връзка на изгонените руски дипломати с хибридни атаки (обновена) ['Petkov alluded to a connection between the expelled Russian diplomats and hybrid attacks (updated)'], *Dnevnik*, 21 March 2022 (https://www.dnevnik.bg/bulgaria/2022/03/21/4326829_desetimata_ruski_diplomati_obiaveni_za_persona_non/?ref=home_layer2).

[33] Between April 2021 and June 2024 Bulgaria held six rounds of elections and a seventh round is expected in the autumn of 2024.

PART II: HACKING MACHINES

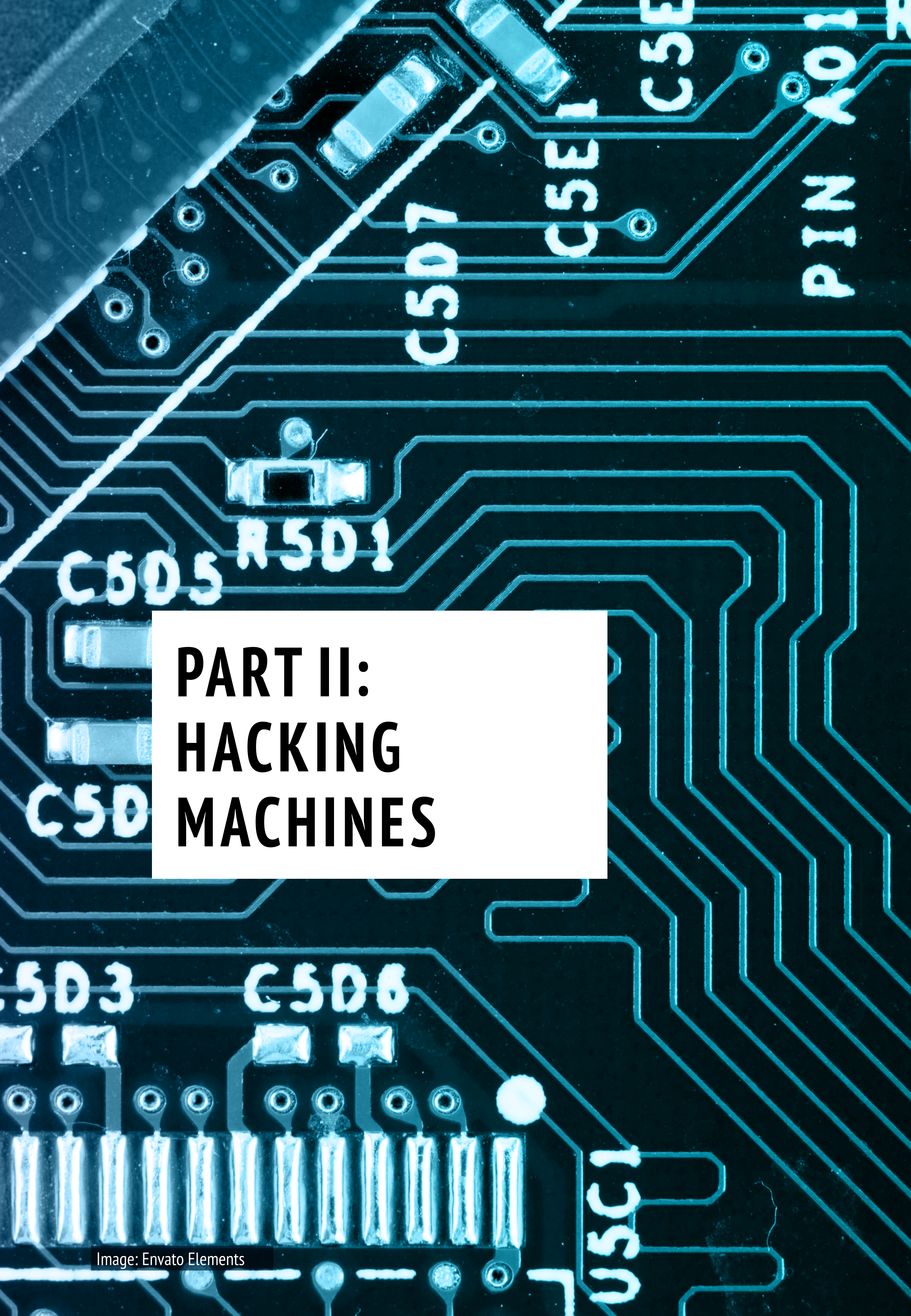


Image: Envato Elements

CHAPTER 3

THE SILENT SIEGE

Cyber-enabled IP theft in the foreign interference landscape

by
BART HOGEVEEN

INTRODUCTION

Almost two and a half years. That is how long a hacker group was able to snoop around undetected on the corporate networks of NXP, Europe's second-largest semiconductors manufacturer. This extensive access gave them ample time to become familiar with the company's IT systems, allowing them to steal chip designs and sift through financial data. The breach was only detected in early 2020, two years after US chip giant Qualcomm launched a bid to acquire NXP[1]. The takeover, worth USD 44 billion, failed when the Chinese regulator withheld approval[2].

Did the Chinese authorities possess information that they should not have had, or gain insights through unfair means, to influence this commercial decision?

Intellectual property (IP) is an essential asset underpinning economies that are driven by research, innovation and access to new technologies. Therefore, IP protection is a fundamental pillar of the trading regime of the World Trade Organization (WTO) through the Agreement on Trade-Related Aspects of Intellectual Property Rights. The European Patent Office estimates that 47 % of the EU's GDP is generated by IP-intensive businesses[3].

As the advanced economies of the 2010s were at the peak of digitisation, Western governments also began to realise that certain industries and sectors were becoming targets of sophisticated cyber-enabled infiltrations. In 2008, amidst broader leadership and management issues, the Canadian telecommunications manufacturer Nortel collapsed after it came to light that design documents of their then world-leading fibre-optics equipment

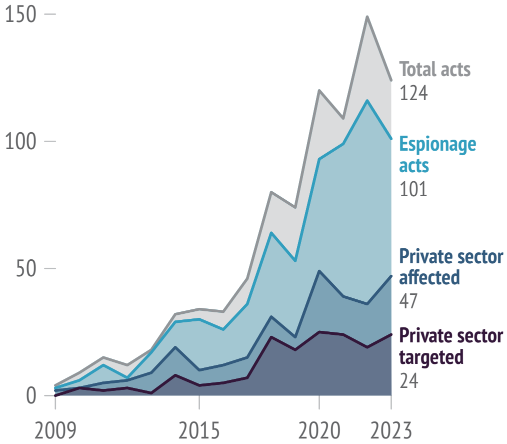
[1] Hijink, M., 'Spionage: Chinese hackersgroep zat jarenlang in het netwerk van de Nederlands chipfabrikant NXP' [Espionage: Chinese hacker group was hiding in the network of Dutch chip manufacturer NXP for years], NRC, 24 November 2023 (https://www.nrc.nl/nieuws/2023/11/24/spionage-chinese-hackersgroep-zat-jarenlang-in-het-netwerk-van-de-nederlandse-chipfabrikant-nxp-a4182149).

[2] Clark, D., 'Qualcomm scraps $44 billion NXP deal after China inaction', *New York Times*, 25 July 2028 (https://www.nytimes.com/2018/07/25/technology/qualcomm-nxp-china-deadline.html).

[3] European Patent Office and EU Intellectual property Office, 'IPR-intensive industries and economic performance in the European Union: Industry-level analysis report', 4th edition, October 2022 (https://documents.epo.org/projects/babylon/eponet.nsf/0/33DCE530D888258BC12588D7004539D1/$File/ipr-intensive_industries_and_economic_performance_in_the_EU_2022_en.pdf).

Spike in state-sponsored cyber operations

Increase in reported incidents of state-sponsored operations from 2009 to 2023



Data: Council on Foreign Relations, Cyber Operations Tracker, 2024

had been exfiltrated a few years previously[4]. In 2014, the US Department of Justice indicted hackers of the People's Liberation Army (PLA) for cyber-enabled economic espionage against companies in the nuclear power, metals and solar industries[5]. Between 2009 and 2022, the number of cases of state-sponsored cyber campaigns quadrupled, of which 20-30 % relate to espionage efforts targeting private firms.

That is the background against which the leaders of the United States and China reached an agreement in 2015[6], and later together with the other members of the G20, stipulating that: 'no country should conduct or support ICT-enabled theft of intellectual property, including trade secrets or other confidential business information, with the intent of providing competitive advantages to companies or commercial sectors'[7].

Theft of IP is not a new phenomenon, nor is economic espionage. While most countries have legislation in effect that criminalises both practices on their territory, international law does not explicitly prohibit state-on-state espionage even when it is of an economic nature[8]. State-on-company cyberespionage, however, is a different ballgame.

The G20 agreement provides a framework for addressing cyber-enabled IP theft by state actors. The agreement condemns state-sponsored cyber operations that steal IP from companies and research institutions in other jurisdictions and when they then provide local companies with an unfair competitive advantage. With illegally acquired IP, local beneficiary companies would be able to leapfrog research and development time and costs, offer services and products below competitive market prices, manipulate trade contracts or dispute settlements and 'squat' patents.

This chapter looks at the phenomenon of state-supported acts of cyber-enabled theft of intellectual property[9] (also known as

[4] Calof, J., 'An overview of the demise of Nortel Networks and key lessons learned: Systemic effects in environment, resilience and black-cloud formation', University of Ottawa, February 2014 (https://sites.telfer.uottawa.ca/nortelstudy/files/2014/02/nortel-summary-report-and-executive-summary.pdf); Cooper, S., 'Inside the Chinese military attack on Nortel', *Global News*, 25 August 2020 (https://globalnews.ca/news/7275588/inside-the-chinese-military-attack-on-nortel/).

[5] US Department of Justice, 'US charges five Chinese military hackers for cyber espionage against US corporation and a labor organization for commercial advantage', Press release, 19 May 2024 (https://www.justice.gov/opa/pr/us-charges-five-chinese-military-hackers-cyber-espionage-against-us-corporations-and-labor).

[6] The White House, 'Remarks by President Obama and President Xi of the People's Republic of China in Joint Press Conference, 25 September 2015 (https://obamawhitehouse.archives.gov/the-press-office/2015/09/25/remarks-president-obama-and-president-xi-peoples-republic-china-joint).

[7] G20 Leaders' Communiqué Antalya Summit, 15-16 November 2015, paragraph 26 (https://www.consilium.europa.eu/media/23729/g20-antalya-leaders-summit-communique.pdf).

[8] Lotrionte, C., 'Countering state-sponsored cyber economic espionage under international law', *North Carolina Journal of International Law and Commercial Regulation*, Vol. XL, 2015, pp 443-538 (https://securitypolicylaw.syr.edu/wp-content/uploads/2015/06/Lotrionte_Countering_State_Sponsored_Cyber_Economic_Espionage.pdf).

[9] Intellectual property (IP) refers to items that are the property of the mind or proprietary knowledge. It involves patents, trademarks, registered designs, geographical indications and copyrights but also sensitive business information and trade secrets as well as foundational (unpublished) research.

economic cyberespionage) for commercial gain and establishes where this is becoming a form of foreign interference. The latter refers to activities of foreign states that are coercive, corrupting, deceptive and clandestine in nature, and seek to advance those states' strategic, political, military, social or economic goals at the expense of the victim nation's sovereignty, values and national interests [10].

THE INCIDENT

To further demonstrate how economic cyberespionage unfolds, the following section describes a campaign of 'massive IP theft' that was reported by a cybersecurity company and that explicitly draws a connection between IP theft, state-supported hackers and financial consequences.

In May 2022 the US-Israeli cybersecurity group Cybereason disclosed Operation CuckooBees: a campaign where hackers managed to exfiltrate hundreds of gigabytes of information from some 30 multinational companies worth trillions of US dollars. The attackers spent years undetected to stealthily 'conduct reconnaissance and identify valuable data'. They targeted intellectual property developed by the victim companies, including 'sensitive documents, blueprints, diagrams, formulas, and manufacturing-related proprietary data' [11].

Cybereason holds the Winnti Advanced Persistent Threat (APT) group responsible for the operation. This group shares a significant number of tactics, techniques and procedures (TTPs) with APT41. Western intelligence agencies consider them to be affiliated with China's Ministry of State Security (MSS) [12]. The designation APT is given to 'continuous, clandestine, and sophisticated hacking techniques' that amount to 'complex and covert cyber-attacks' and which involve 'highly skilled actors, usually targeting high-profile organisations'. They are often sponsored or supported by state agencies [13]. The MSS, China's agency responsible for collecting intelligence and its conglomerate of operating units [14], including at provincial and municipal levels, has been described by the US government as 'an intelligence enterprise that includes contract hackers who also conduct unsanctioned cyber operations worldwide, including for their own personal profit' [15].

In the case of Operation CuckooBees, hackers gained access through basic means of compromise, such as spearphishing [16] and social engineering. After this initial entry they deployed sophisticated TTPs that included selective targeting, relentless persistence and

[10] Based on the definition of foreign interference used by the Australian Government Department of Home Affairs (https://www.homeaffairs.gov.au/about-us/our-portfolios/national-security/countering-foreign-interference/).

[11] Cybereason, 'Operation CuckooBees: Cybereason uncovers massive Chinese intellectual property theft operation', 4 May 2022 (https://www.cybereason.com/blog/operation-cuckoobees-cybereason-uncovers-massive-chinese-intellectual-property-theft-operation). In October 2022, cybersecurity firm Symantec added to this reporting, noting a continuation of the same type of activities targeting organisations in Hong Kong (https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/spyder-loader-cuckoobees-hong-kong).

[12] Mandiant, 'APT41, a dual espionage and cyber crime operation', Report, February 2022 (https://www.mandiant.com/sites/default/files/2022-02/rt-apt41-dual-operation.pdf); US Department of Justice, 'United States of America v Jinag Lizhi, Qian Chuan and Fu Qiang', Grand Jury, 7 May 2019 (https://www.justice.gov/opa/press-release/file/1317206/download).

[13] HackerOne, 'What are Advanced Persistent Threats?' (https://www.hackerone.com/knowledge-center/advanced-persistent-threats-attack-stages-examples-and-mitigation); Kaspersky, 'What is an Advanced Persistent Threat (APT)?' (https://www.kaspersky.com/resource-center/definitions/advanced-persistent-threats).

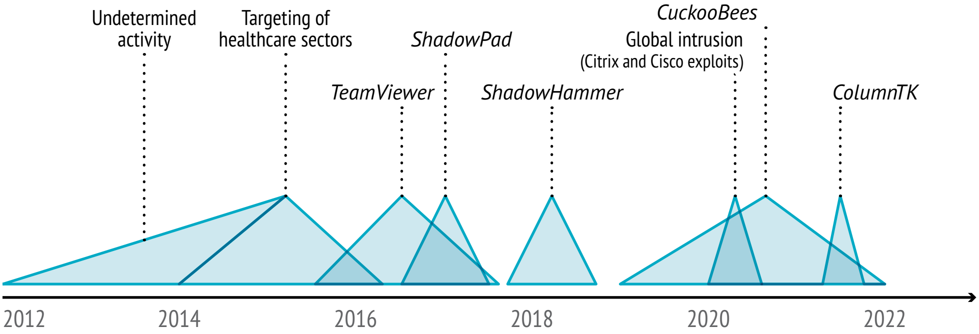
[14] Ibid.

[15] The White House, 'The United States, joined by allies and partners, attributes malicious cyber activity and irresponsible state behavior to the People's Republic of China', Statement, 19 July 2021 (https://www.whitehouse.gov/briefing-room/statements-releases/2021/07/19/the-united-states-joined-by-allies-and-partners-attributes-malicious-cyber-activity-and-irresponsible-state-behavior-to-the-peoples-republic-of-china/). See also: Uren, T., 'Risky Biz Briefing: The i-SOON data leak', *Seriously Risky Business*, 22 February 2024 (https://news.risky.biz/risky-biz-briefing-the-i-soon-data-leak/).

[16] Attempts, often through email, to trick people into revealing sensitive information (personal, financial or account data) or installing malware.

APT 41

Timeline of attacks



Data: The Mitre Corporation, 2022

agility when victim organisations plugged holes in their IT security systems. Their toolset included the misuse of supply chain compromises, compromised digital certificates, and bootkits – assets that require a high level of technical proficiency and expertise to deploy covertly [17].

Targets involved technology and manufacturing companies located in Asia, Europe and North America, and the attacks mirrored the tactics of APT 41 which tends to 'target industries in a manner generally aligned with China's Five-Year economic development plans' [18]. The attackers focused on the healthcare sector, pharmaceuticals, high-tech semiconductors and advanced hardware, electrical vehicles, and telecommunications. Cybersecurity researchers also believe this APT aims to 'gather intelligence ahead of imminent events, such as mergers and acquisitions and political events' [19].

THE EFFECTS

Measuring the impact of cyberespionage and IP theft is notoriously difficult. Cases like the NXP hack and Operation CuckooBees show that the consequences of economic cyberespionage are often not discerned or recognised until years after the incident [20]. Also, most companies – even when aware of a compromise – are often reluctant to disclose breaches to shareholders let alone local authorities. For instance, NXP, after they discovered that they had been compromised, downplayed the significance of the infiltration, maintaining that it had not incurred 'a material adverse effect on our business' [21]. Finally, establishing a causal relationship between the infiltration of IT networks and the loss of profitability and competitiveness is near-impossible. As a result, the true effects of these activities can only be estimated.

[17] Mandiant, 'APT41, a dual espionage and cyber crime operation', op.cit.; TrendMicro, 'Hack the real box: APT41's new subgroup Earth Longzhi', 9 November 2022 (https://www.trendmicro.com/en_us/research/22/k/hack-the-real-box-apt41-new-subgroup-earth-longzhi.html). Kaspersky defines *bootkits* as 'a malicious program designed to load as early as possible in the boot process, in order to control all stages of the operating system start up, modifying system code and drivers before anti-virus and other security components are loaded'. (https://encyclopedia.kaspersky.com/glossary/bootkit/).

[18] Mandiant, 'APT41, a dual espionage and cyber crime operation', op.cit., page 6.

[19] Ibid.

[20] AIVD, MIVD and NCTV, 'Dreigingsbeeld Statelijke Actoren 2' ['Threat Assessment State Actors 2'], November 2022 (https://www.nctv.nl/documenten/publicaties/2022/11/28/dreigingsbeeld-statelijke-actoren-2).

[21] NXP Semiconductors, *Annual Report for the Financial Year ended 31 December 2020*, p. 54, (https://www.nxp.com/docs/en/supporting-information/2020-IFRS-STATUTORY-ANNUAL-REPORT.pdf).

Business continuity risks for victims

Victim organisations are first and foremost confronted with financial risks. In the pre-compromise phase, entities at the coalface of strategic competition are confronted with elevated costs related to hardening business resilience, tightened cybersecurity controls and service packages and topped-up insurance premiums. Some may even opt out from a market or R&D investment if they do not stand a chance of protecting their IP and business operations [22]. Further risks pertain to the aftermath of an actual breach, which could involve loss in royalties from patents and trademarks, payments resulting from liability, litigation actions or regulatory fines; and costs of repairing and restoring compromised IT systems [23]. Investments in time- and resource-intensive R&D can be considered further write-offs.

In the long run, a successful cyberespionage campaign can cause a company to lose its order portfolio, access to markets and disadvantage an entire sector to the detriment of a country's economy. The Nortel case suggests it may indeed contribute to the collapse of a company; a telecommunications company that would be considered a critical infrastructure entity today.

Impact on prosperity and economic competitiveness

The risk surface for the theft of IP and innovation is substantial, and cyberespionage can have a significant economic impact. In the EU, the top-1000 R&D companies together invested some €230 billion in innovation in 2022 [24]. At the macro-level of the economy, estimations suggest that the US economy is losing circa USD 400 billion per annum (2018) and the EU around €60 billion (2020). This amounts to 1-3 % of annual national GDP. Cybereason, in their report of the incident, speak of trillions of US dollars' worth of IP stolen in Operation Cuckoobees. But when successful acts of economic cyberespionage structurally target certain sectors and companies – as we have seen from the incidents described – it diverts away foreign direct investment and venture capital [25]. A country simply becomes a less attractive investment market with adverse consequences for growth, employment and prosperity. Emerging economies and industries that are part of transnational value chains are particularly vulnerable.

A report produced for the European Commission established that manufacturing, information and communication technologies, finance, health and medical technologies are the most impacted sectors of IP theft in the EU and concluded that 'cyber-misappropriation of trade secrets' is clearly aligned with areas of strategic economic competition [26]. Therefore, the effects of economic cyberespionage need to be measured against Europe's objectives of

[22] Ibid., pp. 54-55.

[23] Gelinne, J.P., Francher, D. and Mossburg, E., 'The hidden costs of an IP breach: cyber theft and the loss of intellectual property', *Deloitte Review*, No 19, 25 July 2016 (https://www2.deloitte.com/us/en/insights/deloitte-review/issue-19/loss-of-intellectual-property-ip-breach.html).

[24] European Commission, '2023 EU Industrial R&D Investment Scorecard', 2023, p. 65 (https://op.europa.eu/en/publication-detail/-/publication/1e5c204f-9da6-11ee-b164-01aa75ed71a1/language-en),

[25] OECD, 'Enquiries into Intellectual Property's economic impact', 2015, pp. 8-10 (https://one.oecd.org/document/DSTI/ICCP(2014)17/CHAP1/FINAL/En/pdf); Chen, Y. and Puttitanun, T., 'Intellectual property rights and innovation in developing countries', *Journal of Development Economics*, Vol. 78, 2005, pp. 474 – 493 (https://qed.econ.queensu.ca/pub/faculty/lloyd-ellis/econ835/conf07/sluys.pdf); Guo, D. and Jiang, K. 'Venture capital investment, intellectual property rights protection and firm innovation: evidence from China', *Entrepreneurship & Regional Development*, Vol. 34, Issue 5-6, 2022 (https://www.tandfonline.com/doi/full/10.1080/08985626.2022.2062618).

[26] European Commission, DG internal Market, Industry, Entrepreneurship and SMEs, 'The scale and impact of industrial espionage and theft of trade secrets through cyber', Report by PwC, December 2018 (https://op.europa.eu/en/publication-detail/-/publication/4eae21b2-4547-11e9-a8ed-01aa75ed71a1/language-en).

strategic autonomy which depend on sovereign innovation capabilities and competitive indigenous (academic) research[27]. Successful campaigns of economic cyberespionage by competitors or adversaries of significant scale and targeted at vulnerabilities and dependencies would undermine this ambition.

Impact on global governance

Persistent campaigns of cyber-enabled theft of innovation also undermine the global system of governance that guides state behaviour in cyberspace, international security and international trade and competition. The international community has invested deeply in elaborating a framework for responsible state behaviour in cyberspace which includes the recognition that international law applies in this domain and a set of agreed rules and norms, including the G20 commitment[28]. Systemic forms of IP theft also put the international trade regime under stress, in particular the IP protection and non-technical barriers to trade arrangements of the WTO. This reverberates in bilateral and regional free trade agreements as these rely on WTO standards and dispute settlement mechanisms[29]. The erosion of existing global economic security regimes, mainly the WTO and G20, is already pushing countries towards 'trusted geographies' and de-risking strategies. While these strategies aim to mitigate risk, they also come at a cost and can create new vulnerabilities[30].

Systemic forms of IP theft put the international trade regime under stress.

THE RESPONSE

Responses to malicious state-sponsored cyber activities that target intellectual property have predominantly occurred in domestic or minilateral settings, and include:

Counter-intelligence: Cybereason shared their threat intelligence with the FBI. In response to the growing caseload of state-sponsored economic cyberespionage, the FBI together with their UK and Five Eyes'[31] counterparts, have initiated high-level outreach activities to various business communities. In July 2022, the directors of the UK and US security agencies MI5 and FBI addressed a business audience in London to warn about the risk of doing business with China[32]; and in October 2023, the chiefs of the Five Eyes security services spoke to the Silicon Valley technology community about China's theft of innovation[33].

Criminal investigations: Law enforcement action has been a central part of the US response to 'the ravages of theft and destruction' through cyberespionage[34]. Data from

[27] Powell, C., Tocci, N. and Wolff, G., 'Making European strategic autonomy work', *The Strategist*, Australian Strategic Policy Institute, 27 November 2023 (https://www.aspistrategist.org.au/making-european-strategic-autonomy-work/).

[28] Perlroth.N. and Sanger, D.E., 'Obama calls for new cooperation to wrangle the "Wild West" internet', *New York Times*, 13 February 2015 (https://www.nytimes.com/2015/02/14/business/obama-urges-tech-companies-to-cooperate-on-internet-security.html); Priyandita, G. and Hogeveen,B., 'State-sponsored economic cyber-espionage for commercial purposes: tackling an invisible but persistent risk to prosperity', Australian Strategic Policy Institute, December 2020, p. 17 (https://www.aspi.org.au/report/state-sponsored-economic-cyberespionage).

[29] Australian Strategic Policy Institute, 'Hacking for cash', podcast, 11 August 2023 (https://www.aspi.org.au/news/pgm-hacking-cash-francis-gurry-nigel-corey-and-elizabeth-chien-hale).

[30] Schaus, M. and Lannoo, K., 'The EU's aim to de-risk itself from China is risky…yet necessary', CEPS, 7 September 2023 (https://www.ceps.eu/the-eus-aim-to-de-risk-itself-from-china-is-risky-yet-necessary/).

[31] The intelligence-sharing alliance of Australia, Canada, New Zealand, UK and the US.

[32] MI5, 'Joint address by MI5 and FBI Heads', July 2022 (https://www.mi5.gov.uk/news/speech-by-mi5-and-fbi).

[33] FBI, 'Emerging Technology and Securing Innovation Security Summit', fireside chat moderated by Condoleeza Rice, October 2023 (https://www.fbi.gov/video-repository/101723_fireside_chat_01.mp4/view).

[34] Goldsmith, J. and Williams, T.D., 'The failure of the United States' Chinese-hacking indictment strategy', Lawfare, December 2018 (https://www.lawfaremedia.org/article/failure-united-states-chinese-hacking-indictment-strategy).

Operation CuckooBees is still to appear in criminal indictments related to APTs and other state-supported hackers, but 'cyber sanctions' and offers of rewards for information leading to the identification of APT leaders and affiliates have become part of the response portfolios of the US, EU and others [35].

Public attributions: Operation CuckooBees has not been the subject of a government-led attribution, but other cases of economic cyberespionage have. In 2018, the UK and partners attributed the Cloud Hopper campaign [36] and in 2021 a grand coalition of eight countries, NATO and the EU attributed the large-scale misuse of Microsoft Exchange vulnerabilities – both with references to the G20 commitment [37]. These attributions serve to express a unity of views, signal to the responsible state that their *modus operandi* has been detected and reinforce the continued application of the norm [38].

Cybersecurity research: The forensic analysis of Operation CuckooBees has informed follow-up research and analysis by other researchers and cybersecurity companies. The data Cybereason was able to share has made its way into the service packages and threat detection and mitigation tools that companies worldwide deploy and offer to their customers.

Public advisories: Because of continuous reports of economic cyberespionage, such as Operation Cuckobees, national cybersecurity centres have become more forthcoming in sharing information in the public domain. In particular the US Cybersecurity and Infrastructure Security Agency (CISA) has published a suite of technical advisories related to state-sponsored APTs [39], and several intelligence and security agencies are addressing economic security and cyberespionage in their annual reports. They also publish specific guidance to industry on protecting their business and IP [40]. These advisories are essential early warning instruments for companies to assess their level of risk exposure and take necessary protective steps.

National cyber defence capabilities and updates to legal frameworks. Furthermore, states have updated legal frameworks and strengthened national cyber defence capabilities. Examples include the UK's National Security Law and the establishment of the National Cyber Force [41], and Japan's Economic Security Protection Act in combination with a strengthened role for the National Centre of Incident Readiness and Strategy for Cybersecurity (NISC) [42].

[35] US Department of State, 'Cyber sanctions' (https://www.state.gov/cyber-sanctions/);Australian Department of Foreign Affairs and Trade, 'Significant cyber incidents sanctions regime' (https://www.dfat.gov.au/international-relations/security/sanctions/sanctions-regimes/significant-cyber-incidents-sanctions-regime); EU, Council Regulation 2019/796 concerning restrictive measures against cyber-attacks threatening the Union or its Member States, 17 May 2019 (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019R0796).

[36] UK Government, 'UK and allies reveal global scale of Chinese cyber campaign', December 2018 (https://www.gov.uk/government/news/uk-and-allies-reveal-global-scale-of-chinese-cyber-campaign).

[37] Euronews, 'Microsoft Exchange email hack came from China, say EU and US', 19 July 2021 (https://www.euronews.com/next/2021/07/19/microsoft-exchange-email-hack-came-from-china-say-eu-and-us).

[38] Rupp, C. and Paulus, A., 'Official public political attribution of cyber operations: state of play and policy options', SNV Berlin, October 2023 (https://www.stiftung-nv.de/sites/default/files/official-public-political-attribution-of-cyber-operations.pdf).

[39] US Cybersecurity and Infrastructure Security Agency, 'Chinese state-sponsored cyber operations: Observed TTPs', August 2021 (https://www.cisa.gov/news-events/cybersecurity-advisories/aa21-200b).

[40] UK National Protective Security Authority, 'Secure innovation', October 2023 (https://www.npsa.gov.uk/secure-innovation); Netherlands Intelligence and Security Service, 'Cyber-attacks by state actors: seven moments to stop an attack', November 2021 (https://english.aivd.nl/publications/publications/2021/11/29/cyber-attacks-by-state-actors-seven-moments-to-stop-an-attack).

[41] UK Ministry of Defence, 'Industry Security Notice: the National Security Act 2023', January 2024 (https://assets.publishing.service.gov.uk/media/65ba6951c75d300012ca0ff3/ISN_2024-01_National_Security_Act_2023-O.pdf); UK National Cyber Force, 'The National Cyber Force: responsible cyber power in practice', April 2023 (https://assets.publishing.service.gov.uk/media/642a8886fbe620000c17dabe/Responsible_Cyber_Power_in_Practice.pdf).

[42] European Parliament, 'At a glance: Japan's economic security legislation', July 2023 (https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/751417/EPRS_ATA(2023)751417_EN.pdf); Author's briefings with NISC and JPCERT.

CONCLUSION

While practices of economic cyberespionage are not confined to China, several sources point to its activities as being the most concerning on a global scale. States such as Russia, France, Israel, Vietnam and Iran have also appeared on watchlists[43], but the sheer scale, depth and persistence of China's cyberespionage programme set it apart according to the assessment of numerous intelligence agencies[44].

The key factor in all of this is the institutional environment that Beijing has created where cyber-enabled theft of IP is facilitated, encouraged and directed. Chinese APTs select their targets in technology sectors where China sees a domestic capability shortfall[45]. The Chinese authorities provide hackers – serving members of government, contractors and affiliates – with protection from potential criminal prosecutions[46]. And the introduction of expanded regulations for industry to report software vulnerabilities to the authorities is providing state hackers with advanced access to backdoors and exploits[47]. This creates a significant concern for foreign interference, as it allows China to acquire sensitive information and technological advancements through coercive, corruptive, deceptive and clandestine means.

Chinese APTs select their targets in technology sectors where China sees a domestic capability shortfall.

The policy responses, as described, will not prevent or deter these states from continuing their campaigns, in particular when they are pursued in the name of national security. So, as geostrategic competition intensifies and more nations develop into knowledge economies, the risk that economic cyberespionage will affect more economies – developed and emerging – and across multiple sectors of industry is growing.

Given the invisible character of most theft of economic assets and the long lag in time between compromise and discovery, government policy should be directed towards greater investments in understanding the threat landscape and establishing a qualitative and quantitative appreciation of the risks and costs to the economy.

This would allow legislators and regulators to expand cybersecurity initiatives beyond government agencies and critical infrastructure entities. Such an approach would encompass critical economic sectors as well as those heavily reliant on IP, including (academic) research institutions. This would significantly enhance cybersecurity awareness, maturity, and the adoption of best practices across the entire economy. Furthermore it would encourage greater investment in preventative measures. One could argue that economic crown jewels and critical technological capabilities are not just 'too important to fail', but also 'too important to be stolen'.

Diplomatically, governments should re-energise their G20 commitment and make cybersecurity risks to economic and

[43] 'State-sponsored economic cyber-espionage for commercial purposes: tackling an invisible but persistent risk to prosperity', op.cit.

[44] Ibid, page 17; German Domestic Intelligence Service, *2022 Report on the Protection of the Constitution: Facts and trends*, June 2023 (https://www.verfassungsschutz.de/SharedDocs/publikationen/EN/reports-on-the-protection-of-the-constitution/2023-06-brief-summary-2022-report-on-the-protection-of-the-constitution.pdf).

[45] Mattis, P., 'A guide to Chinese intelligence operations', War on the Rocks, August 2015 (https://warontherocks.com/2015/08/a-guide-to-chinese-intelligence-operations/).

[46] Ibid.

[47] Cary, D. and Del Roso, K., 'Sleight of hand: How China weaponizes software vulnerabilities', Atlantic Council, September 2023 (https://www.atlanticcouncil.org/in-depth-research-reports/report/sleight-of-hand-how-china-weaponizes-software-vulnerability/).

knowledge security part of further international cybersecurity negotiations, either in the G20 or as part of the UN forums such as the Open-ended Working group on ICT security. Given the significance of the interference and the actors behind campaigns of economic cyberespionage, this phenomenon ought to be seen in the same league as offensive cyber operations and attacks against critical infrastructure. Governments should also seek to update current WTO arrangements on intellectual property to reflect the reality of digital trade and cyber-dependent IP, and the distinct cybersecurity threats to trade, economic cooperation and innovation.

The lingering uncertainty surrounding the NXP incident exemplifies the challenges of attributing and quantifying the impact of cyber-enabled espionage. We do not know if the Chinese authorities managed to extract sensitive business information from NXP, and whether that influenced their decisions. If it were the case, we will probably only come to learn years from now. The fact, however, that the incident occurred is already affecting NXP's operating costs and casting a shadow over its reputation. This underscores the need for companies, especially those in critical sectors, to invest heavily in cybersecurity measures, and for national governments to have strong early warning mechanisms in place to counter potential cyber-enabled foreign interference.

CHAPTER 4

THE ATTACK OF THE CLONES

Deepfakes and the evolving landscape of disinformation

by
ANDREA SALVI

INTRODUCTION

As the world braces for the November 2024 US presidential elections, an incident unfolded in New Hampshire early in the year, showing the potential impact of deepfakes in the political arena. Reports came to light that automated calls impersonating President Joe Biden had discouraged Democrats from voting in the state's presidential primary in January 2024[1]. These calls, engineered through AI generative technology, have raised concern among experts. Beyond their immediate impact on voter behaviour, such 'deep-faked' content poses broader threats to the integrity of electoral systems and public discourse and hence to democratic processes.

Deepfakes are broadly defined as AI-generated content mimicking a real human being through digitally manipulated audio, video or images that convincingly depict people doing or saying things they never actually did or said. To use the taxonomy provided by a 2017 Council of Europe report[2], deepfakes encompass several categories such as 'manipulated content', 'imposter content', and 'fabricated content'[3]. The technical foundation for deep fakes is typically identified in Generative Adversarial Networks (GANs). These generative models, which have only recently emerged,[4] are a form of unsupervised machine learning and, in lay terms, are characterised by two 'competing' sub-models: a generator and a discriminator model. The former produces fake imagery, while the latter works as a classifier to assess whether a given image is a fake or not. The rationale behind this adversarial relationship consists in the fact that if the fake is spotted, the generative model will adjust

[*] The author would like to thank Fee-Marie von der Brelie for her research assistance.

[1] Murphy, H., 'Audio deepfakes emerge as a weapon of choice in election disinformation', *Financial Times*, 23 January 2023 (https://www.ft.com/content/bd75b678-044f-409e-b987-8704d6a704ea).

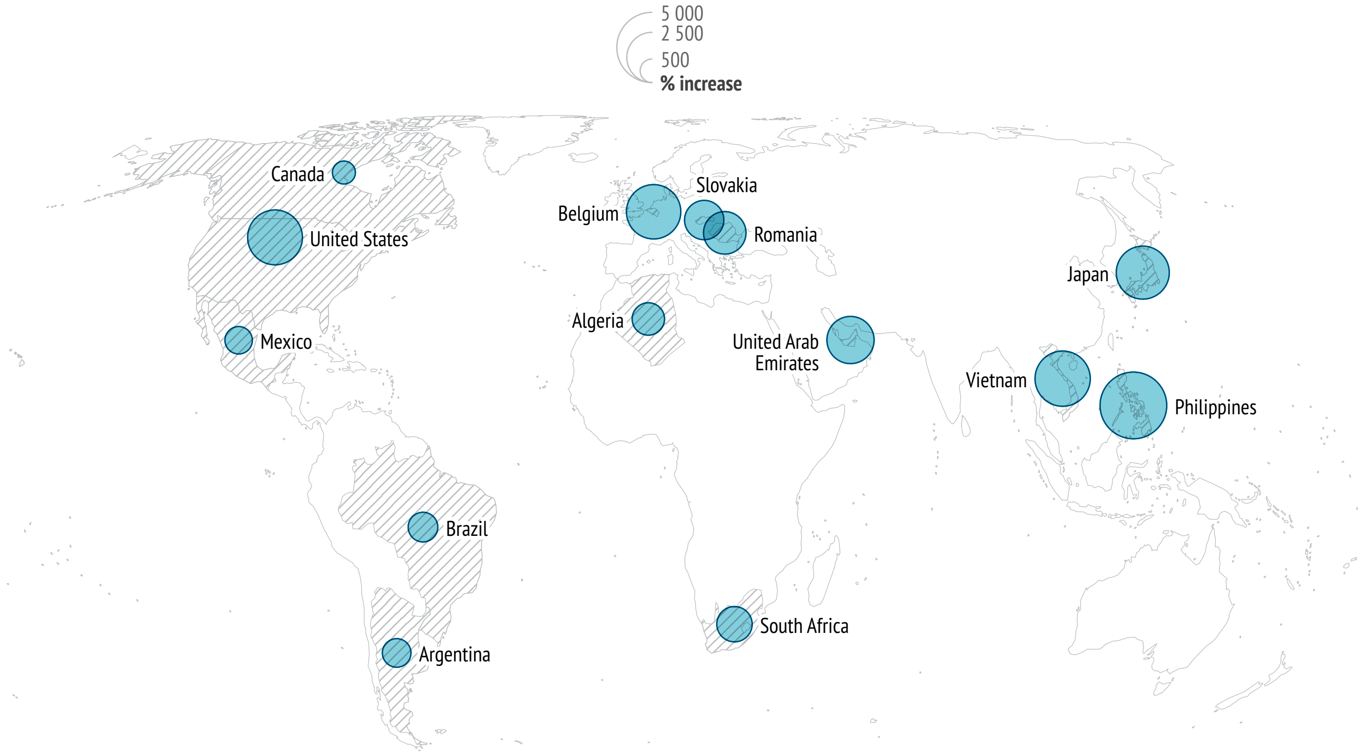
[2] Wardle, C. and Derakhshan, H., 'Information Disorder: Toward an interdisciplinary framework for research and policy making', Council of Europe, September 2017 (https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html#).

[3] Dobber, T., Metoui, N., Trilling, D., Helberger, N. and de Vreese, C., 'Do (microtargeted) deepfakes have real effects on political attitudes?' *The International Journal of Press/Politics*, Vol. 26, No 1, 2021, pp. 69–91.

[4] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 'Generative Adversarial Networks', *Communications of the ACM*, Vol. 63, No 11, 2014, pp.139–144.

AI-driven fraud

Top-ranked countries by region with the biggest increase in deepfake incidents, 2022-2023



Data: Sumsub Identity Fraud Report, 2023

accordingly and improve. In essence, 'competition' and interplay between the sub-models leads progressively to the production of more credible output.

While most technologies related to data and generative content, GANs and similar models are inherently neutral, they can be misused for malicious purposes. There are many benign applications that can dramatically improve the automation of tasks and enhance image analysis in various fields, including in the medical domain with AI-augmented diagnostics[5]. However, the same capabilities can empower malign actors to carry out broader disruptive interference operations.

This chapter explores deepfakes as a prime example of how convergence between machines and information manipulation unfolds. Although the focus here is on deepfakes, which are generative models of videos and images, the analysis can be applied to a broader category of AI-generated content such as texts, audio and other fabricated constructs.

Firstly, the chapter presents and examines incidents caused by deepfakes. Secondly, it discusses the consequences of weaponising AI-generated content, illustrating current applications and their impact. Thirdly, it explores potential and current responses to address this challenge, highlighting existing policy actions.

[5] Sundaram, S. and Hulkund, N., 'GAN-based data augmentation for chest X-ray classification.' *ACM Computing Surveys*, March 2021 (https://arxiv.org/abs/2107.02970v1).

THE INCIDENT

To illustrate the ways in which cyber-enabled deepfake tools augment information manipulation, this section explores specific areas where their effects and implications have significant impacts. Specifically, it will examine three primary categories: (i) the use of deepfake technology in election interference; (ii) its ability to cause reputational harm and damage in contexts of geopolitical tensions; and (iii) its role in armed conflict.

Category 1: Election Interference

Countries are particularly vulnerable to acts of foreign interference in the form of information manipulation during the lead-up to elections. Deepfakes can target key political figures depicting them engaging in scandalous activities or making controversial statements. These manipulated or AI-generated videos spread rapidly on social media and other digital platforms. Some of them are easily identifiable as false and it may be tempting to simply dismiss them for that reason. However, they are becoming increasingly sophisticated and harder to distinguish from reality. A video released by the US Republican Party in 2023 showed an AI-generated projection of the country's future if President Joe Biden were to be re-elected in 2024: it depicted catastrophic crises and a Chinese invasion of Taiwan. While this is a legitimate use, it can give the reader a sense of the potential effect of AI-generated content used for political purposes and its potential impact on shaping public perceptions. Moreover, the proliferation of fake videos and fraudulent phone calls has raised significant concerns regarding the manipulation of information through AI, with potentially harmful effects. These developments raise red flags about the ethical use of AI in disseminating information within political contexts. Aside from fake phone calls, in the context of the US presidential elections mentioned in the introduction to this chapter, numerous deepfake videos have been circulating. They mostly feature President Joe Biden and former Democratic presidential contender Hillary Clinton, making controversial declarations [6]. This content can in fact influence public perception and lead to further polarisation, thereby compromising democratic institutions.

If AI-generated content is a matter of concern for the US and the EU, this is even more the case in countries with a deeply fractured political landscape, where the phenomenon could wreak havoc in elections. As an example, just prior to the crucial parliamentary and presidential election in Türkiye in May 2023, during a political rally President Erdogan presented a video splicing footage of his challenger Kemal Kilicdaroglu and Murat Karayilan, one of the founding members of the Kurdistan Workers' Party (PKK), a Kurdish militant organisation designated as a terrorist group by both the European Union and the United States. The video appeared to show that Karayilan and other PKK militants were participating in Kilicdaroglu's election campaign. Despite later revelations that it had been fabricated, the video proved damaging to Kilicdarogl's campaign, highlighting the disruptive potential of AI-generated content in democratic processes and political contexts.

Category 2: Reputational damage in contexts of geopolitical tensions

As tensions rise between countries both online and in the physical world, malicious actors are increasingly utilising AI-generated fake content to magnify conflicts and stir up sentiment against adversaries and certain groups. Deepfakes make information manipulation, a tactic long used to interfere in other countries'

[6] Ulmer, A. and Tong, A., 'Deepfaking it: America's 2024 election collides with AI boom', *Reuters*, 31 May 2023 (https://www.reuters.com/world/us/deepfaking-it-americas-2024-election-collides-with-ai-boom-2023-05-30/).

affairs [7], much more powerful. Fabricated atrocities, or misleading attribution of serious crimes and wrongdoings, are typical examples of such manipulative practices. Generative technologies can *de facto* enable malicious actors to create convincingly realistic deceptive videos that could potentially instigate cycles of retaliation or escalations of conflict, and sabotage diplomatic efforts. The propagation of deep fakes in such a context can erode the fragile *modus vivendi* between actors and hamper prospects of peaceful coexistence or conflict resolution.

In other words, through the strategic manipulation of emotions and beliefs – exploiting societal vulnerabilities, political or ethnic tensions – falsified content possesses the ability to reinforce existing divisions and polarise communities. In such operations, the quality of the artifact becomes of secondary importance as its content unleashes destabilising dynamics that can have destructive consequences. For example, a controversy was sparked by the alleged video-confession of Phyo Min Thein, former chief minister of the Yangon Region in Myanmar. In this video, the former official declared that he had offered bribes to the Burmese leader Aung San Suu Kyi [8]. Despite the artificial look of the video – due to the voice distortion and static facial expression of the speaker – this was used as one of the key pieces of evidence to add allegations of corruption against the Burmese leader.

The war in Ukraine has been accompanied by the proliferation of deepfakes.

Category 3: Destablisation in situations of armed conflict: The case of Ukraine

The war in Ukraine has been accompanied by the proliferation of deepfakes. For example, several weeks after Russia's invasion of the country in February 2022, a video portraying President Zelensky was widely circulated on Twitter, Facebook and YouTube, and then posted on the Russian social media platform VKontakte. In the short video, the President is shown addressing the soldiers and citizens of Ukraine and imploring them to surrender and lay down their weapons. The video was subsequently dismissed by Zelensky himself as a 'childish provocation'. While the puppeteers behind 'Zelensky's clone' remain unknown, the defence Intelligence office of Ukraine had previously warned about Russia's tactic of disseminating deepfakes as part of its information warfare strategy. As a further demonstration that this episode was part of a broader toolkit of interference, the fake message from Zelensky was briefly broadcast on the scrolling news section of Ukraine 24, a national television channel, and displayed on their website [9]. This was allegedly due to a group of hackers who were able to compromise the cyber architecture of the news outlet. Similar techniques have also been used against Russia. In June 2023 another video depicted Vladimir Putin calling for martial law in the Russian Federation. Once more, the video was allegedly broadcast on several radio stations and TV networks.

At times, forged images have been created to boost morale and weaken the resolve of opponents. It is worth mentioning the so-called 'Ghost of Kyiv' a Ukrainian fighter ace whose heroic deeds were relayed on internet

[7] See for instance: Zhukova, E., 'Image substitutes and visual fake history: historical images of atrocity of the Ukrainian famine 1932–1933 on social media', *Visual Communication*, Vol. 21, No 1, 2022, pp.3–27.

[8] Gregory, S., 'The world needs deepfake experts to stem this chaos', *Wired*, 24 June 2021 (https://www.wired.com/story/opinion-the-world-needs-deepfake-experts-to-stem-this-chaos/).

[9] Allyn, B., 'Deepfake video of Zelenskyy could be "tip of the iceberg" in info war, experts warn', *npr*, 26 March 2022 (https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia).

platforms, broadcasts and even on former president Petro Poroshenko's social media account. The pilot, portrayed in the fashion of a contemporary Red Baron, allegedly shot down six Russian jets on the first day of the war. Subsequently, stories, photos and videos of the pilot and of his combat engagements continued to circulate. Some media outlets reported that the heroic pilot had died after taking down over 40 Russian fighter planes. All these stories were fabricated, or at least they coalesced real stories extrapolated from different pilots as well as pre-war photos, cutscenes from videogames and other content. Although not strictly a 'deepfake' – as most of this material was not AI-generated – this illustrative case reveals the potency of fabricated content.

The reports from Ukraine show how stories orchestrated through AI-generated technology, although quickly debunked, garnered significant attention, were taken up by major news outlets, and had the potential to distort people's perception of the conflict. These techniques can have an even more powerful impact in cases when fabricated 'message(s) of surrender' are circulated at the local level, for instance among isolated troops, first responders and frontline operators. As highlighted, the use of malicious cyber operations to propagate these messages can further amplify negative dynamics[10].

THE EFFECTS

The malicious use of deep fake technology poses a significant threat by exacerbating the challenge of discerning truth from deception in video, audio and photographic content. The most immediate effect materialises in potential 'direct disruptions' that deepfakes may cause. This is especially true in high-pressure and dynamically uncertain environments (or 'extreme contexts'[11]) where sabotage to communications, decoy operations and disinformation may seriously compromise the situational awareness of the actors involved. Front-line organisations operating in environments characterised by limited information – 'informational asymmetries' – are particularly exposed to such incidents. Examples include: first responders, emergency workers and soldiers in conflict zones. These environments are often characterised by the 'fog of war', where information is scarce and debunking malicious fabricated content can be challenging. While most frontline personnel are exposed to such threats, an organisation's ability to respond quickly and its overall resilience[12] are crucial for mitigating disruptions and countering potential interference. In short, direct disruptions through deepfakes can give a tactical edge to the perpetrator.

At a more strategic and subtle level, perpetrators may be playing a long game: deepfakes can generate 'indirect disruptions'. AI-generated visuals – and disinformation mediums in general – are capable of instilling doubt in individuals or communities exposed to them. As recounted in one article devoted to the subject[13], they can inject false news in a media ecosystem or subtly alter people's perceptions

[10] Dreikhausen V. and Salvi A., 'From cyber to hearts and minds: Cyber operations and the battle for global influence', *Brief* No 19, EUISS, November 2023 (https://www.iss.europa.eu/content/cyber-hearts-and-minds).

[11] Hannah, S. T., Uhl-Bien, M., Avolio, B. J. and Cavarretta, F. L., 'A framework for examining leadership in extreme contexts.', *Leadership Quarterly*, Vol. 20, No 6, 2009, pp. 897–919.

[12] Weick, K. E. and Sutcliffe, K. M., 'Mindfulness and the quality of organizational attention', *Organization Science*, Vol.17, No 4, 2006, pp. 514–524.

[13] 'Do (microtargeted) deepfakes have real effects on political attitudes?', op.cit.

regarding the legitimacy of democratic institutions[14], impair the quality of public debate, undermine citizens' ability to participate effectively in the political process[15], erode trust in conventional media[16], and enhance the political capabilities as well as the credibility of the narrative peddled by the perpetrators[17]. Furthermore, deepfakes can build upon and amplify pre-existing disinformation by reinforcing already-held biases, further entrenching false narratives on contentious issues. The low barrier of access enables malicious actors to craft micro-targeted deepfakes, for example by using psychological techniques to tailor messages to appeal to specific groups[18].

The ease with which deepfakes can be created and distributed is clearly a cause for concern[19]. Software and tools for generating deepfakes are now readily available: DeepFaceLab[20], for instance, is one of the most popular packages for creating high-quality face-swapping videos. A provocative use of the package was proposed by the US-based non-profit organisation 'RepresentUS'. They created a series of ads called 'Dictators', featuring Kim Jong Un and Vladimir Putin, to support their 'Save The Vote' campaign. The ads were allegedly set to be aired on prominent outlets such as Fox News and CNN following the 2020 presidential debates, to raise awareness of the dangers of deepfake technology. According to RepresentUS, the broadcasts were rejected at the last minute and without a clear explanation; the series was therefore published on YouTube[21]. Its basic tutorial consists of a video that lasts roughly 30 seconds: while more complicated forgeries require more resources, the barrier of access remains extremely low.

Incidents involving deepfakes have given rise to widespread concern due to both their direct and indirect effects. Structured campaigns of interference that harness the ever-growing capabilities of AI have the potential to cause serious harm both at the individual and corporate level – through psychological and financial damage. Moreover, at the societal level, they can destabilise political ecosystems, mould and influence public opinion, and undermine democratic processes[22].

THE RESPONSE

Addressing the challenges posed by deepfakes and by AI-generated content requires a comprehensive and agile response strategy with engagement from the broader multistakeholder community. Policymakers, IT specialists, researchers and civil society need to actively monitor the broader information and cybersecurity ecosystem to stay ahead of malign actors and adopt reactive institutional

[14] Bennett, W. L. and Livingston, S., 'The disinformation order: Disruptive communication and the decline of democratic institutions', *Sage Journals,* Vol. 33, No 2, 2018, pp. 122–139.

[15] Flynn, D. J., Nyhan, B. and Reifler, J., 'The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics', *Political Psychology*, Vol. 38, 2017, pp. 127–150.

[16] Vaccari, C. and Chadwick, A., 'Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news', *Social Media+ Society*, Vol. 6, No 1, 2020, pp.1–13.

[17] Bradshaw, S. and Howard, P. N., 'The global organization of social media disinformation campaigns', *Journal of International Affairs*, Vol. 71, No 1.5, 2018, pp.23–32.

[18] See for instance: Wheeler, S. C., DeMarree, K. G. and Petty, R. E., 'A match made in the laboratory: Persuasion and matches to primed traits and stereotypes', *Journal of Experimental Social Psychology*, Vol 44, No 4, 2008, pp.1035–1047; Chang, C., 'Seeing the small picture: AD-self versus ad-culture congruency in international advertising', *Journal of Business and Psychology*, Vol. 20, No 3, 2006, pp. 445–465.

[19] Generation of such models requires a powerful 'consumer level' graphic card (GPU), among other components. To date, a top-tier GPU has a retail price of roughly €2 000.

[20] Perov, I., et al., 'DeepFaceLab: Integrated, flexible and extensible face-swapping framework', *ariXiv*, May 2020 (https://arxiv.org/abs/2005.05535v5). The package is readily available at: https://github.com/iperov/DeepFaceLab.

[21] von Mueffling, D., 'First ever use of deepfake technology in major ad campaign', *representUs*, 29 September 2020 (https://act.represent.us/sign/deepfake-release/).

[22] See van Huijstee, M. and van Boheemen, P., 'Tackling deepfakes in European policy', European Parliamentary Research Service, July 2021 (https://doi.org/10.2861/325063).

forms and practices[23]. This section focuses on policy responses, but it is worth mentioning how technical measures – such as monitoring tools and watermarking of AI-generated content – play a crucial role in countering the effect of malicious use of AI-generated content.

The ease with which deepfakes can be created is clearly a cause for concern.

Institutions and Law Enforcement Agencies (LEAs) are aware of the threat constituted by deepfakes. In 2022, Europol published an extensive report on the challenges posed by AI-generated content aimed at disinformation[24]. The report examines the impact of deepfakes on law enforcement and the legal system, highlighting the need for stricter evaluation, verification and detection techniques. It highlights the importance of adapting regulatory frameworks, including laws, policies and practices, for LEAs, service providers, and other organisations both in public and private domains.

The EU has been at the forefront of the effort to combat manipulated media content. In particular it has been reinforcing its defensive toolkit in collaboration with the multistakeholder community and with the private sector. The Strengthened Code of Practice on Disinformation[25] released in 2022 sought to address the shortcomings of the 2018 iteration. It has gained approval from various actors in the information ecosystem, including big tech firms, and civil society. It establishes a code for the industry to combat disinformation and create a transparent, open and safe internet. The Code clearly identifies deepfakes as one of the 'manipulative behaviours' that private entities and institutions must fight against. On AI-generated content, in June 2023, European Commission Vice-President for Values and Transparency Věra Jourová declared that the Code's signatories will need to boost their detection capabilities to clearly label content as generated by a machine.

This new code will be closely tied to the Digital Services Act (DSA)[26]. It establishes severe penalties for disinformation as well as fines for platforms that fail to comply with the obligations. Platforms will have to demonstrate their procedures for reporting and taking down illegal content as well as content identified as information manipulation. The regulatory push aims at introducing transparency obligations since 'self-regulation has not worked' as declared by the Chair of Internal Market and Consumer Protection Committee of the European Parliament Anna Cavazzini[27]. Based on the declaration of a Meta Whistleblower, the MEP highlighted how there is a fundamental clash between the effort to fight disinformation and business models underpinned by algorithms that aim at maximising visibility of specific content.

As a further specific measure targeting deepfakes, the Artificial Intelligence Act mandates disclosure of AI-generated content by its creators under a specific transparency clause (Article 52(3)). In the risk-based classification of AI applications the act frames deepfakes as constituting 'limited risk'. This categorisation is problematic as it results in the imposition of relatively minimal transparency requirements.

[23] Spagnoletti, P., Ceci, F. and Salvi, A., 'Adversarial Evolution: Competing dynamics and reactive institutional forms in financial services ecosystem', *ITASEC*, 2021, pp. 406–413.

[24] Europol, 'Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab', Publications Office of the European Union, April 2022 (https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf).

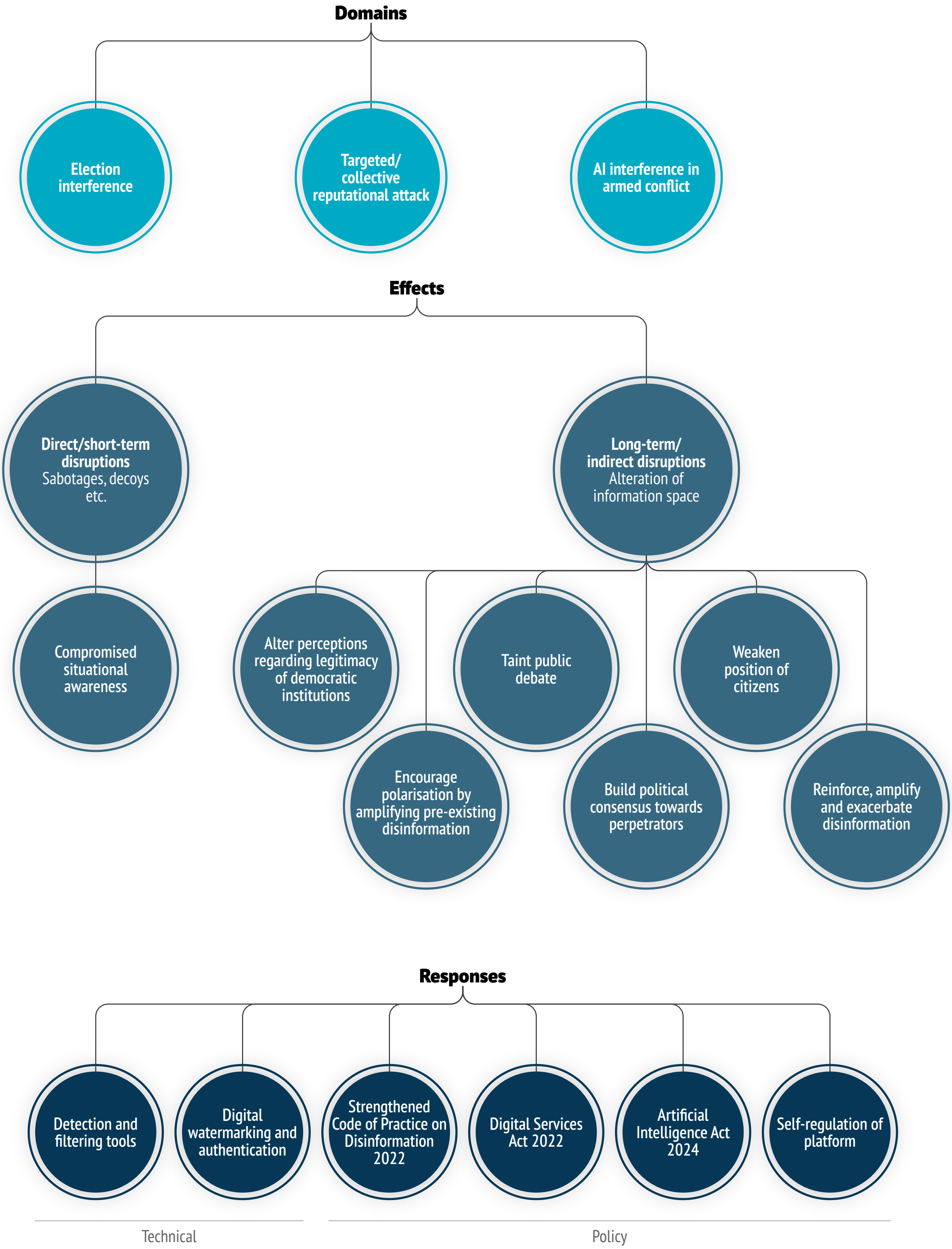
[25] European Commission, 'Strengthened Code of Practice on Disinformation - Shaping Europe's digital future', June 2022 (https://ec.europa.eu/newsroom/dae/redirection/document/87585).

[26] European Parliament and the Council, Regulation on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), Regulation 2022/2065, 19 October 2022 (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R2065).

[27] Nadkarni, I.T., 'Facebook files: MEPs to invite whistleblower Frances Haugen to a hearing', European Parliament, 11 October 2021 (https://www.europarl.europa.eu/news/de/press-room/20211011IPR14619/facebook-files-meps-to-invite-whistleblower-frances-haugen-to-a-hearing).

Framework of AI implications in the information environment

Navigating domains, effects and responses


Domains
Election interference
Targeted/ collective reputational attack
AI interference in armed conflict
Effects
Direct/short-term disruptions Sabotages, decoys etc.
Compromised situational awareness
Long-term/ indirect disruptions Alteration of information space
Alter perceptions regarding legitimacy of democratic institutions
Taint public debate
Weaken position of citizens
Encourage polarisation by amplifying pre-existing disinformation
Build political consensus towards perpetrators
Reinforce, amplify and exacerbate disinformation
Responses
Detection and filtering tools
Digital watermarking and authentication
Strengthened Code of Practice on Disinformation 2022
Digital Services Act 2022
Artificial Intelligence Act 2024
Self-regulation of platform
Technical
Policy

Data: Hannah et al. (2009), Bennett & Livingston (2018), Yiping et al (2019), Flynn et al. (2017), Vaccari & Chadwick (2020), Bradshaw & Howard (2018), Wheeler et al. (2008), Chang (2006), European Commission (2021 & 2022), European Parliament & the Council (2022), Spagnoletti et al (2021)

Furthermore, the AI Act does not spell out obligations with explicit sanctions for non-compliance, resulting in weak incentives for adherence to these transparency rules. The AI act, however, will leave open possibilities to amend the classification under Article 67 if Member States' regulatory authorities identify further risks. The intrinsic dual-use nature of generative technology as well as its polyvalency in terms of creation and detection of deepfake artifacts could complicate the process of assessing their true risks[28].

While progress is being made in addressing the challenges generated by the sudden advent of consumer AI, more reflection is needed on how to tackle deepfake content hosted outside the control of mainstream regulated platforms. Rogue websites, messaging apps as well as darknet markets provide havens for malicious content. LEAs play a critical role in addressing this issue. Increased resources and training are necessary to equip them with the skills to investigate deepfake-related crimes. Investing in the 'light side' of AI – which has been already proven successful in other areas[29] – can boost detection and takedown capabilities.

CONCLUSION

Deepfakes are widely used as instruments of information manipulation. This practice shows the potency of deepfakes in distorting perceptions, moulding public opinion and in polluting the information ecosystem. Given the risk of political spillovers and destabilisation, this chapter has underscored the importance of developing effective and comprehensive resilience against AI-driven information manipulation in the digital space.

As the technology behind deepfakes keeps advancing and becoming more sophisticated, its potential to disrupt counter-interference measures scales up exponentially. To effectively confront this threat, stakeholders need to create collaborative platforms to work on legal, technological and societal resilience aspects. Through its multilateral efforts the EU is well-positioned to advocate for responsible AI use, enhancement of detection capabilities, and deterring information manipulation involving generative technologies. A transparent, collaborative and efficient regulatory process holds the potential to generate momentum in international bodies and institutions and boost EU diplomacy.

At the same time, it is crucial to reduce technical and societal vulnerabilities by investing in the broader cyber ecosystem. This endeavour should not only encompass the cyber domain but also include efforts to foster media literacy and skills. This can be facilitated by developing confidence-building measures and capacity-building measures to empower individuals and organisations to leverage technology effectively. This can be achieved by establishing robust legal frameworks and creating forums for dialogue around shared values that can help build consensus on how to address deepfakes and other online threats. The EU Digital Decade targets, the EU Cyber Solidarity Act, the Cybersecurity Skills Academy, as well as various EU-funded projects on cybersecurity, cyber diplomacy and on countering information manipulation serve as a model for this comprehensive approach. These initiatives demonstrate the importance of coordinated action at all levels to build a more resilient digital environment.

[28] Fernandez A., 'Regulating deep fakes in the proposed AI Act', *MediaLaws: Law and Policy of the Media in a Comparative Perspective*, 23 March 2022 (https://www.medialaws.eu/regulating-deep-fakes-in-the-proposed-ai-act/).

[29] Salvi, A., Spagnoletti, P. and Noori, N. S., 'Cyber-resilience of critical cyber infrastructures: Integrating digital twins in the electric power ecosystem', *Computers & Security*, Vol. 112, No 8, 2022.

PART III: HACKING THE ECOSYSTEM

Image: Shahadat Rahman/Unsplash

CHAPTER 5

OUT OF SIGHT, OUT OF MIND?

A cyber-FIMI nexus in critical infrastructure protection

by
PATRYK PAWLAK

INTRODUCTION

Today, hardly any sector or country has been spared from cyberattacks, whether in peace or war. In February 2024, several US agencies linked the Chinese APT group Volt Typhoon to a series of operations that could disrupt critical communications infrastructure between the United States and the Asia region during future crises. Ukrainian infrastructure operators are constantly targeted by malicious operations, including intelligence collection campaigns aimed at gleaning information about Ukrainian plans regarding strategic entities such as the Zaporizhzhya Nuclear Power Plant (ZNPP) [1]. While cyber interference operations targeting critical infrastructure have been addressed in policy and academic debates, the junction between cybersecurity and manipulation of the information environment remains under-researched, except in the context of hybrid conflicts. This neglect is worrying given the increasing use of information manipulation in conjunction with cyberattacks.

The case of interference operations against Albania by Iran illustrates this point well. In July 2022, Albania was confronted with an unprecedented attack on its systems which rendered government websites and services inaccessible. Later, ransomware and destructive malware were deployed. Interestingly, cyber operations against Albania were accompanied by the extensive use of information manipulation techniques. 'Homeland Justice' – the Iranian state cyber actor who claimed credit for the attack – left an anti-Mujahideen E-Khalq (MEK) [2] message on computers infected with ransomware and created a website and multiple social media profiles with similar content [3]. The social media campaign that lasted several weeks is believed to have had the purpose of undermining authorities in Tirana in retaliation for its sheltering of the MEK

[1] State Service of Special Communications and Information Protection of Ukraine, *Russia's Cyber Tactics – Analytical Report*, 25 September 2023 (https://cip.gov.ua/en/news/yak-zminyuyutsya-taktiki-cili-i-spromozhnosti-khakerskikh-grup-uryadu-rf-ta-kontrolovanikh-nim-ugrupovan-zvit).

[2] The Mujahideen E-Khalq (MEK) is an exiled Iranian opposition group that settled in Albania after being evacuated from Iraq in 2016.

[3] Vicens, A., 'Albania says Iranian hackers hit the country with another cyberattack', *CyberScoop*, 12 September 2022 (https://cyberscoop.com/iranian-cyberattack-albania-homeland-justice/).

members who settled in Albania after being evacuated from Iraq in 2016.

The cross-domain analysis approach presented in the second report published by the European External Action Service (EEAS) on Foreign Information Manipulation and Interference (FIMI) provides a glimpse into how the link between cybersecurity and FIMI could be further operationalised [4], but it does not address a critical point which is that the information and cyber environments cannot be addressed separately from one another. Therefore, building on the existing research and empirical work [5], this chapter calls for a more rigorous approach to designating 'critical information environment infrastructure' and inclusion of the information environment as a key component of the discussion about critical infrastructure protection.

THE INCIDENT

Critical infrastructure systems depend on connectivity for production, distribution and delivery of their services. A growing number of stakeholders in the infrastructure ecosystem – including across the whole supply chain – creates additional risks. In the energy sector, for instance, those risks extend from resource transportation and power plant infrastructure to the precision timing and communication networks essential for grid management. In order to design an effective policy response, it is important to understand connections and dependencies between four categories of incidents in the critical infrastructure environment:

> **Category 1. Critical infrastructure as a target**: Attacks against specific sectors designated by governments as critical due to their strategic importance, such as transportation, health or energy services. In general, any unauthorised attempt to access government or corporate networks that compromises the availability, authenticity, integrity or confidentiality of data is considered an incident. Notable examples of incidents in the energy sector, for instance, include the Brazilian Companhia Energética de Minas Gerais (2020) and the Colonial Pipeline (2021).

> **Category 2. Communications networks and internet infrastructure as a target**: These systems represent a sub-category of critical infrastructure. They are often included as a critical sector and the backbone for other sectors. This is particularly the case of satellites and undersea cables as attacks on them not only undermine communication flows but also jeopardise the delivery of other internet-dependent services. For example, in 2022 the broadband satellite internet access provided by the US firm Viasat's KA-SAT network in Ukraine was disrupted by a cyberattack.

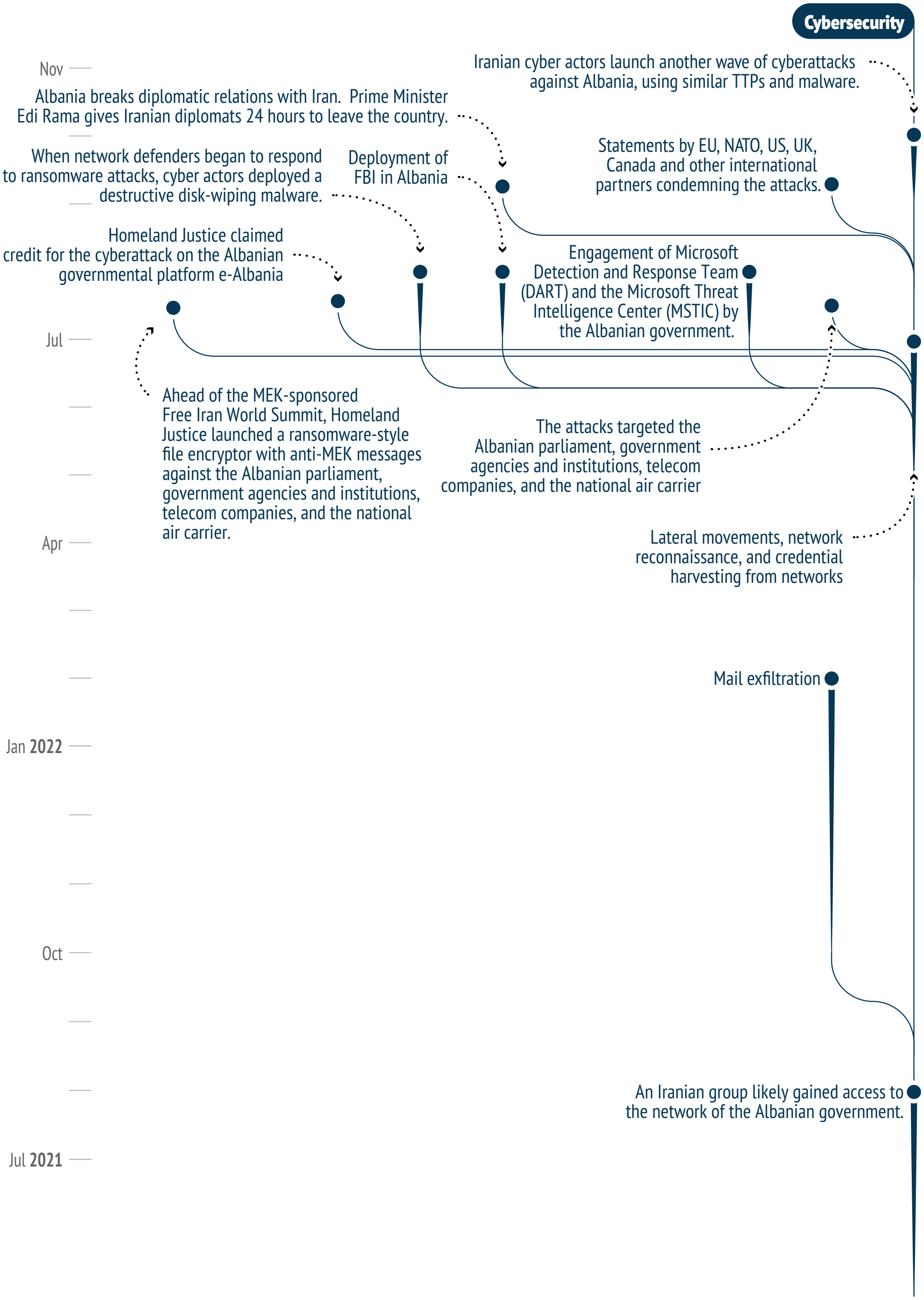
> **Category 3. Communication transmission networks as a tool to intercept information**: Communication infrastructure can be abused by states for cyber surveillance or spying to obtain information of strategic importance in the political, economic or security domains. The use of tools such as Pegasus falls in this category. Submarine cables that constitute the backbone of the global economy and telecommunications are also prone to nation-state sabotage and spying, although not many cyberattacks

[4] European External Action Service, '2nd EEAS Report on Foreign Information Manipulation and Interference Threats – A framework for networked defence', January 2024. (https://www.eeas.europa.eu/sites/default/files/documents/2024/EEAS-2nd-Report%20on%20FIMI%20Threats-January-2024_0.pdf).

[5] Wanless, A. and Shapiro, J. N., 'A CERN model for studying the information environment', Carnegie Endowment for International Peace, 17 November 2022 (https://carnegieendowment.org/2022/11/17/cern-model-for-studying-information-environment-pub-88408).

Case study of Albania

Cybersecurity and information operations



Iranian cyber actors launch another wave of cyberattacks against Albania, using similar TTPs and malware.

Albania breaks diplomatic relations with Iran. Prime Minister Edi Rama gives Iranian diplomats 24 hours to leave the country.

Statements by EU, NATO, US, UK, Canada and other international partners condemning the attacks.

When network defenders began to respond to ransomware attacks, cyber actors deployed a destructive disk-wiping malware.

Deployment of FBI in Albania

Homeland Justice claimed credit for the cyberattack on the Albanian governmental platform e-Albania

Engagement of Microsoft Detection and Response Team (DART) and the Microsoft Threat Intelligence Center (MSTIC) by the Albanian government.

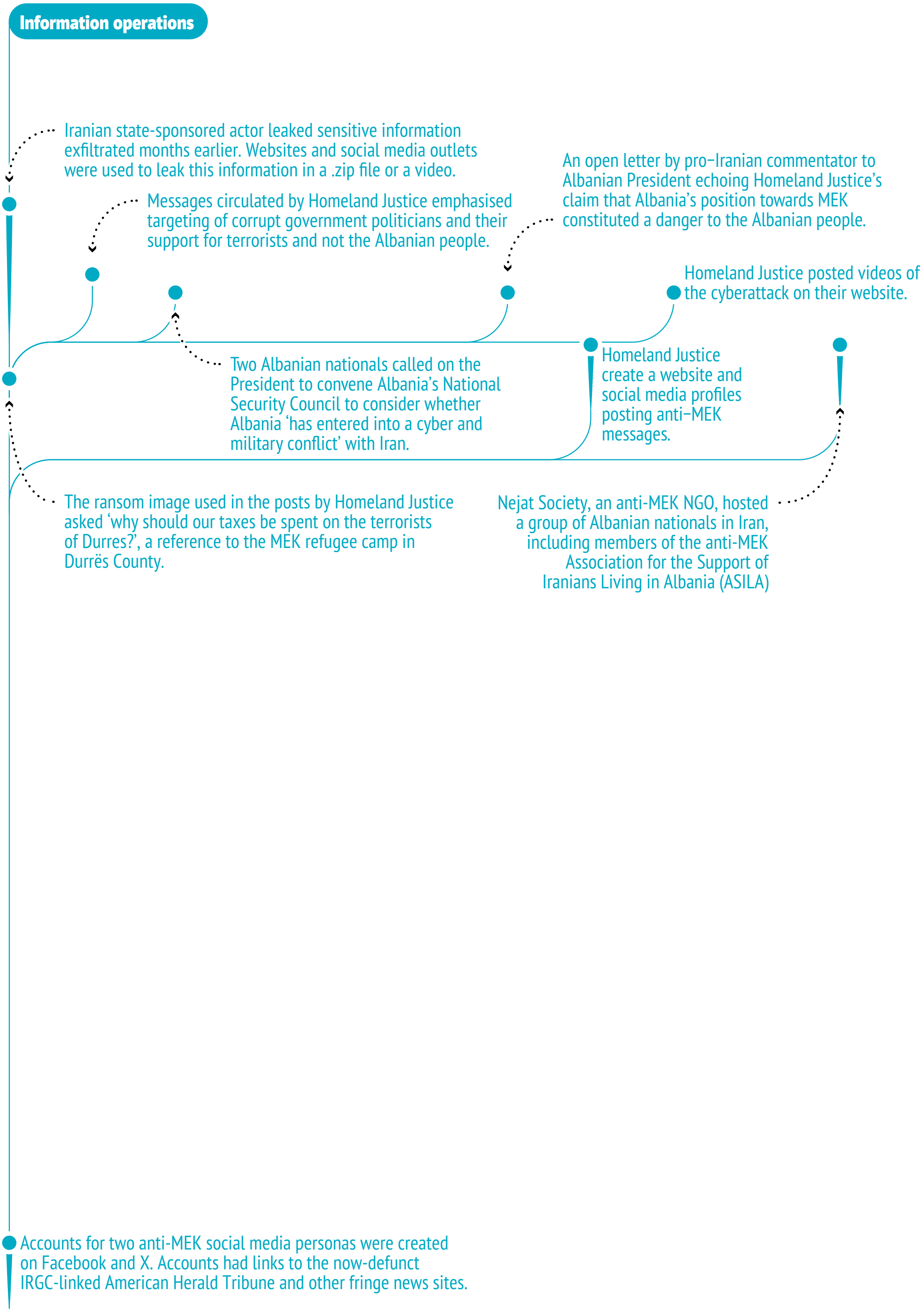
Ahead of the MEK-sponsored Free Iran World Summit, Homeland Justice launched a ransomware-style file encryptor with anti-MEK messages against the Albanian parliament, government agencies and institutions, telecom companies, and the national air carrier.

The attacks targeted the Albanian parliament, government agencies and institutions, telecom companies, and the national air carrier

Lateral movements, network reconnaissance, and credential harvesting from networks

Mail exfiltration

An Iranian group likely gained access to the network of the Albanian government.



Information operations

Iranian state-sponsored actor leaked sensitive information exfiltrated months earlier. Websites and social media outlets were used to leak this information in a .zip file or a video.

Messages circulated by Homeland Justice emphasised targeting of corrupt government politicians and their support for terrorists and not the Albanian people.

An open letter by pro–Iranian commentator to Albanian President echoing Homeland Justice's claim that Albania's position towards MEK constituted a danger to the Albanian people.

Homeland Justice posted videos of the cyberattack on their website.

Two Albanian nationals called on the President to convene Albania's National Security Council to consider whether Albania 'has entered into a cyber and military conflict' with Iran.

Homeland Justice create a website and social media profiles posting anti–MEK messages.

The ransom image used in the posts by Homeland Justice asked 'why should our taxes be spent on the terrorists of Durres?', a reference to the MEK refugee camp in Durrës County.

Nejat Society, an anti-MEK NGO, hosted a group of Albanian nationals in Iran, including members of the anti-MEK Association for the Support of Iranians Living in Albania (ASILA)

Accounts for two anti-MEK social media personas were created on Facebook and X. Accounts had links to the now-defunct IRGC-linked American Herald Tribune and other fringe news sites.

against this type of infrastructure have been reported[6].

> **Category 4. Information operations against critical infrastructure:** Their aim is to shape the narrative and public messaging around critical infrastructure. China has actively engaged in the media landscape in Italy following the signature of the Memorandum of Understanding in support of the Belt and Road Initiative (BRI) in March 2019[7]. China has also targeted the rare earth mining sector with information operations exploiting environmental concerns surrounding US, Canadian and Australian mining projects in order to bolster the competitive advantage of Chinese companies[8]. At the same time, the United States have engaged in anti-Chinese influence operations in Europe to limit the presence of Huawei and ZTE technology in 5G networks across the continent.

When it comes to incidents targeting critical infrastructure, a combination of several factors needs to be considered to better understand the dynamic nature of interference: the *motivation* behind the attack, the *nature* of the perpetrator, and the *importance* of the target. Therefore, designating a cyber incident as interference is a political act, which takes the act out of the hands of the incident response technical community and experts and places it in the hands of politicians and diplomats. A distributed denial of service (DDoS) attack against the Greek gas transmission system operator (DESFA) conducted as part of the Ragnar Locker ransomware operation[9] or cyberattacks against oil terminal-operator companies in Germany, Belgium and the Netherlands that caused distribution difficulties can be characterised as an incident, but not necessarily as interference. On the other hand, cyberattacks against German wind turbine manufacturers Nordex and Enercon and wind farm maintenance company Deutsche Windtechnik carried out after the Russian invasion of Ukraine might be considered as foreign interference aimed at influencing the foreign policy of Germany, if the link to the state institutions is confirmed[10]. In other words, the differences between the political and technical criteria are critical to assessing the effects of an incident.

Designating a cyber incident as interference is a political act.

THE EFFECTS

The disruptive effects of cyberattacks on different critical infrastructure sectors make the latter attractive targets for influence operations. Their immediate and visible impact on the larger population often leaves governments vulnerable to foreign influence, as they may be pressured to prevent or respond to such attacks to avoid public backlash. They are also more difficult to counter given that the control systems often cannot be simply switched off due to the potential cascading effects. This is particularly the case in the energy and telecommunications sectors, where there are large

[6] This type of interference is not discussed in this chapter. See: Recorded Future, *The escalating global risk environment for submarine cables*, Threat Analysis, 27 June 2023 (https://go.recordedfuture.com/hubfs/reports/ta-2023-0627.pdf).

[7] Ghiretti, F. and Mariani, L., 'One belt one voice: Chinese media in Italy', *IAI Papers* No 21, October 20211. (https://www.iai.it/sites/default/files/iaip2143.pdf).

[8] Jamalzadeh, S., Barker, K., Gonzalez, A. and Radhakrishan S., 'Protecting infrastructure performance from disinformation attacks', *Nature Scientific Reports*, 26 July 2022 (https://www.nature.com/articles/s41598-022-16832-w).

[9] Toulas, B., 'Greek natural gas operator suffers ransomware-related data breach', Bleeping Computer, 22 August 2022 (https://www.bleepingcomputer.com/news/security/greek-natural-gas-operator-suffers-ransomware-related-data-breach/).

[10] Mellor, S., 'Germany is trying to transition away from Russian fuel and hackers are now hitting German wind energy companies', *Fortune*, 25 April 2022 (https://fortune.com/2022/04/25/germany-trying-to-transition-away-from-russian-fuel-and-hackers-are-now-hitting-german-wind-energy-companies/).

numbers of dependent entities[11]. For instance, the ransomware attack against the US Colonial Pipeline – a large fuel pipeline that delivers 45 % of fuel supplies to the US East Coast – caused gasoline shortages, shutting down services and causing panic among consumers that led to a surge in gas costs.

The scale of attacks against critical infrastructure and the impossible task of responding to all of them have triggered the policy debate about the need to better define the risk management approach in preparing and responding to such attacks. The most important aspect has been defining the threshold above which incidents are considered significant and linking them to adequate policy responses (see next section). Some of the sectors already respond to this challenge by adopting a common incident classification scale[12] with clearly prescribed parameters and thresholds. Critical infrastructure in each sector is comprised of an ecosystem of different actors, each with their own responsibilities and potential impact on the proper functioning of the overall infrastructure. While in the Nordic/Baltic area the energy markets are well-connected, including dependencies with non-EU Member States such as Norway, the cross-border dependencies of the Netherlands are more limited. Interdependencies within the energy supply chain constitute a special category and cyberattacks against them may affect several states. For instance, the 2022 ransomware attacks against three companies that are part of the supply chain for petroleum products – SEA-Tank, Oiltanking, and Evos – affected the functioning of terminals and distribution of goods in several European and African countries.

The most important aspect has been defining the threshold above which incidents are considered significant.

Finally, not all cyber operations against critical infrastructure have the same impact. Their effects often depend on the country-specific context. A ransomware attack on an energy supplier in Europe, for instance, may differ depending on the energy market structure, preparedness of the country concerned and potential cascading effects across Europe. At the same time, different attack techniques produce different effects: a DDoS attack that disrupts the functioning of infrastructure is not the same as a cyber-espionage operation or destruction of networks. Given the scale and complexity of cyberattacks and limited resources to protect such systems, states rely on risk-based assessments to identify a specific category of public services and their operators – such as 'critical infrastructure', 'operators of critical services' or 'critical entities'. These designate industries or stakeholders where the impact of cyberattacks would be particularly damaging[13]. The inclusion of the trust service providers, top-level domain name registries, and Domain Name System (DNS) services regardless of the size is significant given the key role they play in information flows and spread of disinformation. Malicious practices such as DNS abuse and hijacking or spoofing of the top-level domains (e.g.,.com, .org, .eu, .fr) are particularly important in the context of information manipulation as they potentially offer control over the

[11] NIS Cooperation Group, *Sectorial implementation of the NIS Directive in the Energy sector*, Report – CG Publication 03/2019, October 2019 (https://digital-strategy.ec.europa.eu/en/library/eu-wide-cybersecurity-legislation-report-implementation-eu-rules-energy-sector).

[12] European Network of Transmission System Operators for Electricity (ENTSO-E), Incidents Classification Scale Methodology, 4 December 2019 (https://eepublicdownloads.entsoe.eu/clean-documents/SOC%20documents/Incident_Classification_Scale/200629_Incident_Classification_Scale_Methodology_revised_and_in_use_as_of_2020.pdf).

[13] See for instance the Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L__.2016.194.01.0001.01.ENG&toc=OJ:L:2016:194:TOC).

online information environment of a company, institution or even of the whole country[14].

The potential effect of a cyberattack against such infrastructures or entities is often used as a criterion for designation. EU law establishes clear criteria for the identification of the operators of essential services considering the effect that a disruption of their services would have on public safety, security or health, as well as potential cross-border consequences. It also includes provisions concerning a specific category of 'significant incidents' if they cause or have the potential to cause substantial disruptions or financial losses for the entity concerned or cause considerable material or non-material losses to other natural or legal persons[15].

THE RESPONSE

Recognising the multifaceted nature of the challenge of foreign interference, policy responses related to cyber operations targeting critical infrastructure have focused on addressing key issues for the cyber-FIMI nexus, in particular through strengthening resilience, deterring perpetrators, and consolidating international cooperation and partnerships.

Strengthening resilience is pursued primarily through development of regulatory frameworks and institutions. The cyber domain has been regulated from the very beginning through the system of standards and protocols that ensured interoperability and proper functioning of the internet infrastructure. To that end standard setting and technical organisations like the Internet Engineering Task Force (IETF), International Telecommunication Union (ITU) or the International Organization for Standardization (ISO) have developed specific frameworks, standards and product certification schemes. To further enhance preparedness and response capabilities, governments have moved to adopt specific legislation imposing concrete obligations on different actors within the cyber ecosystem. The NIS 2 Directive, for instance, sets the baseline for cybersecurity risk management measures and reporting obligations across all sectors that are covered by the directive, such as energy, transport, health and digital services that are critical for the free flow of information. The increasing number of sector-specific recommendations underscores the need for a comprehensive approach to improve the resilience of critical infrastructure. This includes building a robust institutional ecosystem spanning technical, operational and policy levels. Examples of such initiatives are Computer Security Incident Response Teams (CSIRTs)[16], Product Security Incident Response Teams (PSIRTs), and sector-specific Information Sharing and Analysis Centres[17].

Strengthening resilience is pursued primarily through development of regulatory frameworks and institutions.

Developing a deterrence posture and tools to dissuade the attackers from engaging in a malicious or illegal activity is another approach. In addition to dealing with potentially catastrophic low probability–high impact incidents, states have increasingly focused on strategies to address high probability–low impact attacks. Even though not significant if taken individually, this type of attack may

[14] Greenberg, A., 'Cyberspies hijacked the internet domains of entire countries', *Wired*, 17 April 2019 (https://www.wired.com/story/sea-turtle-dns-hijacking/).

[15] Directive (EU) 2022/2555 on measures for a high common level of cybersecurity across the Union (NIS 2 Directive), *Official Journal of the EU*, L 333, 27 December 2022.

[16] EU Agency for Cybersecurity, *PSIRT Expertise and Capabilities Development*, ENISA report, 3 June 20221 (https://www.enisa.europa.eu/news/enisa-news/new-light-shed-on-capabilities-in-energy-healthcare).

[17] See the European Energy Information Sharing & Analysis Centre: https://www.ee-isac.eu/

generate a significant impact over a longer period ('death by a thousand cuts'). With this in mind, states have developed cyber postures to increase the costs for potential attackers and make them face consequences. For instance, the United States has deployed hunt-forward teams to countries like Croatia or Albania. In the EU context, the cyber sanctions regime under the umbrella of the Cyber Diplomacy Toolbox is seen as the primary deterrence mechanism. While such tools could be considered in the future under the FIMI Toolbox, the EU has already sanctioned several entities in connection with disinformation operations undermining or threatening the territorial integrity, sovereignty and independence of Ukraine[18]. In a similar vein, the attacks on digital service providers in the EU could be potentially governed by the EU's cyber sanctions regime.

In addition, deterrence is also created through the criminal justice approach which is also an important element of the policy response, especially regarding the cross-border access to electronic evidence. The issues have been addressed through legislation by the United States Clarifying Lawful Overseas Use of Data Act (CLOUD Act) to speed up access to electronic information held by US-based global providers[19] and the EU's rules to make it easier and faster for law enforcement and judicial authorities to obtain the electronic evidence[20]. The international discussion about rules for cross-border access to evidence are laid down in the Second Additional Protocol to the Council of Europe Budapest Convention on Cybercrime adopted in November 2021[21]. Cooperation between law enforcement, intelligence agencies and the private sector has allowed for unusual responses such as botnet takedowns (e.g. Qakbot in 2023[22]) or cyber operations to disable 'troll factories' such as the Internet Research Agency in Saint Petersburg. This example clearly demonstrates that in the absence of a credible deterrence approach in the FIMI context, the tools and instruments available in the cyber context might provide an important (temporary) reinforcement.

Finally, given the global scale of the problem, it has become critical to agree the rules of the road in cyberspace. Since 1998, the UN has been engaged in the finetuning of the framework for responsible state behaviour in cyberspace that is grounded in existing international law and a catalogue of voluntary and non-binding norms, rules and principles regarding what states should or should not do in cyberspace. Recognising the potentially devastating consequences of malicious activities targeting critical infrastructure supporting essential services, governments endorsed a norm that forbids states to 'conduct or knowingly support ICT activity contrary to their obligations under international law that intentionally damages critical infrastructure or otherwise impairs the use and operation of critical infrastructure to provide services to the public. Furthermore, states should continue strengthening measures to protect all critical infrastructure from ICT threats and increase exchanges on best practices regarding critical infrastructure protection'[23]. While

[18] See: Council Implementing Regulation (EU) 2023/1563 concerning restrictive measures in respect of actions undermining or threatening the territorial integrity, sovereignty and independence of Ukraine, *Official Journal of the EU*, L 190l, 28 July 2023; Council Decision (CFSP) concerning restrictive measures in respect of actions undermining or threatening the territorial integrity, sovereignty and independence of Ukraine, *Official Journal of the EU*, L 190l, 28 July 2023.

[19] US Department of Justice, CLOUD Act, 21 March 2018 (https://www.justice.gov/dag/cloudact).

[20] European Commission, *Regulation on European Production Orders and European Preservation Orders for electronic evidence in criminal proceedings and for the execution of custodial sentences following criminal proceedings*, 28 July 2023 (https://commission.europa.eu/law/cross-border-cases/judicial-cooperation/types-judicial-cooperation/e-evidence-cross-border-access-electronic-evidence_en).

[21] Council of the European Union, Second Additional Protocol to the Convention on enhanced co-operation and the disclosure of electronic evidence, 17 November 2021.

[22] US Department of Justice, 'Qakbot malware disrupted in international cyber takedown', Press release, 29 August 2023 (https://www.justice.gov/usao-cdca/pr/qakbot-malware-disrupted-international-cyber-takedown).

[23] United Nations, *Report of the Group of Governmental Experts on advancing responsible state behaviour in cyberspace in the context of international security*, 14 July 2021. (https://front.un-arm.org/wp-content/uploads/2021/08/A_76_135-2104030E-1.pdf).

the discussion about the importance of critical infrastructure protection has proliferated to other international organisations and regional organisations, the international normative framework to address the cyber-FIMI nexus is still underdeveloped. Nonetheless, political debates are gradually moving in this direction. The G7 members reaffirmed their resolve to strengthen coordinated cyber defences and improve shared awareness of cyber threats and expressed their concern about the amplification of Russia's disinformation campaign targeting Ukraine [24]. The implementation of the UN Pact for the Future and the Global Digital Compact will provide an opportunity for addressing the existing gaps in a comprehensive way.

THE IMPLICATIONS

Measuring the impact of interference in the cyber and information environment is difficult, although both policy fields have attempted to generate tools to support adequate policy responses [25]. It is a positive sign that institutions responsible for cybersecurity and strategic communication are increasingly acknowledging the linkages between these two fields. The US Cybersecurity and Infrastructure Security Agency has issued guidance on better preparing critical infrastructure against influence operations. The European Union Agency for Cybersecurity (ENISA) has concluded a Memorandum of Understanding with the EEAS to strengthen cooperation against FIMI.

An effective cyber incident communication strategy reduces the space for information manipulation.

Moving forward, to address risks and threats along the cyber-information continuum, critical infrastructure protection strategies should clearly define their critical information environment and subsequently design adequate critical information environment protection strategies.

The critical information environment comprises the ecosystem of information channels and tools within which a critical infrastructure entity operates. Such an environment encompasses different groups of stakeholders that the entity interacts with, including contractors, regulators, media, or end users. Understanding how manipulation of operations might impact their activities and interaction with the entity providing critical infrastructure is critical. Once identified, the risk assessment for critical infrastructure sectors and entities should encompass adequate risk mitigation strategies and clearly defined communication plans.

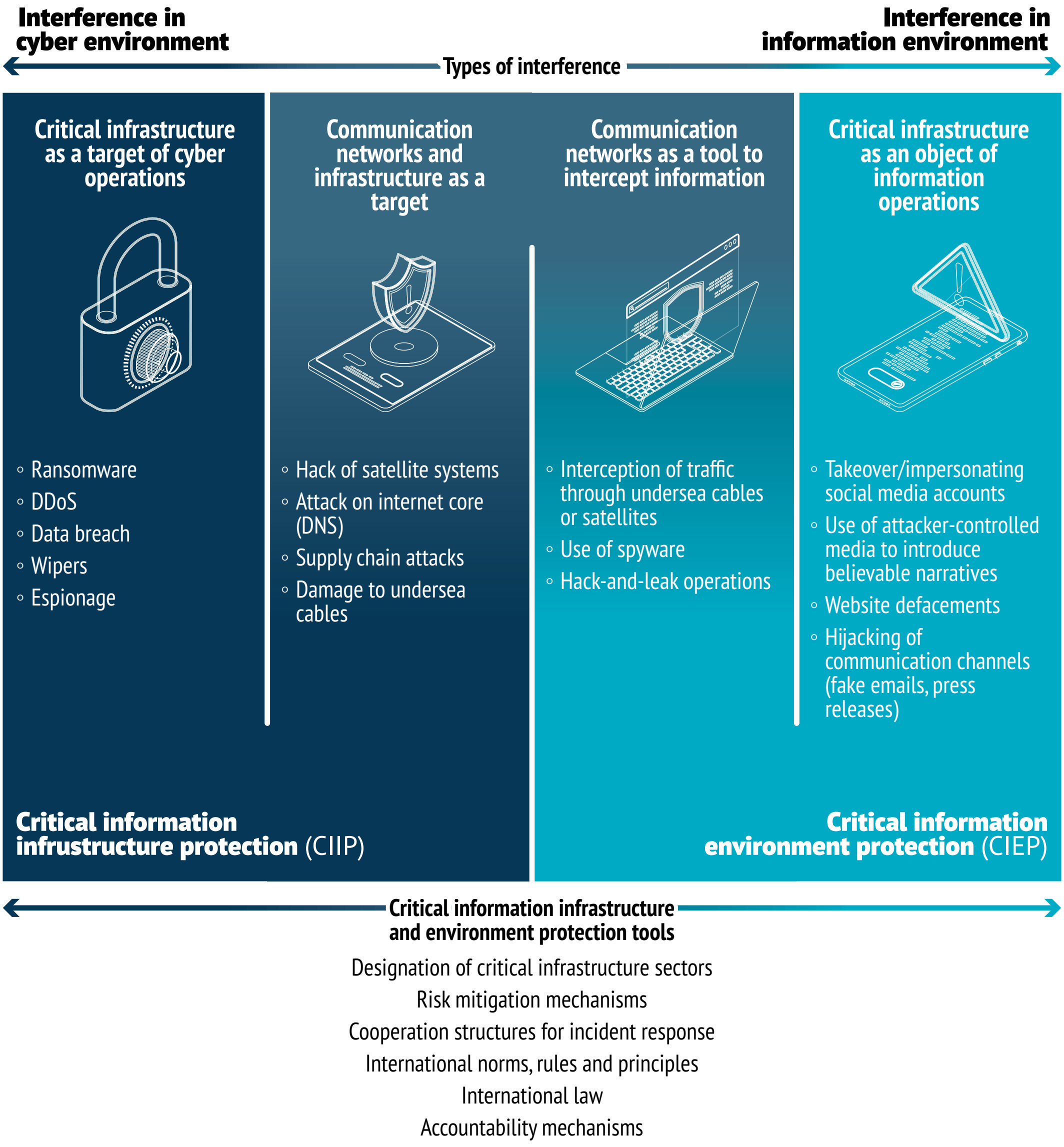
An effective cyber incident communication strategy reduces the space for information manipulation. Part of this approach is development and operationalisation of a common incident classification system that would include both cyber and information operation components, learning about the sources of information used by customers and stakeholders to understand potential threat vectors, as well as mapping of communication channels with stakeholders. For instance, ENISA and the EEAS have developed a dedicated analytical framework with the aim of analysing both FIMI and cybersecurity aspects of disinformation

[24] G7 Leaders' Statement - Brussels, 24 March 2022 (https://www.consilium.europa.eu/en/press/press-releases/2022/03/24/g7-leaders-statement-brussels-24-march-2022/).

[25] European External Action Service, *First EEAS report on foreign information manipulation and interference threats. Towards a framework for networked defence*, February 2023 (https://euvsdisinfo.eu/uploads/2023/02/EEAS-ThreatReport-February2023-02.pdf).

Critical infrastructure at the bridge between the cyber and information environments

Implications and solutions



Interference in cyber environment

Interference in information environment

Types of interference

Critical infrastructure as a target of cyber operations

- Ransomware
- DDoS
- Data breach
- Wipers
- Espionage

Communication networks and infrastructure as a target

- Hack of satellite systems
- Attack on internet core (DNS)
- Supply chain attacks
- Damage to undersea cables

Communication networks as a tool to intercept information

- Interception of traffic through undersea cables or satellites
- Use of spyware
- Hack-and-leak operations

Critical infrastructure as an object of information operations

- Takeover/impersonating social media accounts
- Use of attacker-controlled media to introduce believable narratives
- Website defacements
- Hijacking of communication channels (fake emails, press releases)

Critical information infrustructure protection (CIIP)

Critical information environment protection (CIEP)

Critical information infrastructure and environment protection tools

Designation of critical infrastructure sectors

Risk mitigation mechanisms

Cooperation structures for incident response

International norms, rules and principles

International law

Accountability mechanisms

that can be used also in the context of attacks against critical infrastructure[26].

In implementing this approach, exploring connections between different policy instruments and tools developed for critical infrastructure and information protection is important. For instance, the EU's NIS 2 Directive provides concrete guidance and procedures for mitigating risks to network infrastructure in a similar

[26] EU Agency for Cybersecurity, *Foreign information manipulation and interference (FIMI) and cybersecurity – threat landscape*, 8 December 2022 (https://www.enisa.europa.eu/publications/foreign-information-manipulation-interference-fimi-and-cybersecurity-threat-landscape).

way that the Digital Services Act does in the case of online platforms. However, the links between these two approaches – their complementarities and divergences – are rarely addressed. The NIS 2 Directive also contains specific provisions concerning digital infrastructure – such as DNS, cloud computing service providers or content delivery network providers – that are essential for the integrity of the internet and its continuous and stable operation. Cyberattacks on digital infrastructure undermine not only critical infrastructure but also information flows and transmission that rely on that digital infrastructure. By the same token, the Digital Services Act (DSA) contains specific provisions related to content moderation and other obligations of social media platforms that are particularly relevant for effective cyber incident management. Policy coherence is also relevant at the global level where different mechanisms are discussed. For instance, the principles and actions proposed in the UN Global Principles for Information Integrity[27] and the work undertaken by the General Assembly to strengthen the global framework for responsible state behaviour in cyberspace need to be aligned when it comes to protection of the critical information environment.

Finally, creating a comprehensive approach across the cyber-information continuum calls for a clear definition of roles performed by different stakeholder groups, their resources, and potential capability gaps when dealing with incidents combining both cyber and information environment elements. Whether they are the private sector, government agencies, civil society organisations, or fact-checker networks, it is important to assess the capabilities, strengths and weaknesses of each group and establish mechanisms for capacity building across a broad spectrum of needs. Creating new information exchange channels or strengthening the existing ones between those different groups and clearly defining the procedures for information sharing and cooperation plays an important role. For instance, CERT-EU operates a Social Media Assurance Service (SMAS) to help the EU institutions detect takeovers of social media accounts or impersonation. The body also monitors information manipulation activities, including China's social media mining for intelligence on foreign academics and journalists.

At the same time, strengthening multistakeholder partnerships between cyber and information environment communities is critical. This could be achieved by gaining a better understanding of their shared challenges (e.g., being subject to cyberattacks or disinformation campaigns) and respective resources at their disposal, especially for emergency management for the information environment[28]. Information exchange could be also improved by establishing an Information Sharing and Analysis Center (ISAC) for FIMI[29] and introducing information manipulation as a dimension in the work of other sectoral ISACs and Security Operation Centres (SOCs). To strengthen the resilience of governments and critical infrastructure operators in the evolving information environment it is becoming increasingly important to invest in mapping the existing capacity gaps and developing adequate needs and maturity assessment models.

[27] United Nations, *United Nations Global Principles for Information Integrity: Recommendations for Multi-stakeholder Action*, June 2024 (https://www.un.org/sites/un2.un.org/files/un-global-principles-for-information-integrity-en.pdf).

[28] Adam I., Samantha Lai S., Nelson, A., Wanless A. and Yadav, K., 'Emergency management and information integrity: A framework for crisis response', CEIP Working Paper, 9 November 2023 (https://carnegieendowment.org/files/Adam_et_al_-_Emergency_Management.pdf).

[29] EU Agency for Cybersecurity, 2022.

CONCLUSION

UNMASKING FOREIGN INTERFERENCE AND BUILDING RESILIENCE

by
NAĎA KOVALČÍKOVÁ

The case studies and analyses presented in this *Chaillot Paper* demonstrate that the growing convergence of FIMI and cyber threats across various societal domains amplifies their negative effects. Moreover, it points out the rising interplay between disinformation campaigns, cyberattacks and economic espionage, political subversion and the (mis)use of increasingly more sophisticated AI technologies by hostile actors.

However, it is not enough to simply demonstrate the complexity of foreign interference and the confluence of cyber and information manipulation tactics. To counter these threats effectively, we need to identify patterns and areas of convergence and divergence. This will help us gain a clearer understanding of the interconnected dynamics between *incidents,* their *effects* and the broader *implications* outlined in the five cases studied in this *Chaillot Paper*. Only then can we chart the way ahead for the EU. This knowledge is crucial for developing effective countermeasures as well as preventive and defensive tools and strategies.

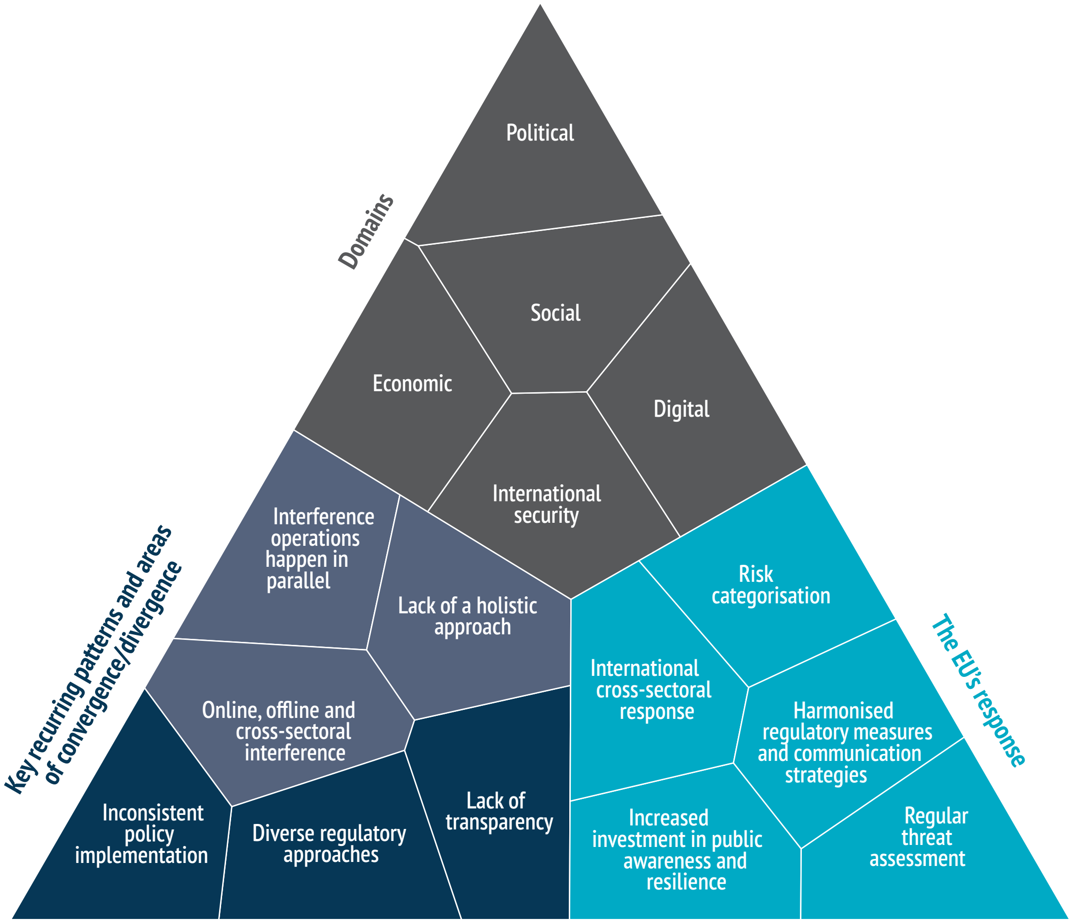
KEY RECURRING PATTERNS

Interference operations happen in parallel

Cyberespionage and intellectual property theft are frequently accompanied by disinformation to deflect attention from the attackers' actual targets and activities. Similarly, disinformation campaigns often coincide with cyberattacks and economic cyberespionage to exacerbate their effects, destabilise democracies and markets or create uncertainty within society. Moreover, at the national level, frequent changes in government or a failure to recognise and address disinformation, economic espionage, or exploitation of critical infrastructure vulnerabilities as national security and/or economic threats may hamper efforts to tackle them systematically and structurally.

Foreign interference untangled

Challenges and solutions


Domains
Political
Social
Economic
Digital
International security
Key recurring patterns and areas of convergence/divergence
Interference operations happen in parallel
Lack of a holistic approach
Online, offline and cross-sectoral interference
Inconsistent policy implementation
Diverse regulatory approaches
Lack of transparency
The EU's response
Risk categorisation
International cross-sectoral response
Harmonised regulatory measures and communication strategies
Increased investment in public awareness and resilience
Regular threat assessment

The lack of a holistic approach

Policy responses to campaigns of economic cyberespionage and political subversion often lack a holistic approach. The same applies to addressing misconduct within closed digital applications, where law enforcement agencies could ensure stronger detection and takedown capabilities. Several countries emphasise the need to integrate responses to both disinformation and cyber threats through national cybersecurity and technology strategies. For example, although transformative technologies such as AI hold great promise for increased efficiency, innovation and industrial progress, rapid advances in this field and the potential for misuse pose significant challenges. Across the globe, individuals and organisations are navigating these complexities, seeking to strike a balance between the opportunities and risks posed by AI [(1)]. EU regulatory measures or declarations following AI summits, such as those recently held in the UK (November 2023) or in Seoul (May 2024),

[(1)] '"Uncharted terrain": how officials, campaigners and fact-checkers tackle AI's influence on elections arounds the world', *Politico*, 21 May 2024 (https://www.politico.eu/article/uncharted-terrain-how-officials-campaigners-and-fact-checkers-tackle-ais-influence-on-elections-around-the-world/).

among other instruments, outline ways forward to address the challenges of AI technology and provide recommendations on how to embrace the opportunities it presents. However, the pace of technological advancement seems to outstrip our ability to address potential misuse. This applies to aggression in both the digital and physical realms, as demonstrated by the growing number of cyber intrusions or when protest activities are accompanied by an aggressive disinformation strategy. As a case in point, when the Chinese state media outlet *Global Times* called the president of the Czech Senate a 'political hooligan'(chapter 1), it set the tone for future bilateral exchanges. However, in this case, the attempt at intimidation also triggered a reaction from other MEPs and parliamentarians from other countries who rallied behind their Czech ally, transforming a bilateral issue into a broader European and transnational concern. In addition, the resort to personal attacks, as seen in this case, mirrored the tactics employed in the deepfake video case study (chapter 4).

Hostile actors deploy social media to further amplify fabricated content.

Online, offline and across sectors

Instability within a domestic political context can create fertile ground for promoting the Kremlin's agenda or exploiting existing anti-EU, anti-Western or pro-Russian sentiment among the population. This fuels further political radicalisation, creating a dangerous dynamic. The hacking of minds and machines does not, however, happen only via cyber-enabled operations. Attackers can also target specific people or sectors offline which are considered vital to a country's functioning and stability. This is well illustrated in the previously mentioned case in the Czech Republic (chapter 1). A 'diplomatic' letter was sent by the Chinese Embassy in Prague addressed to the Office of the Czech President. It aimed to dissuade the country from showing solidarity towards Taiwan by threatening economic and diplomatic repercussions in case of non-compliance with China's demands. This incident, although in the political domain, rippled across the diplomatic, economic, information and societal spheres, exposing the coercive tactics employed by China. Other incidents highlighted in the case studies presented in this volume have had similar cascading effects.

Identifying the linkages between cyber, FIMI and other tools of foreign interference is critical for comprehensive and coherent policy responses. However, despite ongoing efforts in both the disinformation and cyber domains to increase resilience (or, conversely, minimise the effects of the lack thereof), there remain notable differences in the nature and direction of policy and regulatory approaches.

AREAS OF CONVERGENCE AND DIVERGENCE

Lack of transparency

The operations of state and non-state actors have frequently been characterised by a high degree of concealment and opacity. They include *inter alia* cyber-enabled espionage, covert cyberattacks, political and economic coercion, and information manipulation. These operations have often involved outsourcing news manipulation to local partners or leveraging existing opaque networks within the targeted country or region, exploiting cultural or historical connections. Hostile actors deploy social media to further amplify fabricated content. The covert nature of such manipulative efforts makes technical attribution, and even public identification, of these actors extremely difficult. Foreign actors often exploit domestic divisions, instrumentalising local proxies. This not only obscures the identities of the actors behind these malicious operations, but also makes it difficult to identify the wider networks involved in specific contexts, further complicating efforts to counter them. In addition, interference with communication networks can have crippling repercussions across all societal sectors, including health, transportation, national administration, energy and trade. These interceptions create opportunities for cybersurveillance in various strategic domains, that can potentially be weaponised at critical junctures such as elections, trade negotiations, or high-level/high-stakes diplomatic visits.

Diverse regulatory approaches

While the EU has set out clear criteria for identifying essential service operators in certain sectors, particularly those with potential cross-border impact, other sectors and societal domains like politics, information or diplomacy often lack comprehensive criteria to recognise foreign interference and influence operations. Different sectors have tailored regulations, such as specific measures for AI-generated content under the AI Act versus broader cybersecurity measures under the NIS 2 directive. However, countering carefully orchestrated and persistently executed hostile strikes requires a strategic and calibrated response. While some deceptive operations are exposed and attributed relatively quickly, others may only be identified or disclosed years after the incident has taken place and evidence gathered by counterintelligence. In contrast, disinformation campaigns often aim for immediate impact and publicity using open sources and media amplification to destabilise societies as much, as frequently and as visibly as possible. While exposure of disinformation incidents helps build societal resilience to such interference, cyber espionage and IP theft presents a different challenge. Companies are often reluctant to disclose such breaches to shareholders or local authorities, fearing potential reputational damage and legal ramifications, and sometimes downplay the significance of such incidents (chapter 3). Unlike the readily apparent disruption caused by disinformation campaigns, it can take years for the economic consequences of cyberespionage to be discerned and fully appreciated.

Inconsistent policy implementation

Some policies prioritise technological solutions, e.g. AI-powered threat detection, while others focus on human-centric approaches such as education and training. Implementation of legal measures often lacks consistency and attribution may be a political rather than technical act. The case studies analysed in this volume suggest that the criminal justice approach has not been consistently applied across different sectors. Cybersecurity policies often prioritise protecting critical infrastructure, while disinformation policies tend to focus more on media literacy and content moderation. This lack of synergy creates a gap in effectively addressing the convergence of these threats.

Moreover, the spectrum of hostile activities extends beyond direct attacks intended to cause direct harm. Some, like diverting foreign direct investment and venture capital, aim rather to preserve the *status quo*, hindering a rival country's economic development and making it less attractive to investors. Such tactics contrast with overtly offensive cyber and FIMI operations that aim to create chaos, compromise infrastructure, steal data or fuel societal instability during election periods or at other critical junctures.

THE EU'S RESPONSE

Effective responses to foreign interference require a holistic and consistent strategy across different societal domains. The EU could address this challenge in five ways.

International cross-sectoral response

Countering interference requires robust international and cross-sectoral collaboration. Insidious tactics thrive on exploiting societal divisions, legal loopholes, political instability, historical grievances and existing geopolitical tensions. The countermeasures to address them may vary across countries and governments need to tailor them to national circumstances. This complexity can make timely and systematically coordinated responses a significant challenge.

Tailored and strategic partnering with global cybersecurity and disinformation watchdogs, and the expansion of such partnerships, can facilitate the sharing of good practices and contribute to coordinated responses to foreign threats. Enhanced multistakeholder cooperation is critical between counterintelligence services, law enforcement, and high-risk private sector industries, among others. This has worked effectively on previous occasions, as shown in chapter 5. Such a collaborative approach could be bolstered via increased investment in investigative journalism, local projects and public awareness campaigns combined with digital and media literacy programmes. These efforts, tailored to local contexts, would educate citizens about the tactics used in foreign interference and address their specific concerns. Lastly, to deter future violations, it is essential to hold social media companies financially liable for insufficient compliance with existing codes of practice or regulatory measures, and ensure their adequate collaboration with EU and national authorities.

An effective framework for responsible state and non-state behaviour in the digital and broader public sphere requires dedicated cross-sectoral investment from governments and the international community. An example is the partnership between the FBI and their UK and Five Eyes counterparts that tackles economic cyber-espionage and engages with the business community. Moreover, the case studies presented in this *Chaillot Paper,* and specifically the one describing cyber operations to disable 'troll factories' (chapter 5), highlight that while cyber tools can offer temporary solutions in the absence of FIMI deterrence or defence measures in certain contexts, law enforcement authorities should develop a more comprehensive approach.

Harmonised regulatory measures and communication strategies

All the chapters in this volume have underlined the need for the EU to build stronger defences against foreign authoritarian influence. This can be achieved through a comprehensive assessment of vulnerabilities and threats, at national, European and transnational levels, considering the global informational environment's susceptibility to interference. The criminal justice system can also play a role in deterring such activities. The EU should develop a cohesive strategy that addresses the entire cyber-information continuum to counter foreign interference. This includes defining roles and capabilities for various stakeholders, improving early-warning and information-sharing mechanisms, and ensuring policy coherence across different regulatory instruments.

Responses to cyberespionage primarily occur at the domestic or minilateral level, while the broader EU response to coercive tactics in the political sphere, e.g. those employed by China and outlined in the first case study in chapter 1, highlights the potential for a wider mobilisation at the level of the European Parliament. Both cyber and FIMI incidents, as well as other types of interference, require effective incident communication strategies. These strategies should establish clear roles, identify resource and capability gaps and leverage a combination of FIMI and cyber expertise to minimise the impact of interference. As the case study on critical infrastructure demonstrates, collaboration is key. Several government agencies have issued guidance or concluded agreements to strengthen measures and cooperation addressing both the cybersecurity and FIMI aspects.

Risk categorisation and management

To counter foreign interference effectively, the EU should ensure that policies addressing cybersecurity and disinformation are harmonised. This includes improving risk classification and management of AI-generated disinformation within the DSA and AI Act, as well as enhancing risk management strategies to proactively address and respond to rising threats. A more nuanced approach to risk categorisation is needed. For example, deepfakes are categorised as a 'limited risk' under the AI Act, which might not reflect the higher perceived risk that they pose in disinformation contexts. Similarly, addressing the high costs associated with cyber-enabled espionage, such as harming a sector's competitiveness or hardening business resilience, requires closer attention. Thus, defining a common classification scale for risks and incidents needs to take into consideration various criteria, thresholds, and potential cascading effects. For less apparent threats, counterintelligence services can play a crucial role in advancing criminal investigations and law enforcement actions, such as cyber sanctions or expelling diplomatic personnel.

A cross-sectoral effort involving all relevant stakeholders is crucial to assess the evolving hybrid threat landscape.

Increased investment in public awareness and resilience

The public exposure of foreign interference helps foster and alert citizens to the disruptive effects on democracy. The EU needs to prioritise clear and transparent public communication, explaining the economic and value-based reasoning behind political decisions, as seen in the chapter on political coercion in this volume. The EU could also consider further measures to strengthen deterrence through public exposure of foreign interference incidents, in collaboration with mainstream media, thereby raising the costs of interference, as highlighted in the chapter on Chinese interference in Czech politics. Identifying, understanding, and publicly exposing the connections between disinformation and cyber threats can lead to improved defence mechanisms, enhanced societal awareness and resilience building. In addition, 'prebunking' tools that warn citizens about manipulative tactics can reinforce collaboration at the transnational level.

As lack of transparency allows interference to thrive, enhanced sharing of information may further contribute to build stronger societal resilience. This is exemplified in the case of national cybersecurity centres which have become more forthcoming in sharing information in the public domain and may serve as a useful model. In particular, the EU needs to formulate a coherent and unified strategy to counter foreign interference across various domains, including the economic, media and political spheres. Moreover, investing in initiatives like the EU Cybersecurity Skills Academy strengthens public resilience against disinformation and cyber threats by creating a more knowledgeable and empowered citizenry.

Regular threat assessment

Hostile actors employ increasingly aggressive messaging. The EU should therefore ensure regular assessment of evolving threats, foster shared understanding and develop early warning tools. Moreover, to better deter malign influence operations in the long term and raise the costs of interference, the EU could systematically impose severe penalties or fines on companies that fail to comply with regulatory obligations, thus making non-compliance a costly endeavour.

Nevertheless, there is no single or simple solution to address the diverse tools of foreign interference. A collaborative cross-sectoral effort involving all relevant stakeholders is crucial to assess the evolving hybrid threat landscape[2], as policy and regulatory frameworks are updated. While existing cyber diplomacy, FIMI and hybrid threat toolboxes provide enhanced responses to these threats in specific areas, the Member States and the EU should consider further integration or even a dedicated 'counter-coercion toolbox', e.g. specifically designed to protect European politicians (as outlined in chapter 1).

By implementing these steps, the EU can bolster its defences against foreign interference and empower European societies to navigate the digital era with tangible measures, greater confidence and enhanced resilience.

[2] Giannopoulos, G., Smith, H. and Theocharidou, M., 'The landscape of hybrid threats: A conceptual model', European Commission and Hybrid CoE, November 2020 (https://euhybnet.eu/wp-content/uploads/2021/06/Conceptual-Framework-Hybrid-Threats-HCoE-JRC.pdf).

ABBREVIATIONS

5G	Fifth generation *(of wireless mobile telecommunications technology)*
AI	Artificial intelligence
APT	Advanced Persistent Threat
BRI	Belt and Road Initiative
CERT-EU	Computer Emergency Response Team for the EU institutions, bodies and agencies
CSIRT	Computer Security Incident Response Team
DDoS	Distributed denial of service
DNS	Domain Name System
DSA	Digital Services Act
EEAS	European External Action Service
ENISA	European Union Agency for Cybersecurity
FBI	Federal Bureau of Investigation
FIMI	Foreign information manipulation and interference
GANs	Generative Adversarial Networks
GDP	Gross domestic product
ICT	Information and Communication Technologies
IP	Intellectual Property
ISAC	Information-Sharing and Analysis Centre
IT	Information Technology
LEAs	Law Enforcement Agencies
MEK	Mujahedin-e-Khalq
MEP	Member of the European Parliament
MSS	Ministry of State Security
NATO	North Atlantic Treaty Organization
NGO	Non-Governmental Organisation
NIS	Network and Information Security
NISC	National Centre of Incident Readiness and Strategy for Cybersecurity
PRC	People’s Republic of China
R&D	Research and Development
TRIPS	Trade-Related Aspects of Intellectual Property Rights
TTPs	Tactics, techniques and procedures
UK	United Kingdom
UN	United Nations
USD	United States dollars
WTO	World Trade Organization

NOTES ON THE CONTRIBUTORS

Rumena Filipova is Chairperson of the Institute for Global Analytics in Bulgaria. Her main research interests focus on the politics and international relations of Central and Eastern Europe, with particular reference to questions of media and disinformation, identity and the authoritarian influence exercised by Russia and China in the region. Rumena Filipova holds a DPhil and MPhil in International Relations from the University of Oxford, a BA in Politics, Psychology and Sociology from the University of Cambridge and has gained extensive experience in the think tank sector.

Bart Hogeveen is the Acting Director for Cyber, Technology & Security at the Australian Strategic Policy Institute (ASPI). He oversees research, strategic engagements, dialogue and capacity-building initiatives on issues at the nexus of national, cyber and economic security. His contribution to this volume is informed by insights from a multi-year project that looked at the issue of state-sponsored economic cyberespionage and the effects on emerging economies. His research also focuses on security and military cyber capabilities in the Indo-Pacific, practices of cyberwarfare and espionage and initiatives that enhance international rules, norms and standards for responsible state behaviour in cyberspace.

Ivana Karásková is the China Team Lead at the Association for International Affairs (AMO). She is also the founder and coordinator of MapInfluenCE, which analyses China and Russia's influence in Central Europe, and of the China Observers in Central and Eastern Europe (CHOICE) network. Since 2020, she has been a European China Policy Fellow at the Mercator Institute for China Studies (MERICS) in Berlin. In 2023-2024, she worked as a Special Advisor to the Vice-President of the European Commission and Commissioner for Values and Transparency Věra Jourová. She holds a Ph.D in International Relations from Charles University in Prague and other university degrees in Journalism and Mass Communication, European Studies and International Relations.

Nad'a Kovalčíková is the Senior Analyst in charge of the transnational security portfolio at the EUISS and Project Director of the EU-funded initiative 'Countering Foreign Interference'. She is a member of the ESPAS foresight Steering Group; an expert collaborator for Minsait's Ideas for Democracy; and a member of the Steering Committee of Women in International Security. She previously worked at the German Marshall Fund of the United States, NATO, the European Parliament, the French and Canadian embassies, and on several NGO and think tank projects across Europe and the Atlantic. She holds a PhD in International Economic Relations.

Patryk Pawlak is a part-time professor at the Robert Schuman Centre for Advanced Studies at the European University Institute (Florence) and a visiting scholar at Carnegie Europe. He is project director of the Global Initiative on the Future of the Internet funded by the European Union. In December 2023, he was appointed to the UN Advisory Board on Disarmament Matters and the UNIDIR Board of Trustees. Before joining EUI, he worked for the EUISS where he headed the Brussels office and led the Institute's cyber and digital projects. In this capacity, he was also the project director of the EU Cyber Diplomacy Initiative and coordinated the European Cyber Diplomacy Dialogue between senior government officials and scholars. His current research focuses on the impact of technology on foreign and security policy and the EU's cyber and digital diplomacy.

Andrea Salvi is the Senior Analyst responsible for cyber and digital issues at the EUISS. He is also Project Director of the EU Cyber Diplomacy Initiative. Previously, he served as lead project officer for the DRMKC Risk Data Hub at the Joint Research Centre of the European Commission and as Human Rights Statistics Consultant for the UN Mission to South Sudan. He holds a PhD in Political Science from Trinity College Dublin and master's degrees in International Relations, EU Law and Government, and Cybersecurity. He has published in journals such as *International Studies Quarterly* and *Computers & Security*, and has extensive lecturing experience in quantitative methods, political risk, and conflict studies.

This *Chaillot Paper* delves into the phenomenon of foreign interference and the risk it poses to democratic societies. It explores the interplay between information manipulation and disruptive cyber operations, revealing their role as complementary components within a broader strategy. Dedicated chapters examine how interference manifests across various sectors, including social, political, economic, digital and security domains, describing existing tools and evolving policy responses. Each case study follows a clear structure, presenting an incident, its effects and the implemented responses.

The volume concludes by identifying convergences and divergences across the cases studied, and highlights foreign interference as a critical and growing threat to global security. It offers targeted recommendations on how the EU can significantly bolster its defences and resilience against this threat.

euiss European Union Institute for Security Studies

Publications Office of the European Union

ISBN 978-92-9462-351-5
CATALOGUE NUMBER QN-AA-24-004-EN-C