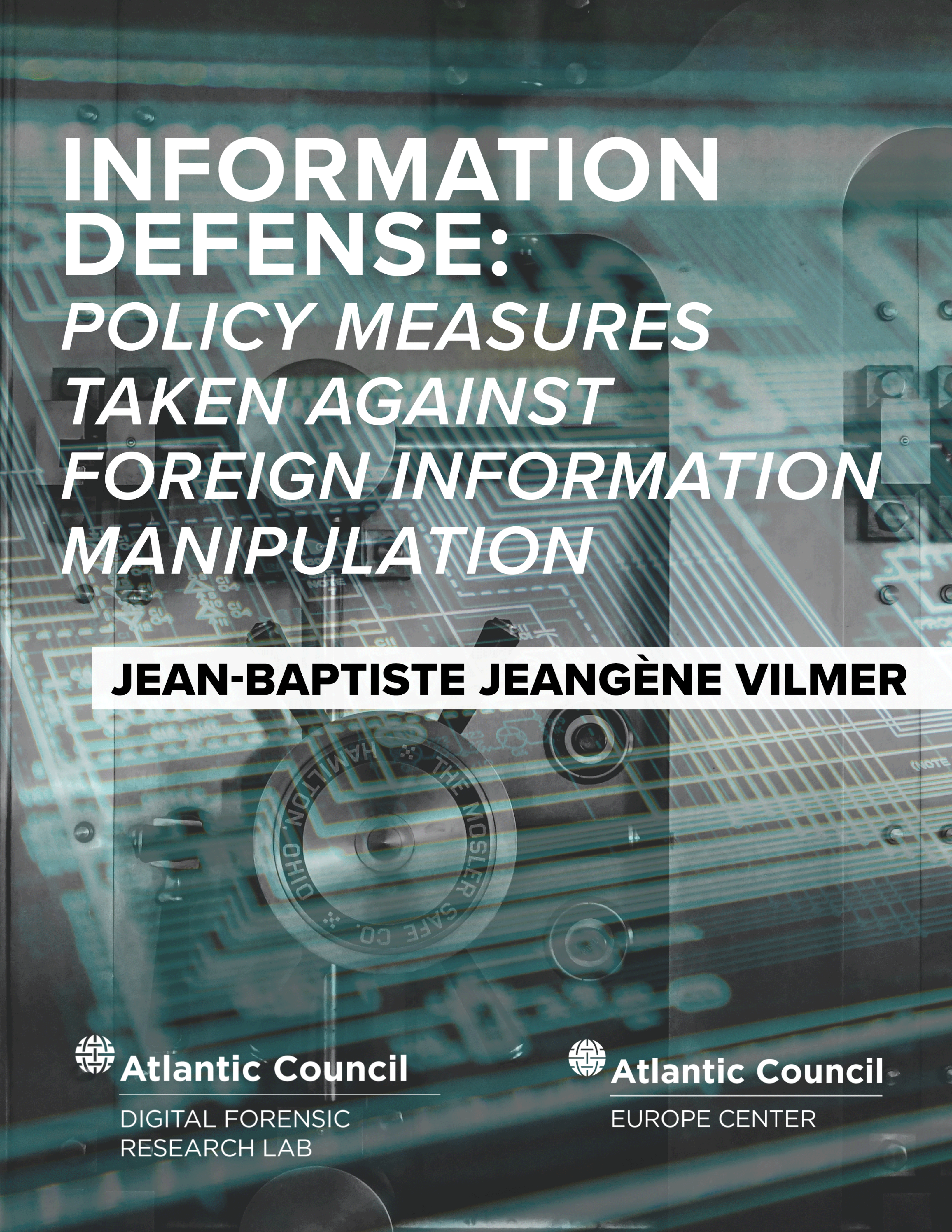

INFORMATION DEFENSE: *POLICY MEASURES TAKEN AGAINST FOREIGN INFORMATION MANIPULATION*

JEAN-BAPTISTE JEANGÈNE VILMER

Atlantic Council
DIGITAL FORENSIC
RESEARCH LAB

Atlantic Council
EUROPE CENTER

Atlantic Council

The Digital Forensic Research Lab (DFRLab) is a start-up incubated at the Atlantic Council, which leads the study of information to promote transparency and accountability online and around the world. DFRLab is known for defining disinformation, documenting human rights abuses, and focusing on facts; it has forged a role as a global leader in bridging the collective responsibility for information among government, media, and the private sector. It leads a large body of work on technology and democracy policy.

The Europe Center conducts research and uses real-time commentary and analysis to guide the actions and strategy of key transatlantic decisionmakers on the issues that will shape the future of the transatlantic relationship and convenes US and European leaders through public events and workshops to promote dialogue and to bolster the transatlantic partnership.

Atlantic Council
1030 15th Street NW, 12th Floor
Washington, DC 20005

For more information, please visit www.AtlanticCouncil.org.

July 2021

Research and Writing
Jean-Baptiste Jeangène Vilmer

Editors
Graham Brookie
Andy Carvin
Zarine Kharazian
Rose Jackson
Iain Robertson

Report and cover design
Eric Baker

COVER PHOTO (BACKGROUND):

Jason Dent, No Title, February 19, 2020, https://unsplash.com/photos/3wPJxh-piRw.

Adi Goldstein, No Title, February 21, 2019, https://unsplash.com/photos/EUsVwEOsblE.

Information Defense:

Policy Measures Taken Against Foreign Information Manipulation

Jean-Baptiste Jeangène Vilmer

Table of Contents

> *Most publications on foreign information manipulation focus on the offense, i.e., on the threat. They expose operations or analyze the strategies and tactics of the attackers. Understanding the threat is indeed a priority, but one should not lose sight of its raison d'être: to prevent and/or counter the attack. For the liberal democracies that are the most vulnerable targets of such operations, the main question is how to respond. That is why this report focuses on the defense. It is not intended to be comprehensive: it cannot cover all responses from all actors in all regions. It therefore focuses on information defense mostly from a governmental perspective, even though private sector efforts will also be mentioned, and mostly from a transatlantic perspective, even though a couple of other examples will also be mentioned. With these limits, this report offers a broad yet concise overview of policy measures taken against foreign information manipulation. Who is doing what?*[1]

Introduction: The three stages of awareness

Information manipulation is understood here as a coordinated campaign disseminating false or consciously distorted information for hostile political purposes.[2] It is "foreign" when it is orchestrated by foreign actors or their representatives, in which case it constitutes deliberate interference. In practice, it is often difficult to disentangle the domestic from the foreign, as well as information manipulation from broader influence efforts: information is only one of many tools in influence campaigns, often combined with other (economic, diplomatic, psychological, etc.) means. Aggressors choose their toolset based on effectiveness and the relative vulnerabilities of the target. For that reason, several policy measures taken against foreign information manipulation presented in this report belong to, or are linked to, broader efforts to counter foreign influence and so-called "hybrid threats."[3]

Aside from a few central, northern, and eastern European countries, where information manipulation from the East did not actually stop when the Cold War ended, all the countries in the West that had implemented defensive measures against Soviet operations dismantled them during the 1990s.[4] They were therefore taken off-guard when, 20 years later, they once again needed to protect themselves from large-scale, state-led attacks. Awareness of this threat has been gradual – and continues to grow throughout the world – but seems to have quickened in pace during the 2010s, over three stages.

The first was the Ukrainian sequence of 2013-2014 with Euromaidan, the annexation of Crimea, and the war in the Donbas region of Ukraine. The Russian offensive in Ukraine has become a textbook case of a so-called "hybrid war," including an information war. It was therefore the first episode to heighten

[1] A previous and shorter version of this report was published in French as a book chapter: "Panorama des mesures prises contre les manipulations de l'information," in C. Marangé and M. Quessard (eds.), *Les guerres de l'information à l'ère numérique* (Paris: Presses universitaires de France, 2021), 365-388. The author thanks Alexandre Alaphilippe, Jakub Kalenský, Ben Nimmo, Graham Brookie, Iain Robertson, and the DFRLab team who provided insightful comments on the first draft of this updated and enriched English version.

[2] Jean-Baptiste Jeangène Vilmer, Alexandre Escorcia, Marine Guillaume, and Janaina Herrera, *Information Manipulation: A Challenge for Our Democracies*, Policy Planning Staff (CAPS), the Ministry of Europe and Foreign Affairs, and the Institute for Strategic Research (IRSEM) of the Ministry of the Armed Forces, August 2018, https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf.

[3] The "hybrid war" concept was introduced in: Lieutenant General James Mattis and Lieutenant Colonel Frank Hoffman, "Future Warfare: The Rise of Hybrid Wars," *Proceedings Magazine*, US Naval Institute, November 2005:131-11, https://www.usni.org/magazines/proceedings/2005/november/future-warfare-rise-hybrid-wars. It can be defined as "the mixture of coercive and subversive activity, conventional and unconventional methods (i.e. diplomatic, military, economic, technological), which can be used in a coordinated manner by state or non-state actors to achieve specific objectives while remaining below the threshold of formally declared warfare" (*Joint framework on countering hybrid threats*, European Commission, April 6, 2016, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016JC0018).

[4] See for example, in the case of the United States: Nicholas J. Cull, *The Decline and Fall of the United States Information Agency: American Public Diplomacy, 1989-2001* (Palgrave Macmillan, 2012).

awareness of state vulnerabilities in that regard. As Ukraine is also a specific case given its history with Russia and the fact that the targeted populations were mainly Russian-speakers, however, this risk particularly struck a chord with countries that shared one or the other of those traits, especially the Baltic nations. But, under pressure from a certain number of countries that were more concerned than the others and also because of the massive export of Kremlin's disinformation about the MH17 crash, this awareness spread to more countries, the North Atlantic Treaty Organization (NATO), and the European Union (EU). The first tangible reaction of the EU came in March 2015, when the European Council "stressed the need to challenge Russia's ongoing disinformation campaigns."[5]

The second stage, which made much of the world conscious of this threat, was Russian interference in the 2016 US presidential election. Whereas the previous actions may have convinced some large countries in Western Europe or North America that they were, in a certain sense, untouchable in terms of both Russia's capacities and ambitions, the US case demonstrated to the world that no one was safe, not even the leading world power. An immediate effect of that affair was to put all the other countries on their guard. Another incidence of electoral interference in a major power, the so-called "Macron Leaks" operation[6] in France the following year, confirmed this shared vulnerability and the need for better protection. Between 2016 and 2018, at least forty-three states "proposed or implemented regulations specifically designed to tackle different aspects of influence campaigns." The body of measures – organizational, legislative, and educational – has since grown exponentially.[7]

The third stage in growing awareness consisted of realizing the global nature of the threat since it is not just Russian or state-led. On the one hand, this has taken the form of a "pivot to Asia," putting the focus of attention on China's growing assertiveness and aggressive posture in information operations, generating a global concern that China is gradually adopting the so-called "Russian playbook" (which does not exist as such, as Camille François notes – "the Russian playbook is akin to a Russian salad: not very Russian, and with different ingredients every time"[8] – so this is mostly a manner of speaking).[9] Again, what has been a familiar threat in Taiwan, Hong Kong, and Singapore for decades started to expand gradually, first through Australia and New Zealand, then to Europe, North America, and the rest of the world. The COVID-19 pandemic made it clear in 2020, but the so-called "Wolf warrior diplomacy" had already been happening for a couple of years. Since 2018 approximately, there is a "China Turn" in disinformation studies, with an ever increasing number of reports and analysis on Chinese operations being produced all over the world, including at the Atlantic Council's Digital Forensic Research Lab (DFRLab).[10]

On the other hand, the same applies to awareness of the issue of non-state actors: it has existed since the early 2000s in regard to the specific case of Jihadist groups, first al-Qaeda and then Daesh (and triggered attempts at coordinated responses), but now extends to populist and nationalist movements and extremist groups of all kinds, as well as to companies that have made disinformation a business. It is not the intention of this report, though, to cover extensively domestic or non-state actor disinformation, as it is focused on state-owned foreign operations. However, the responses presented in this report could be potentially applied more broadly. In any case,

[5] | European Council, *Conclusions of the European Council meeting of 19-20 March 2015*, EUCO 11/15, March 20, 2015, https://www.consilium.europa.eu/media/21888/european-council-conclusions-19-20-march-2015-en.pdf.

[6] | Jean Baptiste Jeangène Vilmer, *The "Macron Leaks" Operation: A Post-Mortem*, IRSEM/Atlantic Council, June 2019, https://www.atlanticcouncil.org/in-depth-research-reports/report/the-macron-leaks-operation-a-post-mortem/.

[7] | Samantha Bradshaw, Lisa-Maria Neudert, and Philip N. Howard, *Government Responses to Malicious Use of Social Media*, NATO Strategic Communications Centre of Excellence, November 2018, https://stratcomcoe.org/pdfjs/?file=/cuploads/pfiles/web_nato_report_-__government_responses-1.pdf.

[8] | Camille François, "Moving Beyond Fears of the 'Russian Playbook,'" *Lawfare*, September 15, 2020, https://www.lawfareblog.com/moving-beyond-fears-russian-playbook.

[9] | Jean Baptiste Jeangène Vilmer and Paul Charon, "Russia as a hurricane, China as climate change: Different ways of information warfare," *War on the Rocks*, January 21, 2020, https://warontherocks.com/2020/01/russia-as-a-hurricane-china-as-climate-change-different-ways-of-information-warfare/.

[10] | Iain Robertson, *Countering Chinese Disinformation Reports*, DFRLab, December 17, 2020, https://www.atlanticcouncil.org/in-depth-research-reports/dfrlab-china-reports/.

there is now a shared conviction that the threat is global, multifaceted, and multi-scale. Only a few rare countries feel this threat does not concern them. From the Faroe Islands to Fiji via Mali, the Philippines and Brazil, awareness is now almost universal, to varying degrees, of course, with some feeling (rightly or wrongly) they are more or less at risk than others.

Against this backdrop, the question this report deals with is: What have the responses been? What measures have states and civil societies taken to defend themselves against information manipulation? This is the sub-field of "information defense", a term advocated by Ben Nimmo, a former Senior Fellow with the DFRLab.[11] Information defense is as much defense *of* as defense *by* information. Jakub Kalenský, a Senior Fellow at the DFRLab, distinguished four lines of defense[12]: (1) "documenting the threat" by producing more knowledge, public or not, of actors and tactics involved in information manipulation; (2) "raising awareness" by communicating (at least some of) this knowledge to large audiences in order to educate them; (3) "repairing the weaknesses that disinformers exploit" by developing media literacy at all level (not only for children) and addressing our societies' weaknesses, which are often divisions, because most of the time attackers do not create new problems, they just exploit and amplify existing societal tensions to further polarize our societies; and (4) "punishing the information aggressor" by imposing a cost, which can be done by sanctions and laws.

This interpretative framework likely covers all existing countermeasures, which fall into one or the other of such categories. However, this report will use a different typology, by actor: it will distinguish between measures taken by states, international cooperation, and civil society (understood here as the aggregate of all nongovernmental organizations and institutions, including the private sector and therefore the digital platforms). The reason is that, contrary to most publications in the field of information defense, this report is mostly descriptive, not normative. It says what the actors are doing in reality, not what they should be doing in an ideal world. While many other publications provide useful recommendations about what should be done, this report focuses on what is being done already.

It is not intended to be comprehensive: it cannot cover all responses from all actors in all regions. It therefore focuses on information defense mostly from a governmental perspective, even though the private sector efforts will also be mentioned, and mostly from a transatlantic perspective, even though a couple of other examples will also be mentioned. The fact that most of the states cited in this report are liberal democracies does not mean that their governments are always or necessarily cooperative. As the cases of the United States under the Trump Administration or Hungary under Viktor Orbán remind us, governments in democracies themselves may be not only uncooperative in countering foreign information manipulation but also participatory in these activities. As for the couple of other states mentioned that are not liberal democracies, their presence in this report does not mean that they are considered in a positive light, or as examples to follow, only that they are active in developing countermeasures, and that some of them developed frequent bilateral relations with the main liberal democracies on that specific issue.

Finally, this report does not assess the impact of each effort, as effectiveness is a complex issue that is the subject of another, complementary report.[13] Effectiveness is always context-based, i.e., in a given situation, at a certain time, and for a certain

[11] Ben Nimmo was then "Senior Fellow for Information Defense" with the DFRLab. In 2019-2021, he was Director of Investigations at Graphika. As of February 2021, he is a Global Threat Intel Lead at Facebook. See: Adam Satariano, "He Combs the Web for Russian Bots. That Makes Him a Target," *The New York Times*, February 9, 2020, https://www.nytimes.com/2020/02/09/technology/ben-nimmo-disinformation-russian-bots.html.

[12] Jakub Kalenský, "Russian Disinformation Attacks on Elections: Lessons from Europe," Testimony to the Foreign Affairs Subcomm. on Europe, Eurasia, Energy, and the Environment, US House of Representatives, July 16, 2019, https://www.congress.gov/116/meeting/house/109816/witnesses/HHRG-116-FA14-Wstate-KalenskJ-20190716.pdf; Thomas Bastianelli, "Disinformation and diverging narratives: an interview with Jakub Kalensky," *geopolitica.info*, May 7, 2020, https://www.geopolitica.info/disinformation-and-diverging-narratives-an-interview-with-jakub-kalensky/.

[13] Jean-Baptiste Jeangene Vilmer, *Effective State Practices Against Disinformation: Four country case studies*, Hybrid CoE Research Report 2, Helsinki: The European Centre of Excellence for Countering Hybrid Threats, July 2021, https://www.hybridcoe.fi/wp-content/uploads/2021/07/20210709_Hybrid_CoE_Research_Report_2_Effective_state_practices_against_disinformation_WEB.pdf.

actor. To paraphrase Robert Cox,[14] effectiveness "is always *for* someone and *for* some purpose." In that sense, because what works here may not work there, this brief and non-exhaustive overview of policy measures is not intended to list replicable recipes against foreign information manipulation. Moreover, there are important limitations of existing measures, that the last part of this report will point out. It is the hope of this report that knowing what is being used, including the limitations, could help designing feasible and realistic improvements.

Measures taken by states

States have implemented a certain number of measures to reactively retaliate or proactively deter influence operations ranging from organization design, i.e., determining the optimal way of organizing government services;[15] to conducting parliamentary inquiries and hearings; passing laws; raising public awareness; and shutting down networks and regulating media.

Organization design

Countries: Finland, Ukraine, Sweden, Netherlands, Latvia, Denmark, Czechia, Canada, Australia, the United Kingdom, and the United States

These countries' first reflex has been to modify internal structures, doing so by a variety of different approaches since they do not have the same administrative and strategic cultures. All of them realized that in dealing with a hybrid threat that combines and sometimes blends the civil and military domains, state-led and non-state actors, and several fields (not only defense and diplomacy but also culture, justice, etc.), the priority was to take a global approach and, therefore, to work in a cross-cutting manner that removes barriers between departments that generally work in silos. At the least, this has involved connecting scattered skills by creating "committees" or "networks," such as Finland's information influencing network, created in December 2014, which brings together approximately thirty government specialists from all ministries, the Office of the President, the police, and the armed forces, to identify, analyze, and respond to hostile attempts at foreign interference.

A number of such networks or task forces were created specifically to secure elections and referenda against foreign influence (defined as including "overt and covert efforts by foreign governments or actors acting as agents of, or on behalf of, foreign governments intended to affect directly or indirectly [the] election – including candidates, political parties, voters or their preferences, or political processes") and interference ("a subset of election influence activities targeted at the technical aspects of the election, including voter registration, casting and counting ballots, or reporting results").[16] Securing elections has been the most powerful driver, especially after 2016-2017, because of the interference activities targeting the United States (2016) and French (2017) presidential elections, and suspicions of interference with Brexit vote in the United Kingdom (2016). This second-stage awareness, as we called it in the introduction, triggered the creation of a number of structures, for instance in Sweden[17] and Estonia.[18] Initially set up as temporary, some of them became permanent and broadened their focus beyond elections, as other challenges arose.

Most European countries now have similar units. Some are temporary working groups or standing committees that only bring together existing staff. Others have gone further by allocating specific

[14] Canadian political science scholar (1926-2018) whose famous dictum was "theory is always *for* someone and *for* some purpose." See: Robert Cox, "Social Forces, States and World Orders: Beyond International Relations Theory," *Millenium 10*, no 2 (1981): 128, https://doi.org/10.1177/03058298810100020501.

[15] Organization design is traditionally defined as "the search for coherence between strategy (domain, objectives and goals), organizing mode (decomposition into subtasks, coordination for the completion of whole tasks), integrating individuals (selection and training of people), and designing a reward system." Jay R. Galbraith, *Organization Design* (Reading [MA]: Addison Wesley Pub. Co., 1977), 5.

[16] National Intelligence Council, Office of the Director of National Intelligence, *Foreign Threats to the 2020 US Federal Elections*, March 10, 2021, https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf.

[17] Gordon LaForge, *Sweden defends its elections against disinformation*, 2016-2018, Innovations for Successful Societies, Princeton University, December 2020, https://successfulsocieties.princeton.edu/sites/successfulsocieties/files/GL_Swedena_Election_FINAL12_23_20_V1_0.pdf.

[18] Tyler McBrien, *Defending the vote: Estonia creates a network to combat disinformation, 2016-2020*, Innovations for Successful Societies, Princeton University, December 2020, https://successfulsocieties.princeton.edu/sites/successfulsocieties/files/TM_Estonia_Election_FINAL%20edited_JG.pdf.

budgets and by creating positions in order to set up structures that work full-time on monitoring information manipulation and/or foreign influence, either within an existing department or by creating one from scratch. Whether this involves networking existing resources or creating a new unit, there is still the thorny issue of the institutional line of command: in concrete terms, where does it belong? If it is within a given ministry, while being cross-departmental by nature, why this ministry and not another? Bureaucratic politics theories teach us that the various agencies and administrations are in constant competition with each other for budget shares, resources, recognition, and territory.[19] In such a context, the creation of a new structure rarely goes smoothly and often generates tensions. If the resistance is strong enough, it can even prevent this creation. Those tensions are easier to resolve in countries that already have horizontal and cross-sectoral cultures, in particular Scandinavian countries, or small ones such as the Baltic nations, in which teams have a greater chance of knowing each other and working together.

Plugging the new unit into the prime minister's or president's office could be a solution in that respect, because countering information and influence activities is cross-departmental by nature. Such an overarching position also presents the benefit of seeing most of the whole-of-government effort and receiving information from all ministries and agencies. In one such example, President Volodomyr Zelenskyy of Ukraine established a Center for Countering Disinformation in March 2021: the head of the Center is nominated by the President and the Center reports to the National Security and Defense Council, an agency under the President.[20] Such a centralized organization may be effective, but it also presents a risk of being (at least perceived as) a political instrument.

In that respect, the Swedish model, based on strong agencies and small ministries, offers a way out. Much of the work on countering disinformation in Sweden comes from "below" the government, at the agencies, municipalities, and civil society levels. In particular, it comes from the Swedish Civil Contingencies Agency (MSB).[21] Since 2016, it has been tasked with identifying and countering information influence campaigns. Starting in 2022, all those efforts will be coordinated by a new Agency for Psychological Defence.[22] The raison d'être of such a bottom-up approach is resilience: it gives agencies the ability to counter foreign influence and disinformation without government support.

In many instances, the main unit in charge of countering foreign information manipulation and/or influence operations lies within a specific ministry: Justice in the Netherlands, Culture in Latvia, etc. In Denmark, since 2017, there is both an inter-ministerial task force (Defense, Foreign Affairs, Justice, intelligence services) and another, internal task force within the Ministry of Foreign Affairs (between the Public Diplomacy, Security Policy, and European Neighborhood and Russia departments).[23] In Czechia, the Ministry of the Interior has hosted a Centre against Terrorism and Hybrid Threats (CTHT) since 2017.[24] The unit of approximately fifteen people, most of them from a policy-making background, also has a strategic communication function, including an official Twitter account, @CTHH_MV. In Canada, Global Affairs Canada in the Ministry of Foreign Affairs hosts the Centre for International Digital Policy (CIDP), itself having two teams: on the one hand, the Rapid Response Mechanism Unit (leading the G7 Rapid Response Mechanism, working with international partners and monitoring the digital information ecosystem for foreign state sponsored

[19] | Morton H. Halperin and Priscilla A. Clapp, with Arnold Kanter, *Bureaucratic Politics and Foreign Policy* (Washington, DC: Brookings Institution Press, 2006), 2nd Edition.

[20] | "Zelensky approves regulation on Center for Countering Disinformation," Unian, May 8, 2021, https://www.unian.info/politics/center-for-countering-disinformation-zelensky-approves-regulation-11413858.html.

[21] | LaForge, *Sweden defends its elections against disinformation*.

[22] | Regeringskansliet, *Summary of Government bill 'Totalförsvaret 2021–2025' (Total defence 2021–2025)*, December 17, 2020, https://www.government.se/4af8fa/globalassets/government/dokument/forsvarsdepartementet/ip-2021-2025/summary-of-government-bill-total-defence-2021-2025-final.pdf.

[23] | Andreas Baumann and Andreas Reinholt Hansen, "Danmark får ny kommandocentral mod misinformation," *Tjekdet*, September 11, 2017, https://www.tjekdet.dk/danmark-faar-ny-kommandocentral-mod-misinformation.

[24] | Ministry of the Interior, Government of Czechia, *Centre Against Terrorism and Hybrid Threats*, https://www.mvcr.cz/cthh/clanek/centre-against-terrorism-and-hybrid-threats.aspx.

disinformation related to Government of Canada priorities); on the other hand, the Digital Inclusion Lab (looking at the intersection of foreign policy and digital technology more broadly: all things related to platforms, content moderation, artificial intelligence, digital inclusion, etc.). The CIDP also participates in a Security and Intelligence Threats to Elections Task Force with the Canadian Security Intelligence Service, the Royal Canadian Mounted Police, and the Communications Security Establishment.[25] Another example is Australia where, since 2018, a National Counter Foreign Interference Coordinator's office,[26] cross-departmental by nature but hosted at the Department of Home Affairs, coordinates the whole-of-government response to counter foreign interference (broadly defined as including all activities, carried out by a foreign actor, that are "coercive, deceptive, clandestine or corrupting that are contrary to [Australia's] sovereignty, values and national interests").[27] Since 2020, the Department of Foreign Affairs and Trade also hosts a counter-disinformation team within its International Security division and, during elections, there is an Electoral Integrity Assurance Task Force also dealing with counter-disinformation.

As for France, a major announcement was made in June 2021: a national agency dedicated to the fight against foreign information manipulation, specifically against "foreign digital interference," will be established by September 2021. Of significant size, with a staff of 60 people, it will operate under the Secretariat-General for National Defense and Security (SGDSN), itself under the Prime Minister's authority.[28] This is something the author has been advocating for since 2017 in internal memos and the CAPS-IRSEM report.[29] The build-up has been gradual. First, in 2018, France created a cross-departmental network, coordinating whole-of-government efforts against information manipulation. Then, for a couple of months in 2020-2021, it experimented with a temporary small cell, called the "Honfleur Task Force."[30] Both were under-the-radar, discreet initiatives. This decision to create a permanent, publicly acknowledged structure, of such a size, is therefore a significant step. This is all the more timely as there are two important electoral deadlines coming, likely carrying high risks of foreign information manipulation: the New Caledonian independence referendum in December 2021 and the Presidential election in April 2022.

Two countries in particular, the United Kingdom and the United States, have many teams across departments. In the United Kingdom, it is worth mentioning the National Security Communications Team under the joint authority of the Cabinet Office and the Prime Minister's Office (No. 10); the Rapid Response Unit also based in No. 10 and the Cabinet Office, as well as the Media Monitoring Unit. The Foreign, Commonwealth, and Development Office is an important actor in this ecosystem with at least two major teams: the Open-Source Unit and the Russia unit, which implements a £29.75 million Counter Disinformation and Media Development (CDMD) program, launched in April 2016. Additionally, a cross-departmental counter-disinformation unit, housed in the Department for Digital, Culture, Media, and Sport, was set up in March 2020 "to provide a comprehensive picture of the extent, scope and the reach of disinformation and misinformation linked to COVID-19, and to work with partners to stamp it out."[31]

[25] | Democratic Institutions, Parliament of Canada, *Combating foreign interference*, January 2019, https://www.canada.ca/en/democratic-institutions/news/2019/01/combatting-foreign-interference.html.

[26] | Department of Home Affairs, Government of Australia, "National Counter Foreign Interference Coordinator," https://www.homeaffairs.gov.au/about-us/our-portfolios/national-security/countering-foreign-interference/cfi-coordinator.

[27] | Select Comm. on Foreign Interference through Social Media, Parliament of Australia, Hearing transcript, December 11, 2020, https://www.aph.gov.au/Parliamentary_Business/Hansard/Hansard_Display?bid=committees/commsen/700fa9a5-68eb-4360-899a-1801a80d9494/&sid=0001.

[28] | "Que sait-on de la future agence de lutte contre les manipulations numériques venues de l'étranger?," France Inter, June 2, 2021, https://www.franceinter.fr/monde/que-sait-on-de-la-future-agence-de-lutte-contre-les-manipulations-numeriques-venues-de-l-etranger.

[29] | Vilmer et al., Information Manipulation, 170.

[30] | "La France va créer une agence nationale de lutte contre les manipulations de l'information," *Le Monde*, June 2, 2021, https://www.lemonde.fr/pixels/article/2021/06/02/la-france-va-creer-une-agence-nationale-de-lutte-contre-les-manipulations-de-l-information_6082561_4408996.html.

[31] | MP Caroline Dinenage, "Internet: Disinformation," Question for DCMS, December 2, 2020, https://questions-statements.parliament.uk/written-questions/detail/2020-12-02/124329. For a more detailed presentation of the British model, see: Vilmer, Effective State Practices Against Disinformation.

In the United States, a main entity dedicated to monitoring and countering foreign information manipulation is the Global Engagement Center (GEC), established in 2016 within the Department of State. The GEC's mission is to "'direct, lead, synchronize, integrate, and coordinate efforts of the Federal Government to recognize, understand, expose, and counter foreign state and foreign non-state propaganda and disinformation efforts aimed at undermining or influencing the policies, security, or stability of the United States and United States allies and partner nations."[32] Cross-departmental by nature, it employed 118 personnel in 2020,[33] including detailees from other departments, notably the Department of Defense, as well as contractors from the private sector. Its funding level is $138 million USD for 2021 (more than twice its 2020 budget).[34] The GEC has several internal teams, including an Analytics and Research team of around "20 individuals with expertise in data collection, analytics and research methodologies such as Social Network Analysis, polling, and artificial intelligence," a Technology Engagement team "tasked with facilitating the use of a wide range of technologies and techniques in our efforts," a Monitoring and Evaluation team, and three country-specific threat teams focused on Russia, China, and Iran, respectively.[35] In practice, the GEC is having a hard time to "direct, lead and coordinate" because of the competition, from both within the State Department (powerful regional bureaus, but also transversal ones like the Bureau of Intelligence and Research and the Bureau of Global Public Affairs) and at the interagency level. Indeed, there are several other units dealing with foreign disinformation and influence in other branches of government, including the Departments of Defense, Justice (the Federal Bureau of Investigation [FBI], in particular), and Homeland Security (the Cybersecurity and Infrastructure Security Agency [CISA]), and several intelligence agencies, in particular the National Counterintelligence and Security Center in the Office of the Director of National Intelligence (ODNI). ODNI also publishes coordinated assessments of seventeen US intelligence agencies, several of them related to electoral influence and interference.[36] During the Trump Administration, the disconnect between the presidency and the growth in resources to address foreign influence contributed to a lack of coordination between all these teams because no clear guidance came from the National Security Council or the White House. In its first months, the Biden Administration showed much more interest and determination to better coordinate these efforts.

Those multiple national examples demonstrate the wide variety of focal areas: information manipulation in the strict sense, influence and/or interference (one difficulty being determining where one ends and the other begins),[37] or hybrid threats, which can be associated with terrorist threats, as is the case with Czechia's CTHT. This common association should not lead to conflating the two since the purpose of a so-called hybrid attack is to generate ambiguity – if properly carried out, it will take time for the target to understand it is under attack and by whom – whereas terrorists usually claim their attacks, as their objective is not to impede identification but rather to allow it and assume responsibility for such attacks. That does not mean that, in practice, a department that combines the two will generate confusion, since it may have

[32] Subcomm. on State Department and USAID Management, International Operations, and Bilateral International Development, US Senate, The Global Egagement Center: Leading the United States Government's Fight Against Global Disinformation Threat, S.HRG. 116–275, March 5, 2020, https://www.foreign.senate.gov/imo/media/doc/03%2005%2020%20--%20The%20Global%20Engagement%20Center%20Leading%20the%20United%20States%20Governments%20Fight%20Against%20Global%20Disinformation%20Threat.pdf.

[33] Subcomm. on State Department and USAID Management, International Operations, and Bilateral International Development, *The Global Engagement Center*.

[34] Subcomm. on State Department and USAID Management, International Operations, and Bilateral International Development, *The Global Engagement Center*.

[35] Subcomm. on State Department and USAID Management, International Operations, and Bilateral International Development, *The Global Engagement Center*.

[36] See, for example: Office of the Director of National Intelligence and the National Intelligence Council, "Background to 'Assessing Russian Activities and Intentions in Recent US Elections': The Analytic Process and Cyber Incident Attribution," January 6, 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf; and ODNI, *Foreign Threats to the 2020 US Federal Elections*.

[37] Malcolm Turnbull, "Speech introducing the National Security Legislation Amendment (Espionage and Foreign Interference) Bill 2017," December 7, 2017, https://www.malcolmturnbull.com.au/media/speech-introducing-the-national-security-legislation-amendment-espionage-an.

teams working in parallel. In the case of Czechia, specifying that CTHT fights against "terrorism *and* hybrid threats" does imply a distinction, even if the two are together, undoubtedly for organizational expediency.

Among the countries that have networked their existing expertise without creating ongoing dedicated bodies and, therefore, additional positions, some are paralyzed by bureaucratic and budget considerations. For others, like Singapore, this is a deliberate choice to show that the fight against foreign information manipulation is "everybody's business" and not run the risk that focusing resources in a special unit makes the other agencies no longer feel concerned about this issue.[38]

Parliamentary inquiries and hearings

Countries: the United States, Canada, the United Kingdom, Australia

Parliaments help fight information manipulation through inquiries and hearings. Those inquiries are carried out by judicial police (in the United States, the FBI began investigating the "Russian affair" in July 2016) and parliamentary committees and groups, which generally produce detailed reports. The largest US example in that respect is the Senate Select Committee on Intelligence (SSCI)'s inquiry and five volume report on Russia's activities in 2016.[39] The whole committee, which is comprehensive and bipartisan, conducted the inquiry, with full access to all the relevant intelligence and with the power to subpoena (i.e., they could compel past and present officials to testify under oath as a part of the inquiry). A number of other parliamentary reports have been released, including those from the US Democratic Senators in January 2018,[40] Canada's House of Commons Standing Committee on Access to Information, Privacy and Ethics in December 2018,[41] or the UK House of Commons' Digital, Culture, Media and Sport Committee in February 2019[42] and July 2020.[43] In December 2019, the Australian Parliament set up its own Select Committee on Foreign Interference through Social Media, which is supposed to submit a report by 2022.[44]

Being only informative, not legislative, these reports are not supposed to be "implemented." They do however contribute in many other ways. Some of those inquiries managed to obtain data and contributed to public attribution of important disinformation campaigns. Being published and widely reported on by the media, they can also inform and educate the general public and, by exposing the perpetrators, they may act as a deterrent. It is by definition difficult to demonstrate if naming and shaming has had any effect, as threats that have been deterred are difficult to spot, but these public investigations can certainly help in building sentiment in policy-making circles and among potential adversaries that some countries are aware and prepared.

Parliamentarians also contribute to international cooperation by working together. In November 2018, members of the national parliaments of nine countries – Argentina, Belgium, Brazil, Canada, France, Ireland, Latvia, Singapore, and the United Kingdom – created the International Grand

[38] | Meeting with a Singapore official, Singapore, November 5, 2019.

[39] | Select Comm. on Intelligence, US Senate, *Russian Active Measures Campaigns and Interference in the 2016 U.S. Election*, Vol.I-V (November 10, 2020), https://www.intelligence.senate.gov/publications/report-select-committee-intelligence-united-states-senate-russian-active-measures.

[40] | Comm. on Foreign Relations, US Senate, *Putin's Asymmetric Assault on Democracy in Russia and Europe: Implications for U.S. National Security*, S.Prt. 115-21 (January 10, 2018), https://www.foreign.senate.gov/imo/media/doc/SPrt_115-21.pdf.

[41] | Standing Comm. on Access to Information, Privacy and Ethics, Parliament of Canada, *Democracy under threat: risks and solutions in the era of disinformation and data monopoly*, December 2018, https://www.ourcommons.ca/Content/Committee/421/ETHI/Reports/RP10242267/ethirp17/ethirp17-e.pdf.

[42] | Digital, Culture, Media and Sport Comm., Parliament of the United Kingdom, *Disinformation and 'fake news': Final Report*, P.R. 2017-19:8, February 14, 2019, https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf.

[43] | Digital, Culture, Media and Sport Comm., Parliament of the United Kingdom, *Misinformation in the COVID-19 Infodemic*, P.R. 2019-21:2, July 21, 2020, https://committees.parliament.uk/publications/1954/documents/19089/default/.

[44] | Parliament of Australia, "Select Committee on Foreign Interference through Social Media," https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Foreign_Interference_through_Social_Media.

[45] | "'International Grand Committee' on Disinformation and Fake News," November 27, 2018, https://www.parliament.uk/globalassets/documents/commons-committees/culture-media-and-sport/Principles-of-the-Law-Governing-the-Internet.pdf.

Committee on Disinformation and Fake News.[45] They first met in London, then in Ottawa and Dublin, with representatives from other countries.[46] Since June 2020, the European Parliament also has maintained a Special Committee for Foreign Interference in all democratic processes in the EU, including disinformation (INGE). As Raphaël Glucksmann, INGE president, explains, the organization works "to assess the level of these threats in different spheres: major national and European elections across the EU; disinformation campaigns on traditional and social media to shape public opinion; cyber-attacks targeting critical infrastructure; direct and indirect financial support and economic coercion of political actors and civil society subversion."[47] They have been conducting hearings (with experts, academics, platforms, nongovernmental organizations [NGOs], etc.)[48] and preparing reports, including what is still – at the time of writing – a *Working document on foreign interference using online platforms - threats, risks and remedies.*[49]

Legislation

Countries: France, Germany, Israel, Taiwan, Malaysia, Cambodia, Kenya, Belarus, Egypt, Russia, Singapore, Morocco, Algeria, Kyrgyzstan, Brazil

Some countries already had legislation in that regards, some of which was quite longstanding. In France, for example, the 1881 Act on the freedom of the press already provided for sanctions against "the malicious publication, dissemination and reproduction, by whatever means, of false news and documents that have been fabricated or falsified or mendaciously attributed to third parties, when this has disturbed the peace, or was capable of disturbing it."[50] But such older provisions are sometimes incomplete or poorly adapted to our current digital era, which has not changed the nature of information manipulation (which has always existed) but the means and speed for its propagation, since it is now possible to reach millions of people in a few minutes. In response to the need to update, France enacted a law against information manipulation in November 2018. Under this law, information manipulation is defined as the "inexact or misleading allegation of a fact that could alter the sincerity of an upcoming vote and that is spread deliberately, artificially or automatically and massively to the online public through a communication service."[51] The law, limited to electoral periods, has generated strong opposition, from journalists and NGOs in particular, and it has been invoked only once, somewhat ironically, against a tweet from the Minister of the Interior (though the court found the Minister not guilty of spreading a false information). The complainants acknowledged that their "objective was to demonstrate by the absurd that the law is useless."[52]

The French law came after a German one, the "Netzwerkdurchsetzungsgesetz" (NetzDG) that came into effect on January 1, 2018, forcing digital platforms to take down "manifestly illegal" messages within 24 hours or face fines of up to €50 million euros.[53] Under the auspices of NetzDG, in July 2019, the German Federal Office of Justice issued a €2 million fine against Facebook for failing to fulfill its reporting duty.[54] The law also

[46] | Centre for International Governance Innovation, "The International Grand Committee Timeline," https://www.cigionline.org/igc/timeline.

[47] | European Parliament, "About: Welcome to INGE by President Raphaël Glucksmann," https://www.europarl.europa.eu/committees/en/inge/about.

[48] | See the meeting documents: https://emeeting.europarl.europa.eu/emeeting/committee/en/archives/INGE.

[49] | Special Comm. on Foreign Interference in all Democratic Processes in the European Union, including Disinformation, European Parliament, "Working Document on foreign interference using online platforms – threats, risks and remedies," May 12, 2021, https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/INGE/DV/2021/05-10/1231331EN.pdf.

[50] | French Republic Act of July 29, 1881, on the freedom of the press, Art. 27, translated by the French government, *A bill against the manipulation of information*, June 7, 2018, https://www.gouvernement.fr/en/a-bill-against-the-manipulation-of-information.

[51] | For a detailed analysis of the French law, see: Marine Guillaume, *Combating the manipulation of information – a French case, Strategic Analysis* 2/2019, Hybrid Center of Excellence (CoE) (Helsinki), May 3, 2019, https://www.hybridcoe.fi/wp-content/uploads/2020/07/HybridCoE_SA_16_manipulation-of-information_.pdf.

[52] | Géraldine Delacroix, "Le mensonge de Castaner sur la Pitié-Salpêtrière n'est pas une 'fake news', selon la justice," *Mediapart*, May 21, 2019, https://www.mediapart.fr/journal/france/210519/le-mensonge-de-castaner-sur-la-pitie-salpetriere-n-est-pas-une-fake-news-selon-la-justice.

[53] | Heidi Tworek and Paddy Leerssen, "An Analysis of Germany's NetzDG Law," Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression, University of Amsterdam, April 15, 2019, https://pure.uva.nl/ws/files/40293503/NetzDG_Tworek_Leerssen_April_2019.pdf.

[54] | Thomas Escritt, "Germany fines Facebook for under-reporting complaints," *Reuters*, July 2, 2019, https://www.reuters.com/article/us-facebook-germany-fine-idUSKCN1TX1IC.

faced important criticism from civil society, which the German government attempted to address in an amendment of the law in 2020.[55] Other examples of such legislation in liberal democracies include Israel, where a ruling by the Supreme Court prohibited anonymous advertising as from March 2019 and mandated identification of fake accounts used to spread propaganda and bots, and Taiwan with its "Anti-Infiltration Law" promulgated in January 2020, which includes disinformation control measures.

The fact remains, however, that most legislation has been adopted in countries that cannot be considered liberal democracies and where the fight against "fake news" is often a pretext for censorship. When Malaysia (April 2018), Cambodia and Kenya (May 2018), Belarus (June 2018), Egypt (July 2018), and Russia (March 2019) enacted legislation in this area, they were strongly criticized by human rights organizations. A study also found that such laws doubled between 2016 and 2020 in Sub-Saharan Africa, and that most of those punished have been journalists and political opponents.[56] Indeed, under such circumstances, where freedom of the press is already threatened or does not exist, such legislative measures further undermine that freedom while bolstering the authorities' control over the population.

Similarly, in Singapore, the Parliament enacted a law in May 2019, the Protection from Online Falsehoods and Manipulation Act (POFMA), after the Select Committee on Deliberate Online Falsehoods, created in January 2018, published its report.[57] POFMA allows ministers to order a correction or removal of online information if it is deemed false and affecting the public interest; this aspect of the law has been denounced by Western NGOs and media as unduly restricting freedom of expression.[58] It also generated a debate within Singapore, with local politicians and journalists denouncing the law.[59]

The COVID-19 pandemic provided an additional pretext to justify censorship in some of those countries and many others. Russia passed a new "fake news" law in March 2020,[60] like Morocco (March 2020), Algeria (April 2020), Kyrgyzstan (June 2020), and Brazil (July 2020), among others. A DFRLab study found at least twenty-four countries having "issued measures that affect free expression to address the infodemic occurring parallel to the pandemic,"[61] including Turkey, Thailand, Honduras, India, South Africa, and Hungary.

Raising public awareness

Countries: France, Lithuania, Sweden, Canada, Netherlands, the United Kingdom, Taiwan, Finland, Denmark, Estonia, Italy, the United States

Liberal democracies quickly understood that the above measures would fall short if the community as a whole was not aware of the dangers of information manipulation. So, in order to raise awareness, states – some more than others – implemented a range of measures. One of them is raising internal awareness within the government through the production and distribution of

[55] | Amélie Heldt, "Germany is amending its online speech act NetzDG... but not only that," *Internet Policy Review*, April 6, 2020, https://policyreview.info/articles/news/germany-amending-its-online-speech-act-netzdg-not-only/1464.

[56] | Peter Cunliffe-Jones, Assane Diagne, Alan Finlay and Anya Schiffrin, "Bad Law – Legal and Regulatory responses to misinformation in Sub-Saharan Africa 2016-2020", in P. Cunliffe-Jones et al., *Misinformation Policy in Sub-Saharan Africa: From Laws and Regulations to Media Literacy*, London: University of Westminster Press, 2021, https://www.uwestminsterpress.co.uk/site/chapters/m/10.16997/book53.b/.

[57] | Select Comm. on Deliberate Online Falsehoods, Parliament of Singapore, "Report of the Select Committee on Deliberate Online Falsehoods – Causes, Consequences and Countermeasures," September 19, 2018, https://sprs.parl.gov.sg/selectcommittee/selectcommittee/download?id=1&type=subReport.

[58] | See, for example: "Singapore: Free Expression Restrictions Tighten," Human Rights Watch, January 14, 2020, https://www.hrw.org/news/2020/01/14/singapore-free-expression-restrictions-tighten; and "Singapore 'falsehoods' law shows perils of fake news fight," *Financial Times*, February 3, 2020, https://www.ft.com/content/e50eb042-3db3-11ea-a01a-bae547046735.

[59] | Rachel Au-Yong, "Parliament: Workers' Party opposes proposed law on fake news, says Pritam Singh," *The Straits Times*, May 7, 2019, https://www.straitstimes.com/politics/parliament-workers-party-opposes-proposed-law-on-fake-news-pritam-singh; Danisha Hakeem, "Prominent journalists express concern over POFMA's impact on their practices in a letter to Comms and Info Minister S Iswaran," *The Online Citizen*, April 25, 2019, https://www.theonlinecitizen.com/2019/04/25/prominent-journalists-express-concern-over-pofmas-impact-on-their-practices-in-a-letter-to-comms-and-info-minister-s-iswaran/.

[60] | "New 'fake news' law stifles independent reporting in Russia on COVID-19," International Press Institute, May 8, 2020, https://ipi.media/new-fake-news-law-stifles-independent-reporting-in-russia-on-covid-19/.

[61] | Jacqueline Malaret and John Chrobak, "The criminalization of COVID-19 clicks and conspiracies," Digital Forensic Research Lab, Atlantic Council, May 13, 2020, https://medium.com/dfrlab/op-ed-the-criminalization-of-covid-19-clicks-and-conspiracies-3af077f5a7e7.

dedicated newsletters. For instance, since January 2019, the French Government Information Service (SIG), under the prime minister, has been sending a weekly newsletter called #Desinfox with a selection of research articles, newspaper articles, and think tank reports on disinformation. Initially sent to government administrators only, its distribution was broadened in March 2020 to include external researchers, journalists, and experts, in an effort to fight COVID-19 related disinformation.[62] Similar internal newsletters, and other regular products based on social network monitoring, are also produced by several French ministries.

To raise external awareness among the general population, several French ministries and public institutions also adapted their communication, in particular their websites. For example, the Ministry of Higher Education, Research and Innovation created a new section on its website entitled "Detox: the word to science," the objective of which is to "fight against disinformation about the epidemic and put an end to fake news" by decrypting false information and misconceptions related to the virus.[63] Several public health institutions, such as the National Institute of Health and Medical Research (INSERM)[64] and the Pasteur Institute,[65] took similar initiatives.

States also implemented a number of measures to raise awareness more broadly. One of these has been the publication of doctrines or national strategies highlighting the dangers of informational threats and, in some cases, exposing at least some of the countermeasures put in place. The Lithuanian National Security Strategy adopted this approach as early as 2017.[66] Another important measure has been to provide training for various audiences, including civil servants, political parties, and journalists, in particular. Between 2016 and April 2021, the Swedish MSB trained 16,000 civil servants with an awareness program on information influence activities, ranging from half a day to two days,[67] among other kinds of training.[68] During the 2017 presidential campaign in France, the SGDSN warned political parties as early as late summer 2016 and, in October, the French National Cybersecurity Agency held a workshop on cybersecurity for the parties.[69]

Similarly, ahead of Canada's 2019 federal election, as part of the objective of "improving organizational readiness,"[70] Canadian government agencies provided technical advice to political parties and election administrators on how to better protect their cyber installations, sensitized decision-makers to the risk of foreign interference, provided classified briefings to political party leaders, and organized whole-of-government simulations and table-top exercises on a regular basis to prepare for potential incidents or scenarios. Canadian Heritage, comprising a portfolio of departmental agencies, Crown corporations, and administrative tribunals,[71] also funded the training of approximately seventy journalists on disinformation and digital literacy: in order to preserve press freedom, however, the journalists were trained by an academic intermediary, McGill University's Media Ecosystem Observatory, instead of government representatives.[72] A similar program is ongoing in Sweden, where MSB funds trainings for journalists, and these too are conducted through

[62] | "Le SIG affine sa stratégie sur les fake news," *La Lettre A*, April 24, 2020, https://www.lalettrea.fr/action-publique_executif/2020/04/24/le-sig-affine-sa-strategie-sur-les-fake-news,108402671-bre.

[63] | "Désintox : la parole à la science," Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, https://recherchecovid.enseignementsup-recherche.gouv.fr/desintox-la-parole-la-science.

[64] | "Canal Détox, la série qui lutte contre les fausses informations," *Inserm*, https://presse.inserm.fr/canal-detox/?cat=109.

[65] | Institut Pasteur, "Coronavirus : attention aux fausses informations sur la COVID circulant sur les réseaux sociaux," October 22, 2021, https://www.pasteur.fr/fr/journal-recherche/actualites/coronavirus-attention-aux-fausses-informations-covid-19-circulant-reseaux-sociaux.

[66] | Seimas of the Republic of Lithuania, "National Security Strategy of the Republic of Lithuania," January 17, 2017, https://kam.lt/en/defence_policy_1053/important_documents/strategical_documents.html.

[67] | Interview with an MSB official, April 2021.

[68] | See: Vilmer, *Effective State Practices Against Disinformation*.

[69] | Vilmer, *The 'Macron Leaks' Operation*, 31 32.

[70] | Democratic Institutions, Government of Canada, *Improving organizational readiness*, January 2019, https://www.canada.ca/en/democratic-institutions/news/2019/01/improving-organizational-readiness.html.

[71] | Government of Canada, "Portfolio organizations – Canadian Heritage," https://www.canada.ca/en/canadian-heritage/corporate/portfolio-organizations.html.

[72] | Homepage, Media Ecosystem Observatory, https://mediaecosystemobservatory.com/.

an intermediary, the Fojo Media Institute.[73]

Other measures include public campaigns on television, such as the Dutch program *Nieuws of Nonsens (News or Nonsense)*, a ninety-minute documentary on the danger of disinformation and how to detect it, originally broadcast by public news organization NOS in March 2018;[74] the issuance of handbooks, such as the *Swedish Countering Information Influence Activities*: *A Handbook for Communicators* (March 2019),[75] a product of a collaboration between the MSB and the Department of Strategic Communication at Lund University; or the British *RESIST Counter-Disinformation Toolkit*, a collaboration between Lund University and the United Kingdom's Government Communications Service (April 2019). RESIST stands for "Recognise disinformation, Early warning, Situational insight, Impact analysis, Strategic communication, Track outcomes."[76] States can also distribute dedicated leaflets, including printed ones, distributed by postal mail in an attempt to inform everyone or, at least, another, less connected part of the population: in Sweden, in 2018, the MSB mailed 4.8 million households the brochure *If War or Crisis Comes* explaining what to do in the event of a crisis, whether that be a terrorist attack or some kind of information manipulation.[77] At the other end of the technological spectrum, states also took innovative steps to respond to disinformation online, on social media platforms, using similar means as the attackers, such as memes. In 2019, Audrey Tang, Taiwan's digital minister, revealed that each department in the government was prepared to respond to disinformation within sixty minutes by creating a "clarification" meme so funny that it would go viral: "it acts as an inoculation, as a memetic vaccine."[78]

Another preventative measure has been adding or enhancing media and information literacy (MIL) into school curricula. Finland, which ranks first in Europe for media literacy, is definitely a model to follow in that field.[79] In Finland, MIL "is seen as civic competence, important to every citizen from an early age" and is promoted through a number of national policies, as detailed in a 2017 report by the Finnish National Audiovisual Institute.[80] Sweden has been building "digital skills" in its schools since July 2018 by integrating digital education in compulsory subjects (history, geography, mathematics, etc.) with the objective of understanding "the impact of digital transformation on individuals and society."[81] Denmark released its 2019 publication *Trolls in your feed*, which focused on Russian disinformation, in both Danish and English. Also in 2019, Estonia organized a Media Literacy Week ("Think before you share") and a 35-hour course on "Media and Manipulation" in high schools.[82] Even if other countries also developed similar initiatives – in 2017, Italy added the objective of being able to "recognize fake news" to school curricula; in 2018, several US states, including California and Massachusetts, adopted laws in that regard – northern and eastern European countries are generally ranked the best in the world ("Finland, Denmark, the Netherland, Sweden, and Estonia top the Media Literacy Index 2019").[83] There are

[73] | Homepage, FOJO Media Institute, Linnæus University, https://fojo.se/en/.

[74] | *Nieuws of Nonsens*, March 5, 2018, https://www.npostart.nl/nieuws-of-nonsens/05-03-2018/POW_03787753.

[75] | MSB, "Countering information influence activities: A handbook for communicators," March 2019, https://www.msb.se/RibData/Filer/pdf/28698.pdf, based on the 2018 report "Countering Information Influence Activities: The State of the Art" by James Pamment.

[76] | Government Communication Service, Government of the United Kingdom, *RESIST Counter Disinformation Toolkit*, https://gcs.civilservice.gov.uk/publications/resist-counter-disinformation-toolkit/.

[77] | The brochure, entitled *Om Krisen eller Kriget Kommer*, is also available in English: https://www.msb.se/sv/publikationer/om-krisen-eller-kriget-kommer--engelsk-version/

[78] | Audrey Tang, "2019-09-25 Finding facts in a world of disinformation," YouTube, posted by PDIS, October 22, 2019, https://www.youtube.com/watch?v=l0uR4_dctTg&t. On Taiwan's innovative approach and how it can be an inspiration for other countries, see: Jude Blanchette, Scott Livingston, Bonnie S. Glaser, and Scott Kennedy, *Protecting Democracy in an Age of Disinformation: Lessons from Taiwan*, Center for Strategic and International Studies, January 2021, https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/210127_Blanchette_Age_Disinformation.pdf.

[79] | Eliza Mackintosh, "Finland is winning the war on fake news. What it's learned may be crucial to Western democracy," *CNN*, May 2019, https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/.

[80] | *Finnish Media Education: Promoting Media and Information Literacy in Finland*, Kansallinen Audiovisuaalinen Instituutti, 2017, https://kavi.fi/sites/default/files/documents/mil_in_finland.pdf.

[81] | European Commission, *Digital skills enter into Sweden schools*, September 5, 2018, https://eacea.ec.europa.eu/national-policies/eurydice/content/digital-skills-enter-sweden-schools_en.

[82] | McBrien, *Defending the vote*, 14.

now many governmental initiatives aiming at strengthening MIL and critical thinking in schools, but there is room for improvement for similar initiatives targeting adults, especially seniors, who studies have shown are more likely to share fake news.[84]

Some governments also created online services allowing users to report "fake news" to the police: in 2018, Italy created such a program but quickly abandoned the idea after it generated some controversy because of the risk of censorship and violation of press freedom. Indeed, this measure turned policemen into fact-checkers: they had the authority to pursue legal action against any "false and tendentious news" that "could disturb the public order," a vague characterization that could be used to weaponize such a service toward government or police critics. As Arianna Ciccone, founder of the International Journalism Festival, said: "It is not the job of the state to establish the truth… That they do in authoritarian regimes."[85] This "red button" portal, as it was named, launched in January 2018 but stopped a few days ahead of the March 4 elections.[86]

Finally, many states are also funding or supporting civil society initiatives. In the United Kingdom, the previsouly mentioned CDMD program, one goal of which is to counter "disinformation directed at the UK and its Allies from Russia,"[87] funds many civil society initiatives aiming at exposing disinformation, both in the United Kingdom and abroad. In Canada, the government launched its Digital Citizen Initiative in 2019, dedicating $7 million CAD "to support digital, news and civic literacy programming… skills development, awareness sessions, workshops and learning material."[88] It also invested $19.4 million CAD over four years in a Digital Citizen Research Program led by Canadian Heritage. In 2020, the effort funded fifty specific projects, and "$4.3 million was dedicated specifically to counter COVID-19 disinformation, misleading information, and the racism and stigmatization that are often the result."[89] In Sweden, MSB funds research from their crisis management fund. Since 2017, the standard number for research financing is approximately €1.2 million per year, with an additional €50,000 per year for short-term studies.[90] This budget is likely to increase with the new Agency for Psychological Defence. With these funds, MSB regularly commissions reports, like the ones previsouly mentioned and, more recently, a report on *Conspiracy theories and COVID-19: the mechanisms behind a fast-growing societal challenge*, which they commissioned from Andreas Önnerfors, a professor in intellectual history at Uppsala University and which MSB published on April 21, 2021.[91]

Support for civil society initiatives can also happen through collaboration and joint activities. For example, in Lithuania, the three StratCom units of the Ministry of Defence, Ministry of Foreign Affairs, and Armed Forces collaborated with journalists and local civil society to create the website Demaskuok.lt ("debunk.lt" in English), which is funded by the Google Digital Innovation Fund and the Baltic internet portal Delfi. The platform uses algorithms to analyse 10,000 articles in Lithuanian and Russian languages per day to spot disinformation, which can then be debunked in

[83] | Open Society Institute Sofia, *Just think about it. Findings of the Media Literacy Index 2019*, Policy Brief 55, November 2019, 2, https://osis.bg/?p=3356&lang=en.

[84] | See, for instance: Andrew Guess, Jonathan Nagler, and Joshua Tucker, "Less than you think: Prevalence and predictors of fake news dissemination on Facebook," *Science Advances*, 5:1, January 9, 2019, https://advances.sciencemag.org/content/5/1/eaau4586.full; Nadia M. Brashier and Daniel L. Schacter, "Aging in an Era of Fake News," *Current Directions in Psychological Science*, May 19, 2020, https://journals.sagepub.com/doi/10.1177/0963721420915872.

[85] | Daniel Funke, "Italians can now report fake news to the police. Here's why that's problematic," *Poynter*, January 19, 2018, https://www.poynter.org/fact-checking/2018/italians-can-now-report-fake-news-to-the-police-heres-why-thats-problematic/.

[86] | *Freedom on the Net 2018* – Italy, Freedom House, https://freedomhouse.org/country/italy/freedom-net/2018.

[87] | Sir Allan Duncan, MP, "Foreign and Commonwealth Office: Russian Language," Question for Foreign and Commonwealth Office, December 4, 2018, https://questions-statements.parliament.uk/written-questions/detail/2018-12-04/198813.

[88] | Democratic Institutions, *Enhancing citizen preparedness*.

[89] | Canadian Heritage, Government of Canada, "Ongoing Support for Research and Media Literacy Projects as Canada Continues to Fight Online Disinformation," February 9, 2021, https://www.canada.ca/en/canadian-heritage/news/2021/02/ongoing-support-for-research-and-media-literacy-projects-as-canada-continues-to-fight-online-disinformation.html.

[90] | Interview with a MSB official, April 2021.

[91] | Andreas Önnerfors, *Konspirationsteorier och covid-19: mekanismerna bakom en snabbväxande samhällsutmaning*, MSB, April 2021, https://www.msb.se/contentassets/555542e57381475cb26d6862dc7a543a/msb-studie.pdf.

only two hours of time.[92] In France, in June 2019, the Ambassador for Digital Affairs (in the Ministry of Europe and Foreign Affairs) and the SGDSN, organized a two-day event on countering online information manipulation with approximately fifty people from civil society, including journalists, academics, developers, and NGOs, but also private companies, including social media platforms, other members of governmental agencies, and representatives from four other countries.[93]

Shutting down networks and regulating the media

Countries: India, China, Myanmar, Venezuela, Indonesia, Ukraine, Latvia, Lithuania, Estonia, the United Kingdom, France, the United States

The most brutal, undemocratic, and controversial way of stopping or slowing down the spread of disinformation and malicious rumors online are temporary network shutdowns. These shutdowns sometimes affect not only the internet, but also mobile and landline telephone services, cable television, or even newspaper production, imposing a de facto information blackout. India employed this tactic repeatedly in the northern state of Jammu and Kashmir no less than 307 times between 2012 and May 2021, which is more than half of the 528 internet outages that India has seen during this period.[94] These shutdowns can last a long time: the 2019-2020 shutdown lasted 213 days, making it (as of its 134th day) by far "the longest ever imposed in a democracy."[95] Turning off the internet is frequent in authoritiarian states, such as China, Myanmar, and Venezuela. A less radical option is to block not the internet itself but all or some social media platforms and messaging apps. In 2019, during riots in Indonesia, the government blocked access to social media. Similarly, one of the first things the Myanmar military did after seizing power in a coup on February 1, 2021, was to shut down the internet. After resuming services, it quickly organized nightly shutdowns and blocked social media.[96] The justifications given for these tactics vary (to stop the spread of "fake news," to quell unrest, to "protect" public order, etc.), but they often are just pretexts for the government to better control the population and avoid accountability. As Poynter notes, "around the world, governments have been turning to network shutdowns with increasing frequency."[97]

Other countries have chosen outright bans on certain media, such as Ukraine did with Russian media, banning seventy-three television channels between 2014 and 2016; several Russian websites in May 2017, including VKontakte, Odnoklassniki, Yandex, Mail.ru; and three additional "pro-Kremlin" television channels in February 2021. The three Baltic states also banned Russian TV channels: Latvia banned ten of them in 2019, all *RT* channels in 2020,[98] and imposed a temporary ban on *Rossija RTR* in 2019 and 2021;[99] Lithuania followed Latvia on banning *RT*, doing so one week later.[100] Estonia has yet to ban *RT* and has "not exclud[ed] the possibility" of doing so, but it did ban *Sputnik* in 2019,[101] which resurrected itself as

[92] | Vaidas Saldžiūnas and Viktoras Daukšas, "How to Spot Disinformation within 2 Minutes in Real Time?," European Political Strategy Centre, October 12, 2018, https://medium.com/election-interference-in-the-digital-age/how-to-spot-disinformation-within-2-minutes-in-real-time-df4e7c50a5b8.

[93] | See: Ambassadeur pour le numérique, "Disinformation unconference," Slide deck, 2019, https://disinfo.quaidorsay.fr/assets/2019_disinformation_unconference_digest_HD.pdf.

[94] | According to statistics compiled by https://internetshutdowns.in/, as of June 3, 2021.

[95] | Niha Masih, Shams Irfan, and Joanna Slater, "India's Internet shutdown in Kashmir is the longest ever in a democracy," *The Washington Post*, December 16, 2019, https://www.washingtonpost.com/world/asia_pacific/indias-internet-shutdown-in-kashmir-is-now-the-longest-ever-in-a-democracy/2019/12/15/bb0693ea-1dfc-11ea-977a-15a6710ed6da_story.html.

[96] | Hannah Beech and Paul Mozur, "A Digital Firewall in Myanmar, Built With Guns and Wire Cutters," *The New York Times*, February 23, 2021, https://www.nytimes.com/2021/02/23/world/asia/myanmar-coup-firewall-internet-china.html.

[97] | Daniela Flamini, "The scary trend of internet shutdowns," *Poynter*, August 1, 2019, https://www.poynter.org/reporting-editing/2019/the-scary-trend-of-internet-shutdowns/.

[98] | "Latvia Bans Russia's RT TV Over EU-Sanctioned Chief," *Radio Free Europe/Radio Liberty*, July 1, 2020, https://www.rferl.org/a/latvia-bans-several-of-russian-rt-channels-over-eu-sanctioned-chief/30700356.html.

[99] | "The National Electronic Mass Media Council (NEMMC) has imposed a one-year ban on retransmitting and distributing channel 'Rossija RTR' in the territory of Latvia," Nacionālā elektronisko plašsaziņas līdzekļu padome (NEPLP), February 16, 2021, https://www.neplpadome.lv/en/home/news/news1/the-national-electronic-mass-media-council-(nemmc)-has-imposed-a-one-year-ban.html.

[100] | Lauren Chadwick, "Lithuania follows Latvia in banning Russian broadcaster RT," *Euronews*, July 9, 2020, https://www.euronews.com/2020/07/09/lithuania-follows-latvia-in-banning-russian-broadcaster-rt.

Sputnik-media.ee, making it a good example of how difficult it is to *really* ban a media in a democratic society where alternative routes to reentry are abundant. Different reasons were invoked for these bans: security reasons in the case of Ukraine; incitement to hatred, violence, and military conflict in the case of *Rossiya RTR* in Latvia; and the implementation of the European sanctions against Dmitry Kiselyov for the rest.[102] In any case, these choices have generally been criticized by human rights organizations, in particular Reporters without Borders (RSF).[103]

A less radical way of controlling the spread of disinformation is regulating the media, which, with digital platforms, are the main vectors of information manipulation. Such measures generally serve to strengthen the powers of media regulatory authorities such as the UK's Ofcom or France's Conseil supérieur de l'audiovisuel ("High Audiovisual Council") and require media to be transparent in their financial relations with foreign states, as with the Foreign Agent Registration Act in the United States. At the US Department of Justice's demand, *RT* and *Sputnik* registered as "foreign agents." Russia enacted a similar law, in a country where the press has been determined to be much less free than in Western democracies (as of June 3, 2021, Russia ranks 150 out of 180 in RSF's 2021 World Press Freedom Index)[104] in such a way that, in reality, foreign agent status can be used to force a certain number of media outlets to shut down.

Retaliating and deterring

Countries: the United States, France, Sweden

Is the best defense a good offense? Imposing a cost to the aggressor by retaliating is potentially the most efficient, but also the most sensitive, method. It implies being absolutely certain of the attribution: one does not want to attack an innocent actor by mistake. In case of a simple disinformation operation, attribution can be pretty straightforward. It is, for instance, when it comes from a "white" propaganda source like state media or officials – Russian or Chinese Ministry of Foreign Affairs spokespersons, for instance. Yet, in the case of a complex operation with a cyber dimension, such as hack-and-leak operations like the 2016 Democratic National Committee leaks in the United States or the 2017 Macron Leaks in France, attribution can be more of a challenge. It can also be difficult because an operation may be segmented between several different actors that do not appear to be obviously coordinated (e.g., the ones hacking may not be the ones leaking).

Retaliation is a difficult issue also because of a number of other reasons: states do not have the capacity to respond to all attacks, so they have to select which ones are worth the risk, including the risk of backfiring or escalation or the risk of losing the moral high ground. It can of course be tempting to go into the lion's den and fight the adversary with his own weapons. It is, however, also very risky. Liberal democracies should certainly not use the same methods than the ones they are denouncing, e.g., trolls, bots, fake personas, AI-generated profile pictures, doctored documents, etc. As mentioned in the author's 2018 CAPS-IRSEM report, "Clandestine operations, aiming for instance at manipulating the manipulators, are risky because, if exposed (and it is becoming increasingly difficult to prevent this in the long-run), they can jeopardize the very credibility of the source and invigorate conspiratorial actors—which would end up strengthening the very actors

[101] | "Estonia 'does not exclude the possibility' of banning *RT,*" *ERR News*, July 8, 2020, https://news.err.ee/1110560/estonia-does-not-exclude-the-possibility-of-banning-rt.

[102] | Kiselyov is the head of the media group Rossiya Segodnya, to which *Sputnik* belongs. According to Latvian authorities, and despite the fact *RT* does not formally belong to *Rossiya Segodnya*, Kiselyov also "ensures full control over RT." For more, see: "NEPLP: Kiselyov, who is on the EU sanction list, ensures full control over RT according to the decree signed by Putin", *NEPLP*, July 7, 2020, https://www.neplpadome.lv/en/home/news/news1/neplp-kiselyov,-who-is-on-the-eu-sanction-list,-ensures-full-control-over-rt.html.

[103] | "Baltic countries: Misusing EU sanctions to ban Russian TV channels is not a legitimate tool for promoting reliable information," Reporters without Bordres (RSF), July 10, 2020, https://rsf.org/en/news/baltic-countries-misusing-eu-sanctions-ban-russian-tv-channels-not-legitimate-tool-promoting; "Ukraine escalates 'information war' by banning three pro-Kremlin media," RSF, February 26, 2021, https://rsf.org/en/news/ukraine-escalates-information-war-banning-three-pro-kremlin-media.

[104] | "World Press Freedom Index," RSF, accessed June 3, 2021, https://rsf.org/en/ranking.

one aimed at undermining."[105] As Alina Polyakova, president and CEO of the Center for European Policy Analysis (CEPA), and Dan Fried, a fellow at the Atlantic Council, also wrote, "defense against disinformation has to be rooted in democratic principles and values: transparency, accountability, and respect for freedom of expression. We must not become them to fight them."[106]

That being said, there are other ways to retaliate. One is by conducting cyberoperations, like when the US Cyber Command blocked internet access to the Internet Research Agency (IRA) in St. Petersburg in 2018 and sent "direct messages to the operatives behind the influence campaigns"[107] to let them know that their identity was known and that they were under surveillance. In doing so, the objective was to "prevent Russian interference in the midterms."[108] It was a show of force, a signal sent to Moscow that Washington was aware of the IRA's activities and willing and able to impose a cost. However, it is difficult to assess the efficacy of such measures, and the very notion of deterrence in cyberspace is disputed, to say the least.[109]

Another preferred action in the US toolkit has been sanctions against Russian entities and officials. A number of other actors also imposed sanctions against Russia since 2014, including the EU, the United Kingdom, Norway, Canada, and Australia, but they have been in response to the Russo-Ukrainian war, to cyberattacks (the first time for the EU in 2020),[110] or, more recently, to the poisoning of Alexei Navalny. The United States seems to be the only country to use sanctions in response to election interference, including other country's elections, which is the closest thing to foreign information manipulation, and it does so with the explicit aim of deterring Russia and other potential adversaries, as the so-called DETER ("Defending Elections from Threats by Establishing Redlines") Act, a bill initially introduced in US Senate in 2018, illustrates.[111] In April 2021, the Biden Administration imposed sanctions on Russia for interfering in the 2020 US presidential election.[112] In October 2020, the US Department of Justice charged a Russian GRU officer named Anatoliy Sergeyevich Kovalev,[113] who was allegedly assigned to Unit 74455 (also known as the hacker group "Sandworm"), for "interference in the 2017 French elections."[114] As pointed out by some observers at the time, it is paradoxical that the United States not only attributed but also indicted for cyberattacks on the French elections before France itself. However, it should be recalled that France has, if not a policy, a custom of non-(public) attribution,[115] meaning that action may have been taken, just not publicly.

As retaliation can be clandestine, it is difficult to know for certain what states are doing in this

[105] | Vilmer et al., *Information Manipulation*, 172.

[106] | Alina Polyakova and Dan Fried, Democratic Offense Against Disinformation, CEPA and DFRLab, December 2020, 1, https://cepa.org/wp-content/uploads/2020/12/CEPA-Democratic-Offense-Disinformation-11.30.2020.pdf.

[107] | Julian E. Barnes, "US Begins First Cyberoperation Against Russia Aimed at Protecting Elections," *The New York Times*, October 23, 2018, https://www.nytimes.com/2018/10/23/us/politics/russian-hacking-usa-cyber-command.html.

[108] | Ellen Nakashima, "US Cyber Command operation disrupted Internet access of Russian troll factory on day of 2018 midterms," *The Washington Post*, February 27, 2019, https://www.washingtonpost.com/world/national-security/us-cyber-command-operation-disrupted-internet-access-of-russian-troll-factory-on-day-of-2018-midterms/2019/02/26/1827fc9e-36d6-11e9-af5b-b51b7ff322e9_story.html.

[109] | See, for instance: Joseph S. Nye Jr., "Deterrence and Dissuasion in Cyberspace," International Security, 41:3, 2016-2017, 44-71; Mariarosaria Taddeo, *How to Deter in Cyberspace*, Hybrid CoE, *Strategic Analysis* June-July 2018, https://mariarosariataddeo.files.wordpress.com/2018/07/pdf.pdf; Jonathan William Welburn, Justin Grana, and Karen Schwindt, *Cyber Deterrence or How we learned to stop worrying and love the Signal*, RAND Corporation, July 2019, https://www.rand.org/content/dam/rand/pubs/working_papers/WR1200/WR1294/RAND_WR1294.pdf; Brett Winterford, "Deterrence in cyberspace isn't working. What next?," risky.biz, April 21, 2020, https://risky.biz/solarium1/.

[110] | European Council, "EU imposes the first ever sanctions against cyber-attacks," Press release, July 30, 2020, https://www.consilium.europa.eu/en/press/press-releases/2020/07/30/eu-imposes-the-first-ever-sanctions-against-cyber-attacks/.

[111] | Defending Election from Threats by Establishing Redlines Act of 2018, S. 2313-115 (2018), https://www.congress.gov/bill/115th-congress/senate-bill/2313/text. See also: Ed Stein, "The Deter Act: Congress's Latest Effort to Discourage Election Interference," *Lawfare*, August 1, 2018, https://www.lawfareblog.com/deter-act-congresss-latest-effort-discourage-election-interference.

[112] | Executive Order on Blocking Property with Respect to Specified Harmful Foreign Activities of the Government of the Russian Federation, Exec. Order 14,024, 86 CFR 20249 (April 15, 2021), https://www.federalregister.gov/documents/2021/04/19/2021-08098/blocking-property-with-respect-to-specified-harmful-foreign-activities-of-the-government-of-the.

[113] | See his profile on the FBI's website: https://www.fbi.gov/wanted/cyber/anatoliy-sergeyevich-kovalev/@@download.pdf.

[114] | Alvaro Marañon, "Justice Department Charges Six Russian GRU Officers for Widespread Hacking Efforts," *Lawfare*, October 19, 2020, https://www.lawfareblog.com/justice-department-charges-six-russian-gru-officers-widespread-hacking-efforts.

[115] | Vilmer, The "Macron Leaks" Operation, 19.

field. What can be better known is what they are doing in terms of deterrence (dissuading another state from doing something it has not yet done) and compellence (coercing the state to stop doing what it is already doing) because, most of the time, it involves public declarations. Sanctions can be either retaliatory or deterring ex ante, as the threat of sanctions itself can often be a means of discouraging an activity.

Deterrence can also be obtained through public statements and diplomacy. For instance, in January 2017, the French defense minister, aware that the presidential campaign was under attack, declared that "France reserves the right to retaliate by any means it deems appropriate. This could be through our cyber arsenal but also by conventional armed means."[116] A similar message was conveyed privately by the minister to his Russian counterpart and by then-President François Hollande to Russian President Vladimir Putin. Democrats in the US Senate, drawing lessons from the French elections in their January 2018 report for the Foreign Relations Committee, concluded that "direct diplomatic engagement clearly pointing to malicious actors and the consequences of their actions can act as a deterrent."[117] "Deterrent" may be too strong a word, as these precautions obviously were not enough to deter the attackers behind the Macron Leaks, but, given the amateurism of the attack, it can safely be assumed that the foreign power behind it exercised restraint in the face of the hard stance taken by the French authorities.

In that sense, there is a link between transparency and deterrence. A few countries, such as the United Kingdom, Canada, or Sweden, often share details of what they do to counter foreign disinformation or influence. They publish national doctrines and regular reports, they acknowledge who is doing what within the government and how they are organized, they have dedicated websites, and the topic is frequently covered by politicians and public officials in speeches and interviews. That is for several reasons: not only because the same countries also defend a rights-based approach (according to which it is important to be transparent to their population and parliament about what methods they use, so everyone can evaluate to what extent they respect fundamental rights like privacy, freedom of speech, and freedom of the press); and because it is a means of raising awareness among the population. It is also about signaling to potential adversaries the high degree of preparedness and determination of the society. Being transparent is explicitely a "part of the Swedish counterstrategy – an example of deterrence": the objective is "to deter actors from contemplating interference in the Swedish elections"[118] but also more generally in the democratic life of the country.

◆

[116] | "Jean-Yves Le Drian: 'Face à une cyberattaque, la France peut riposter par tous les moyens,'" *Le Journal du Dimanche*, January 8, 2017, https://www.lejdd.fr/Politique/Jean-Yves-Le-Drian-sur-le-cyberespionnage-Le-risque-sur-la-vie-democratique-est-reel-838111.

[117] | Comm. on Foreign Relations, US Senate, *Putin's Asymmetric Assault*, 125.

[118] | "Jean-Yves Le Drian: 'Face à une cyberattaque, la France peut riposter par tous les moyens,'" Le Journal du Dimanche, January 8, 2017, https://www.lejdd.fr/Politique/Jean-Yves-Le-Drian-sur-le-cyberespionnage-Le-risque-sur-la-vie-democratique-est-reel-838111.

International cooperation

As the threats raised by information manipulation are transnational in nature, not only because an attack may come from another country but also, even more so, because the internet has no borders, such attacks are a global challenge for the international community, a challenge requiring coordinated responses. International cooperation is therefore vital. It has continued to grow, especially since 2014. There are several layers to this cooperation, e.g., bilateral cooperation, mostly in the form of intelligence sharing, and multilateral formats (EU, NATO, the Group of Seven [G7]).

Bilateral cooperation

The first layer, which has always existed, is sharing intelligence, but this has remained limited, in this area as in the cyber domain, since sharing information also means sharing vulnerabilities. States are thus often reticent to do so. Apart from a few exceptions, including the Five Eyes (an alliance of the intelligence services of the United States, United Kingdom, Canada, Australia, and New Zealand), this type of cooperation is usually bilateral, between two countries that trust each other. It is difficult to open such arrangements up to more countries, and the level of information shared (what you say) depends on the level of trust in your partner (who you say it to). For example, countries hesitate to share too much with agencies they feel their adversary has too big a foothold in or states whose political leadership is deemed too ambivalent with regard to that adversary. And, obviously, the more partners there are, the stronger the reluctance is: the risk that, among the twenty-seven EU countries, there may be a couple of Russian Trojan horses, for instance, is precisely what discourages many actors from sharing sensitive information within the EU Intelligence and Situation Centre (EU INTCEN).

Nevertheless, at the bilateral level, especially between longtime allies, intelligence sharing works well and is indeed an important lever in fighting information manipulation. During the 2017 presidential campaign, France benefited from operational cooperation with the US authorities. France's then-Defense Minister Jean-Yves Le Drian acknowledged that "our services have the necessary exchanges on this subject, if only to draw lessons for the future."[119] Admiral Michael S. Rogers, the US National Security Agency director, told the US Congress in May 2017, "if you take a look at the French election [...] we had become aware of Russian activity. We had talked to our French counterparts prior to the public announcements of the events publicly attributed this past weekend and gave them a heads-up: 'Look, we're watching the Russians, we're seeing them penetrate some of your infrastructure.'"[120]

Bilateral cooperation to counter foreign information manipulation is not limited to intelligence sharing. The relevant national teams mentioned above, in various departments and agencies, are usually well-connected to their counterparts in allied nations. They share ideas and good practices, they develop joint projects, and they sometimes exchange personnel. Some countries even have dedicated services to that effect. For example, the UK Government Communication Service (GCS) has a dedicated branch, GCS International, precisely to work with foreign governments,[121] mainly to help them build their communications capability but also to work with "peer" counterparts to design joint campaigns (e.g., to fight COVID-19 vaccination hesitancy) and establish research partnerships.

Multilateral formats

The second layer involves multilateral formats, with those most active in the fight against information manipulation being the EU, NATO, and the G7.

[119] | "Jean-Yves Le Drian: 'Face à une cyberattaque,'" *Le Journal du Dimanche*.
[120] | "Admiral Rogers Says Intel Community Warned of Russian Hacking Ahead of Macron Leak," *C-SPAN*, May 9, 2017, https://www.c-span.org/video/?c4668917/admiral-rogers-intel-community-warned-russian-hacking-ahead-macron-leak.
[121] | UK Government Communication Service, "What we do," https://gcs.civilservice.gov.uk/about-us/what-we-do/.

European Union

The EU's interest in the matter was, at first, a reaction to Russian disinformation produced in 2014 to justify the annexation of Crimea and deny the invasion of Ukraine and the shoot down of Flight MH17. As a consequence, the European Council on March 19-20, 2015, stressed "the need to challenge Russia's ongoing disinformation campaigns" and invited the High Representative to prepare an action plan on strategic communication. Then, the European External Action Service's (EEAS) Strategic Communication Division created three task forces: an "East" StratCom Task Force active since September 2015, which has sixteen staff currently, focused on Russia, and which shares its work on a dedicated website "EU vs Disinformation" (EUvsDisinfo.eu), in a weekly *Disinformation Review*, and on social networks under the name EU Mythbusters; a "South" StratCom Task Force active since 2017 with six staff currently that combats Jihadist rhetoric; and a "Western Balkans" Task Force also active since 2017, with seven staff currently, that focuses on defending the EU's image in the Balkan region. On top of that, a horizontal team was later created "with a focus on emerging threats, data analysis, policy development and international cooperation, including the EU's Rapid Alert System on Disinformation,"[122] as detailed below. In February 2020, the overall Division for Strategic Communications and Information Analysis was composed of about thirty-five people.[123] In practice, the East Task Force is the only one fighting disinformation, which means that EEAS's efforts in that field are mostly, if not exclusively, focused on Russia.

In 2016, the European Commission adopted a joint framework on countering hybrid threats[124] and created a Hybrid Fusion Cell within the EU INTCEN of the EEAS. Between 2015 and 2020, the EU at large (the Council of the EU; the European Parliament; the Directorate-General for Communications Networks, Content and Technology; the Directorate-General for Communications; EEAS) produced over seventy public documents dealing with disinformation and foreign influence.[125] Russian interference in the 2016 US presidential campaign was a turning point in EU's awareness – paradoxically more than European democratic processes that were also the target of Russian disinformation operations like the Dutch Ukraine-European Union Association Agreement referendum in March 2016 and the Brexit referendum in June 2016.

Following a resolution by the European Parliament in June 2017 requesting that the Commission study the possibility of "legislative intervention to limit the dissemination and spreading of fake content,"[126] the Commissioner for Digital Economy and Society formed a group of experts that issued a report in March 2018 containing a certain number of recommendations. Followed by public hearings, the report served as the basis for the communication on tackling online disinformation the Commission published the next month. The communication proposed a "Code of Practice on Disinformation," (CPD) which was published in July 2018 and which constitutes "the first time worldwide that industry has agreed, on a voluntary basis, to self-regulatory standards to fight disinformation."[127] In October, the CPD was signed by Facebook, Twitter, Google, Mozilla, and several professional associations, all of whom committed to step up their efforts to tackle online disinformation, particularly by taking down fake accounts and limiting the visibility of sites that promote disinformation. In May 2019, Microsoft also signed the Code, and TikTok joined in June 2020. In October 2019, the signing parties published the first self-assessment reports on the implementation

[122] | European External Action Service, European Union, "Questions and Answers about the East StratCom Task Force," April 28, 2021, https://eeas.europa.eu/headquarters/headquarters-homepage/2116/-questions-and-answers-about-the-east-.

[123] | "A diverse toolbox: the instruments of disinformation, Online," *EU Monitor*, February 20, 2020, https://www.eumonitor.eu/9353000/1/j9vvik7m1c3gyxp/vl4kf0ennfxy?ctx=vg9pj7ufwbwe&start_tab1=5.

[124] | *Joint framework on countering hybrid threats*, European Commission, April 6, 2016, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016JC0018.

[125] | I would like to thank Jakub Kalenský for providing this information. See his to-be-published report analyzing these documents.

[126] | Online platforms and the Digital Single Market, European Parliament, P. Res. 2016/2276(INI) (June 15, 2017), https://www.europarl.europa.eu/doceo/document/TA-8-2017-0272_EN.html.

[127] | European Commission, *Code of Practice on Disinformation*, https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation.

of their commitments. This procedure does push these platforms to be more transparent with regard to the measures taken both quantitatively and qualitatively and highlights a real effort on their part that also has certain failings, particularly in terms of measures to be taken to "empower consumers and the research community."[128] Moreover, the intrinsic limitation of the CPD is of course its voluntary nature, i.e., the fact that it is nothing but a self-regulation.[129] That is why, in May 2021, the European Commission published an additional *Guidance on Strengthening the Code of Practice on Disinformation* setting up "how the signatories (platforms and other relevant stakeholders) should strengthen the Code of Practice and how the Code implementation and impact should be monitored."[130] One of the measures proposed is the creation of a permanent task force chaired by the Commission.[131]

In the lead-up to the 2019 European elections, the EU took additional measures to mitigate the risk of information manipulation designed to interfere in the electoral process, including, notably, the "elections package" announced by EU President Jean-Claude Juncker in his 2018 State of the Union address. In December 2018, the European Commission released an Action Plan against Disinformation, which remains the most detailed document it has produced on the topic. Mentioning that "[a]ccording to the EU Hybrid Fusion Cell, disinformation by the Russian Federation poses the greatest threat to the EU," the Plan is based on four pillars: "improving the capabilities of Union institutions to detect, analyse and expose disinformation; strengthening coordinated and joint responses to disinformation; mobilising private sector to tackle disinformation; raising awareness and improving societal resilience."[132] Following the Plan, a Rapid Alert System was set up among EU institutions and member states in March 2019. Similar to the G7's RRM (described below) and set up approximately at the same time, the Rapid Alert System's objective is to facilitate and accelerate information sharing, allowing faster and better coordinated responses. In concrete terms, it is made of a dedicated digital platform and a network of national contact points.[133]

During the 2019 European elections campaign, there was "evidence of coordinated inauthentic behavior aimed at spreading divisive material on online platforms, including through the use of bot and fake accounts."[134] However, the measures taken by civil society, especially journalists and fact-checkers, digital platforms, national authorities, and, finally, European institutions, kept such attempts below the nuisance threshold and the elections went well overall.

Another turn for Europe was the COVID-19 pandemic in 2020. The European commission issued multiple statements on fighting disinformation related to the coronavirus, including a report on *Tackling COVID-19 disinformation – Getting the facts right* published in June 2020,[135] with the EEAS providing an update in December.[136] The European Commission also published monthly reports from

[128] | European Commission, *Code of Practice on Disinformation one year on: online platforms submit self-assessment reports*, October 29, 2019, https://ec.europa.eu/commission/presscorner/detail/en/statement_19_6166.

[129] | Ethan Shattock, "Self-regulation 2:0? A critical reflection of the European fight against disinformation," Misinformation Review, Harvard Kennedy School, May 31, 2021, https://misinforeview.hks.harvard.edu/article/self-regulation-20-a-critical-reflection-of-the-european-fight-against-disinformation/.

[130] | European Commission, *Guidance on strengthening the Code of Practice on Disinformation*, March 31, 2021, https://op.europa.eu/en/publication-detail/-/publication/c1a4094a-921a-11eb-b85c-01aa75ed71a1/language-en.

[131] | European Commission, *Guidance on Strengthening the Code of Practice on Disinformation*, May 2021, https://digital-strategy.ec.europa.eu/en/library/guidance-strengthening-code-practice-disinformation.

[132] | European Commission, *Action Plan against Disinformation*, December 5, 2018, 4-5, https://ec.europa.eu/info/sites/info/files/eu-communication-disinformation-euco-05122018_en.pdf.

[133] | EEAS, "Rapid Alert System: Strengthening Coordinated and Joint Responses to Disinformation," March 2019, https://eeas.europa.eu/sites/default/files/ras_factsheet_march_2019_0.pdf.

[134] | A Europe that protects: EU reports on progress in fighting disinformation ahead of European Council," Press Corner, June 14, 2019, https://ec.europa.eu/commission/presscorner/detail/en/MEX_19_3011.

[135] | High Representative of the Union for Foreign Affairs and Security Policy, European Commission, *Tackling COVID-19 disinformation – Getting the facts right*, Joint communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the regions, June 10, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020JC0008&from=EN.

[136] | "EEAS special report update: short assessment of narratives and disinformation around the Covid-19 pandemic (update May-November)," EUvsDisinfo, December 2, 2020, https://euvsdisinfo.eu/eeas-special-report-update-short-assessment-of-narratives-and-disinformation-around-the-covid-19-pandemic-update-may-november/.

digital platforms (Twitter, TikTok, Google, and Microsoft) on measures they took against COVID-related disinformation as signatories of the Code of Practice on Disinformation;[137] and it funds a number of research projects on COVID-related mis- or disinformation.[138] Overall, the "infodemic" of 2020 made Europe aware, for the first time, that not only Russia but also China was a serious threat in terms of disinformation, indicating a need for Europe to reorganize its efforts.

NATO

NATO observed "a significant increase in disinformation and propaganda since Russia illegally annexed Crimea, Ukraine, in 2014,"[139] and it has to deal on a regular basis with a large range of fake or biased news about the organization. Those attacks – at least the publicly available ones[140] - mainly target its intentions, presented as conquest-driven and aggressive; its expansion and the volume of its troops, particularly those deployed in the Baltic countries and Poland; or even crimes its soldiers allegedly commit.

In the 2018 Brussels Declaration, allied heads of state and government acknowledged they were facing "hybrid challenges, including disinformation campaigns and malicious cyber activities"[141] and, in the 2019 London Declaration, they committed to strengthen NATO's "ability to prepare for, deter, and defend against hybrid tactics that seek to undermine our security and societies."[142] The organization presents its approach to counter disinformation as "a twin-track model," based on two functions: on the one hand, "Understand" the information environment; for example, its evolution during the COVID-19 pandemic, which has been used in anti-NATO narratives (such as "NATO troops are bringing the virus to Baltic states, pursuing large-scale exercises spreading the virus, and therefore putting local civilians at risk" or "NATO created the virus in secret US/NATO laboratories"). On the other hand, "Engage" by providing fact-based communication and exposing disinformation, including in the Russian language, as well as by funding research ("independent NGOs, think tanks, academics, fact-checking organizations and other civil society initiatives to promote debate and to build resilience").[143]

Two major units are in charge of responding to such attacks and, more broadly, studying instances of information manipulation. On the one hand, the international secretariat's Public Diplomacy Division (PDD) in Brussels, which regularly issues rebuttals of Russian accusations, particularly on a dedicated page of its website that responds to the main "myths." The PDD also plays a coordination role between NATO's various units and between the Allies. On the other hand, NATO's Centre of Excellence on Strategic Communication in Riga, created in 2014, publishes a large number of analyses. It also holds an annual "StratCom Summit" on such issues, which is one of the most important occasions for sharing and exchanging information between state and non-state actors from across the world. Additionally, another unit, the Joint Intelligence and Security Division (JISD), also contributes to identifying and countering hybrid operations, which may include disinformation.

[137] | See, for example: European Commission, *Latest set of reports and the way forward - Fighting COVID-19 Disinformation Monitoring Programme*, March 8, 2021, https://digital-strategy.ec.europa.eu/en/library/latest-set-reports-and-way-forward-fighting-covid-19-disinformation-monitoring-programme; European Commission, "Coronavirus vaccine disinformation: new reports from online platforms to inform Code of Practice revamp," Press release, March 26, 2021, https://digital-strategy.ec.europa.eu/en/news/coronavirus-vaccine-disinformation-new-reports-online-platforms-inform-code-practice-revamp.

[138] | European Commission, *Funded projects in the fight against disinformation*, https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/fighting-disinformation/funded-projects-fight-against-disinformation_en.

[139] | NATO, *NATO's approach to countering disinformation: a focus on COVID-19*, July 17, 2020, https://www.nato.int/cps/en/natohq/177273.htm.

[140] | See, for example: Dan Sabbagh, "Russia-aligned hackers running anti-Nato fake news campaign – report," *The Guardian*, July 30, 2020, https://www.theguardian.com/technology/2020/jul/30/russia-aligned-hackers-running-anti-nato-fake-news-campaign-report-poland-lithuania; *'Ghostwriter' Influence Campaign: Unknown Actors Leverage Website Compromises and Fabricated Content to Push Narratives Aligned with Russian Security Interests*, FireEye/Mandiant Solutions, 2020, https://www.fireeye.com/content/dam/fireeye-www/blog/pdfs/Ghostwriter-Influence-Campaign.pdf.

[141] | NATO, *Brussels Summit Declaration*, issued by the Heads of State and Government (Brussels: NATO, July 11-12, 2018), https://www.nato.int/cps/en/natohq/official_texts_156624.htm.

[142] | NATO, *London Declaration*, issued by the Heads of State and Government (London: NATO, December 3-4, 2019), https://www.nato.int/cps/en/natohq/official_texts_171584.htm.

[143] | NATO, *NATO's approach to countering disinformation*.

G7

The G7 – only a Group of Seven (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States) since it suspended Russia's membership in the group in 2014 in response to latter's annexation of Crimea – is also, perhaps in part for that reason, very much preoccupied by the growing weaponization of information since 2014. In the *Charlevoix Commitment on Defending Democracy from Foreign Threats*, taken at the June 2018 G7 Summit in Charlevoix (Canada), the leaders of the G7 committed to "[e]stablish a G7 Rapid Response Mechanism [RRM] to strengthen our coordination to identify and respond to diverse and evolving threats to our democracies, including through sharing information and analysis, and identifying opportunities for coordinated response."[144] The RRM has since been established, coordinated by Canada, and allows information to be rapidly transmitted between G7 countries. The focal points for each country know each other and meet on a regular basis, and the coordination unit of Global Affairs Canada collects and shares a large amount of relevant information, particularly publications and events.

Other forums for cooperation

Among other forums for cooperation, and within the European framework, the European Centre of Excellence for Countering Hybrid Threats in Helsinki (Hybrid CoE) is also notable. Created in 2017, it is a resource hub in this area, running networks of researchers, holding seminars, and issuing publications on a regular basis. Among other nongovernmental initiatives, the Transatlantic Commission on Election Integrity, founded in 2018 by the Alliance of Democracies Foundation, also constitutes a useful bridge between the United States and Europe focused on the issue of foreign electoral interference.

Civil society

Whatever the measures described in the preceding pages and implemented by countries, a society's degree of resilience, its capacity to resist information manipulation, depends first and foremost on the mobilization of its citizens, "because not everyone trusts what governments say. We need other opinion leaders to act as trusted messengers on these issues to their own audiences, which government often cannot reach."[145] Hence the importance of at least three categories of actors: "hunters" (journalists, fact-checkers, researchers); norm entrepreneurs; and digital platforms.

"Hunters": journalists, fact-checkers, researchers

Members of civil society, particularly journalists, are on the front lines. Besides the fact that they all have the responsibility to provide reliable and accurate information, some are working to improve journalism standards, have specialized in hunting down trolls and other influence networks, or – like Bellingcat, a model of its own kind – focus on investigative journalism and open-source tools.

Fact-checkers are also important, even though their efficacy is limited for a number of reasons that will be mentioned in the next part, including the fact that fact-checks rarely achieve the same reach as the rumor they are checking. Fact-checkers are often journalists, but not necessarily. Not all fact-checking organizations are new (one of the best known, Snopes in the United States, has been around since 1994) but their prevalence has shot up everywhere in the world since the 2000s, another indication of the growing awareness at work. Some fact-checking units are operated by well-known media, which, in addition to producing information, are increasingly taking care of checking it (AFP's Fact Check, Reality Check at the BBC, Decodex at *Le Monde*, etc.). But there are not just respectable

[144] | Government of Canada, Charlevoix commitment on defending democracy from foreign threats, June 5, 2018, https://www.international.gc.ca/world-monde/international_relations-relations_internationales/g7/documents/2018-06-09-defending_democracy-defense_democratie.aspx?lang=eng.
[145] | Kalenský, "Russian Disinformation Attacks on Elections."

media involved: many websites spread fake or biased news while pretending to do fact-checking. For example, as in Turkey, government-connected "fact checkers" often distort or ignore facts to present pro-government spin or disinformation.[146] For that reason, in 2015, the Poynter Institute created an International Fact Checking Network and adopted a "code" of common principles for guaranteeing transparent nonpartisan verification. The fact-checkers that Poynter "approves" are supposed to meet those methodological criteria. Such a label could also help them get social traction, while pointing journalists and other writers toward reliable information sources.

Some initiatives aim at coordinating those various efforts and actors, like the non-profit coalition First Draft, created in 2015 by nine organizations[147] and which has expanded since to become a major network linking media, researchers, and civil society organizations; the organization has headquarters in London and offices in New York and Sydney. Having the mission "to protect communities from harmful misinformation," First Draft works "to empower society with the knowledge, understanding, and tools needed to outsmart false and misleading information."[148] That may prove to be a difficult objective as "society" does not always care (or rather, its willingness of being "empowered" depends on a number of factors, including education and political context).

Researchers at think tanks are also an essential piece of that puzzle. First, because they build bridges: think tanks are non-state actors but are often at least partly state-funded and have on their teams members who either formerly worked in national ministries, agencies, or services in the defense and security areas or continue as staff there while being seconded to the "outside." In the gray area between two worlds, they have both access to resources and information and the ability to dispatch them more freely. They regularly hold more or less closed workshops, including "tracks 1.5," i.e., meetings involving both officials and civil society. This cross-fertilization allows officials to "get a bit of fresh air," gather ideas from "outside the box," and members of civil society to better understand how states work and hope to perhaps be able to influence them. Of the better-known meetings of this type, mention can be made of those organized by the Atlantic Council's DFRLab, by the Czech think tank European Values in Prague (Stratcom Summits), and by the S. Rajaratnam School of International Studies' (RSIS) Centre of Excellence for National Security in Singapore, which has the advantage of bringing together differing geographic perspectives, in particular Euroatlantic and Asian viewpoints.

Second, a couple of key think tanks and research centers, most of them in the United States (Atlantic Council's DFRLab, Brookings, the German Marshall Fund, the Center for European Policy Analysis, Carnegie Endowment for Peace's Partnership on Countering Influence Operations, etc.), but also in Europe (EU DisinfoLab, European Values, Oxford University's Computational Propaganda Project), Asia (RSIS), and Australia (Australia Strategic Policy Institute [ASPI]), produce a substantial portion of "operationalizable" research on information manipulation. To that list should be added responsible private sector actors[149] contributing to the research innovation, such as Graphika, which also contribute by identifying and revealing information manipulation in frequent and thorough investigations and reports. Those think tanks, research centers, and responsible private sector actors also organize regular meetings with political decision-makers and also digital platforms, which are increasingly taking part in such activities. An example would be the Brookings High-Level Transatlantic Working Group on Disinformation and Emerging Technology (to which the author was a member): composed of experts, researchers, government officials, and representatives from

[146] | Efe Kerem Sözeri, "These fake 'fact-checkers' are peddling lies about genocide and censorship in Turkey," *Poynter*, May 31, 2017, https://www.poynter.org/fact-checking/2017/these-fake-fact-checkers-are-peddling-lies-about-genocide-and-censorship-in-turkey/.

[147] | Bellingcat, Dig Deeper, Emergent, Eyewitness Media Hub, Google News Initiative, Meedan, reported.ly, Storyful, and Verification junkie.

[148] | First Draft News, *About*, accessed June 4, 2021, https://firstdraftnews.org/about/.

[149] | More generally, information defense is perceived by the private sector as an opportunity to make potential profit. Several private security companies, like Symantec, offer their expertise to political parties and campaigns to bolster electoral security, for instance.

Twitter, Facebook, and Google, it met a couple of times in 2019-2020.

Research conducted in universities is also valuable and contributes to the first line of defense (to use Kalenský's terminology), which is "documenting the threat." An increasing number of books and articles are being published in all languages on this issue, and more and more conferences are being organized all over the planet. Information manipulation has become a trendy, cross-disciplinary research field. Teaching is also extremely important, as it contributes directly to raising awareness, and courses on "fake news" or mis- and disinformation are increasingly common in colleges and universities.[150] Some university libraries also set up helpful webpages with teaching resources.[151]

Mention should also be made of grassroots initiatives like the so-called "elf" movement, which started during the 2014 Russo-Ukrainian war in response to growing Russian propaganda. It originated in Lithuania,[152] quickly spread in the Baltic states, and now has chapters in many countries, mainly in central, eastern, and northern Europe. Elves are debunkers of false or biaised information but, above all, they are hunters: they hunt down, identify, expose, and "blame and shame" pro-Kremlin trolls. For example, the Swedish Facebook group #Jagärhär (#Iamhere) gathers about 75,000 people, aiming "to do what government and social media companies have failed to do: defend people being attacked online by trolls and push back against the spread of misinformation."[153] Similar groups exist in other countries, including Czechia and Slovakia. They largely function independently, but their members occasionally share good practices.

Overall, detecting and exposing the attackers – naming and shaming them – is one of the most powerful and efficient methods, not only because it can have a deterrent effect, but also because it can help capture the public attention by turning the attack itself into a story. In the case of the 2017 Macron Leaks, "a handful of open-source researchers, by their reactivity and the quality of their analysis, helped to derail the attackers' narrative."[154] Ben Nimmo, then a Senior Fellow of the Atlantic Council's DFRLab, was one of them. As Jenni Sargent, managing director of First Draft, said: "It doesn't matter how much money you throw at the problem, or how many technological advances you have. Without the human layer of someone like Ben [Nimmo] dissecting the way that people use the internet, then we wouldn't be as far ahead as we are in terms of understanding the problem and the scale."[155]

Of course, it does not mean that technological advances are not useful. Camille François, Chief Innovation Officer at the social-media monitoring company Graphika, leads the company's efforts to detect and expose information manipulation by using machine learning to map out online communities. Machine-learning algorithms, like data science and big data visualization, are useful to identify and map the spread of disinformation, while artificial intelligence can help detect manipulated content.[156]

Norm entrepreneurs

There are also normative initiatives – labels, ratings, rankings – to make it possible to differentiate reliable sources of information from unreliable ones. Some have advocated "negative" rankings, similar to

[150] One such example is a course on "Information operations and influence in the digital age" at Georgetown University, taught by Olga Belogolova, Lead analyst and policy manager for influence operations at Facebook.

[151] See, for instance: University Libraries, Temple University, "'Fake news,' Misinformation, & Disinformation," updated January 21, 2021, https://guides.temple.edu/fakenews/teaching; HACC Library, *Fake News: Faculty Resources*, updated May 19, 2021, https://libguides.hacc.edu/fake-news/faculty-resources.

[152] Michael Weiss, "The Baltic Elves Taking on Pro-Russian Trolls," *The Daily Beast*, March 20, 2016, https://www.thedailybeast.com/the-baltic-elves-taking-on-pro-russian-trolls.

[153] Makana Eyre and Martin Goillandeau, "Here, here: the Swedish online love army who take on the trolls," *The Guardian*, January 15, 2019, https://www.theguardian.com/world/2019/jan/15/the-swedish-online-love-army-who-battle-below-the-line-comments.

[154] Vilmer, *The "Macron Leaks" Operation*, 39.

[155] Satariano, "He Combs the Web for Russian Bots."

[156] United Nations Interregional Crime and Justice Research Institute (UNICRI), *Stop the Virus of Disinformation: the risk of malicious use of social media during COVID-19 and the technology options to fight it*, November 2020, 18 20, http://www.unicri.it/sites/default/files/2020-11/SM%20misuse.pdf.

Freedom House's freedom of the press rating and Transparency International's corruption index but for disinformation websites.[157] There are already "positive" rankings or certifications, such as RSF's Journalism Trust Initiative, aimed at "reversing this trend by giving a real advantage to all those who reliably produce news and information, whatever their status,"[158] in a view to having digital platforms assign a "bonus" to that quality in their algorithms, thereby giving such sources increased visibility.

A number of organizations rate or list websites. The Global Disinformation Index, a not-for-profit organization based in the United Kingdom, provides disinformation risk ratings. NewsGuard, a US startup, uses journalists to rate the reliability of news websites. It provides "trust ratings" on thousands of websites and developed artificial intelligence (AI) tools like "Misinformation Fingerprints" to catalog hoaxes spread online. It provides various services to journalists, analysists, researchers, and government. There are also regional initiatives, such as the website konspiratori.sk, a Slovakian initiative by private companies, journalists, and academics, providing a "list of websites with controversial content."[159]

Another important actor in that field is the Forum on Information and Democracy, founded in 2019 by eleven NGOs and research centers, including RSF (Christophe Deloire, Secretary General of RSF, is the chair of the Forum's Board of Directors). In June 2020, it created a Working Group on infodemics that published a report the following November with 250 recommendations in four categories: transparency of platforms; meta-regulation of content moderation; platform design and reliability of information; and mixed private and public spaces and closed messaging services.[160]

Digital platforms

Under the pressure of both states and civil society, social media platforms themselves, which had long hesitated to acknowledge the problem, have been forced to react. As Polyakova and Fried wrote, the platforms "have moved from an initial and unsustainable denial of the problem to a stance of willingness to help deal with it, though the depth of this commitment (and the effectiveness of the responses) has yet to be determined."[161] The turning point in cooperation was in 2018, the year the platforms began to share information (Reddit in April about 944 accounts linked to the IRA, Facebook in July in regards to another IRA operation, Twitter in October about 9 million tweets also attributed to the IRA, etc.).[162]

Since then, they have published educational documents, such as Google's *How Google Fights Disinformation* in February 2019, and are now accountable to the European Commission as part of the Code of Practice on Disinformation. The Founder and Chief Executive Officer (CEO) of Facebook, Mark Zuckerberg, also faced questions from the EU Parliament in May 2018, including uncomfortable ones on the Cambridge Analytica scandal[163] and the company's role in spreading disinformation.[164] He much more regularly testifies before the US Congress, including four times in 2020 alone; also in 2020, the CEOs of Twitter and Google spoke before the US Congress three times each.[165] Not all of the hearings to date have

[157] | Peter Pomerantsev and Michael Weiss, *The Menace of Unreality: How the Kremlin Weaponizes Information, Culture and Money, The Interpreter*, 2014, 40, https://imrussia.org/media/pdf/Research/Michael_Weiss_and_Peter_Pomerantsev__The_Menace_of_Unreality.pdf.

[158] | Christophe Deloire, quoted by François Bougon in, "Un label pour redonner confiance dans le journalisme," *Le Monde*, April 3, 2018, https://www.lemonde.fr/actualite-medias/article/2018/04/03/un-label-pour-redonner-confiance-dans-le-journalisme_5280004_3236.html.

[159] | "List of websites with controversial content," Konspiratori.sk, accessed July 12, 2021, https://www.konspiratori.sk/zoznam-stranok/en.

[160] | Forum on Information & Democracy, *Working Group on Infodemics: Policy Framework*, November 2020, https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID_Report-on-infodemics_101120.pdf.

[161] | Alina Polyakova and Daniel Fried, *Democratic Defense Against Disinformation 2.0*, Atlantic Council, June 2019, 1, https://www.brookings.edu/research/democratic-defense-against-disinformation-2-0/.

[162] | Ben Nimmo, "Investigative Standards for Analyzing Information Operations," draft presented at the Brookings High-Level Transatlantic Working Group on Disinformation and Emerging Technology, February 19-21, 2020, forthcoming.

[163] | Nicholas Confessore, "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far," *The New York Times*, April 4, 2018, https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.

[164] | See the follow-up questions here: https://www.europarl.europa.eu/resources/library/media/20180524RES04208/20180524RES04208.pdf.

[165] | Cristiano Lima, "Facebook, YouTube, Twitter execs to testify at Senate hearing on algorithms," *Politico*, April 23, 2021, https://www.politico.com/news/2021/04/23/facebook-youtube-twitter-testify-senate-484456.

focused on counter-disinformation – the one held in April 2021, for instance, was about algorithmic transparency[166] – but this is still a recurrent issue.

These platforms now invest significant resources into detecting and eliminating information manipulation on their sites. At the Munich Security Conference in February 2020, Zuckerberg mentioned that at Facebook 35,000 people scan the platform for problematic content (although he did not specify that most of those content moderators are contractors and that many suffer from mental health issues because of the disturbing posts they have to monitor all day,[167] while receiving "little support" from the company[168]). In Munich, Zuckerberg added that, with the assistance of AI, more than a million fake accounts are taken down every day.[169] Platforms also have dedicated units, like Google's Jigsaw, a technology incubator developing tools to detect and counter manipulated media (for example, an application programming interface [API] called Perspective "using machine learning to reduce toxicity online,"[170] and a service called Project Shield defending "news, human rights and election monitoring sites from DDoS attacks"[171]). Jigsaw also supports research, such as sponsoring the report Improving Machine Learning to Detect and Understand Online Conspiracy Theories, published by the RAND Corporation in April 2021.[172] Social media platforms also set up ad hoc teams or "war rooms" to monitor certain events, as Facebook did ahead of the European Parliament elections of May 2019 and ahead of the Taiwanese general election of January 2020, for instance.

Last but not least, social media platforms not only remove but also publicly disclose inauthentic activity and other activity that violates platform rules that is detected on pages, accounts, groups, or events, before taking them down, especially if it relates to state actors or foreign influence. Facebook regularly explains what it calls "coordinated inauthentic behavior" (CIB)[173] and, since March 2020, has published monthly CIB reports.[174] In May 2021, in a Threat Report on *The State of Influence Operations 2017-2020*, it looked back at more than 150 covert influence operations, demonstrating a CIB, from over fifty countries worldwide.[175] Twitter also regularly discloses "networks of state-linked information operations" on their blog,[176] as well as Google (its Threat Analysis Group documents "government-backed attacks" in a quarterly bulletin and regular posts)[177] and, on the cybersecurity side, Microsoft (which regularly detects and exposes threat actors, especially state-sponsored ones).[178] Some of those platforms also collaborate with certain researchers

[166] | Kate Kaye, "Cheat Sheet: Senators want more transparency into 'addictive' Facebook, Twitter and YouTube algorithms," *DigiDay*, April 28, 2021, https://digiday.com/media/cheat-sheet-senators-want-more-transparency-into-addictive-facebook-twitter-and-youtube-algorithms/.

[167] | Casey Newton, "'The Trauma Floor': The secret lives of Facebook moderators in America," *The Verge,* February 25, 2019, https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona; Casey Newton, "Bodies in Seat: At Facebook's worst-performing content moderation site in North America, one contractor has died, and others say they fear for their lives," *The Verge*, June 19, 2019, https://www.theverge.com/2019/6/19/18681845/facebook-moderator-interviews-video-trauma-ptsd-cognizant-tampa.

[168] | Haley Messenger and Keir Simmons, "Facebook content moderators say they receive little support, despite company promises," *NBC News*, May 11, 2021, https://www.nbcnews.com/business/business-news/facebook-content-moderators-say-they-receive-little-support-despite-company-n1266891.

[169] | Mark Zuckerberg, quoted by Munich Security Conference (@MunSecConf), "'We take down now more than a million fake accounts a day across our network,' @Facebook CEO Mark #Zuckerberg says at #MSC2020. 'The vast majority are not connected to state actors trying to interfere in elections, but certainly part of that is a state effort,'" Tweet, https://twitter.com/MunSecConf/status/1228704286089064448.

[170] | Homepage, Perspective API, accessed June 4, 2021, https://perspectiveapi.com/.

[171] | Homepage, Project Shield, accessed June 4, 2021, https://projectshield.withgoogle.com/landing?hl=en.

[172] | William Marcellino, Todd C. Helmus, Joshua Kerrigan, Hilary Reininger, Rouslan I. Karimov, and Rebecca Ann Lawrence, *Detecting Conspiracy Theories on Social Media: Improving Machine Learning to Detect and Understand Online Conspiracy Theories*, RAND Corporation, April 2021, https://www.rand.org/pubs/research_reports/RRA676-1.html.

[173] | See, for example: Nathaniel Gleicher, "Coordinated Inauthentic Behavior Explained," December 6, 2018, Facebook, https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/.

[174] | See: https://about.fb.com/?s=%22Coordinated+Inauthentic+Behavior+Report%22.

[175] | Facebook, *Threat Report. The State of Influence Operations 2017-2020*, May 2021, https://about.fb.com/wp-content/uploads/2021/05/IO-Threat-Report-May-20-2021.pdf.

[176] | See, for example: Twitter Safety, "Disclosing networks of state-linked information operations," Twitter, February 23, 2021, https://blog.twitter.com/en_us/topics/company/2021/disclosing-networks-of-state-linked-information-operations-.html; Twitter Safety, "Disclosing networks of state-linked information operations we've removed," Twitter, June 12, 2020, https://blog.twitter.com/en_us/topics/company/2020/information-operations-june-2020.html.

[177] | See: Threat Analysis Group Blog, Google, accessed June 4, 2021, https://blog.google/threat-analysis-group/.

[178] | See, for instance: Tom Burt, "New nation-state cyberattacks," Microsoft, March 2, 2021, https://blogs.microsoft.com/on-the-issues/2021/03/02/new-nationstate-cyberattacks/.

and analysts with whom they have developed relationships of trust. Their data and information have notably enhanced studies from the DFRLab and Graphika, both of which maintain information sharing agreements with Facebook, or the Stanford Internet Observatory, with which Twitter has also shared data.

Limits to the measures taken

Taken as a whole, the measures presented in the preceding pages may seem impressive. It is undeniable that over the space of a few years, awareness has grown significantly and that states, international organizations, civil society, and even digital platforms have taken significant strides in combating information manipulation in response to a broader society that is more prone to spreading manipulated rather than verified factual information. However, for the reasons listed below, there are still considerable challenges ahead.

One: while fact-checking is necessary, it is not enough, for at least two reasons. The first is that its effectiveness is disputed. Corrective in nature, it is by definition post facto, i.e., it happens once the damage has been done without being able to erase the psychological impact of the false or biased information. In addition, studies have shown that the human brain is resistant to correction (misinformation has a "continued influence" effect)[179] and that people continue to spread information despite knowing it to be false.[180] Correction has great difficulty reaching its target because audiences exist in parallel worlds, since the people most likely to read "fake news" on questionable sites and be convinced it is true are generally not the same people who consult legitimate sites that check facts. Moreover, false or biaised information will always be faster to produce than any correction, because they do not need to be justified, verified, grounded in facts, which takes time. This is not to say that debunking – exposing disinformation as false or biased – does not work: it does work, "to varying degrees with different audiences," depending on how it is being done.[181] But it is obviously not a "silver bullet," and we should be aware of its intrinsic limits.

[179] | Stephan Lewandowsky, Ullrich K.H. Ecker, Colleen M. Seifert, Norbert Schwartz, and John Cook, "Misinformation and Its Correction: Continued Influence and Successful Debiasing," Association for Psychological Science: *Psychological Science in the Public Interest* 13:3 (2012), 106–131; Colleen M. Seifert, "The continued influence effect: The persistence of misinformation in memory and reasoning following correction," in D. N. Rapp and J. L. G. Braasch (eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (Cambridge: MIT Press, 2014), 39-71.

[180] | Tom Buchanan, "Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation," *Plos One*, October 7, 2020, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0239666.

[181] | Global Engagement Center, US Department of State, "What Works in Debunking," *GEC Counter-Disinformation Dispatches #4*, April 14, 2020, https://e.america.gov/t/ViewEmail/i/3C498F364AD56A2B2540EF23F30FEDED/E849266E89035643A7F290B8E8FDC6A0.

The second reason is that what is important is not so much facts as how information is presented. It matters very little that we are right: as long as our adversaries tell better stories, and more quickly, which they are able to do because they do not have to ground their claims in facts – they will win. Correcting false information will not keep it from spreading because what a certain segment of the population is looking for is "good stories," whether or not they are true. As Nimmo explains so well, we are not so much engaged in information warfare as in narrative warfare.[182] From that point of view, correcting false information is not enough. We have to successfully supplant the false or misleading narrative with another narrative, this one factual, by telling a good story. But which story? The story behind the attack, the "whodunit," explains Nimmo. That is what worked in the case of the "Macron Leaks": real-time analyses made it possible to redirect the attention of the public, who were less interested in the contents of the "leaks" as in where the attack was coming from and in the ties these mysterious players might have with a French political party or with foreign powers.[183] Since then, the DFRLab, Graphika, EU DisinfoLab, ASPI, and others have captured the global attention with many excellent reports telling "good stories" of information operations from all over the globe.

Two: The fight against information manipulation can undeniably pose a threat to democratic or liberal values, particularly the freedom of expression and freedom of the press. The fact that outwardly similar countermeasures are taken in both liberal democracies and autocratic regimes is proof of this. For the former to maintain a healthy balance between security and liberty, which is at the heart of the social contract, and protect themselves from comparisons with authoritarian regimes, they must adopt the guiding principle that the state is not and must not be the front line of defense against information manipulation: it is citizens, civil society, particularly journalists, fact-checkers, and researchers, but also platforms, who form the front line. That is why the first recommendation to states in the author's 2018 CAPS-IRSEM report was to have "as light a footprint as possible"[184] – an approach that was notably adopted in the Canadian plan presented in January 2019, which did not assume that the state has to protect its passive citizens but rather endeavors to train them so as to equip them with the means to detect such manipulation themselves.

Three: The virtues of a cross-sector approach are now recognized – in order to defend ourselves against information manipulation, different agencies and ministries across a given government have to work together, but not all administrative cultures are equally receptive to that approach. Certain countries (those in Scandinavia, the Baltic states, Canada) are better at it than others whose bureaucracies are more complex and still very siloed, sometimes with interdepartmental rivalries that paralyze the work. They do not all have the flexibility to create a Centre for International Digital Policy like the one at Global Affairs Canada, whose team is a mixture of political officers and data analysts.

Four: Many countries focus on elections, which is understandable not only because they are one of the embodiments of democracy (and so should be protected in the same way as the "critical infrastructures" that allow our society to function, e.g. electricity, bridges, railways, telecommunications, hospitals, etc.) but also because it is consensual and bipartisan to want to protect them – unless domestic political incentives overwhelm any ability to achieve consensus, as the 2020 US presidential election and their aftermath demonstrated. Nevertheless, we must not lose sight that attacks – whether cyber or informational – occur on a daily basis, so measures cannot be taken only at election time.

[182] | David A. Wemer, "How to kill a disinformation narrative: Make it a whodunit," Atlantic Council, March 8, 2019, https://www.atlanticcouncil.org/blogs/new-atlanticist/how-to-kill-a-disinformation-narrative/.
[183] | Vilmer, *The "Macron Leaks" Operation*, 39 40.
[184] | Vilmer et al., *Information Manipulation*, 169.

Five: In a similar way, many countries focus on foreign interference, expending fewer resources on the increasingly prevalent problem of home-grown information manipulation. “We don’t touch internal disinformation actors, even if they’re extremists,” explains Mikael Tofvesson from Sweden’s MSB, for instance. “The MSB is technically a national defense capability, and we don’t use defense capabilities against our own population.”[185] Many liberal democracies have the same restrictions, which are perfectly legitimate per se. And yet, although they were deliberately not the focus of the present report, which mostly looks at the various intiatives through the lens of foreign interference or influence, there are many cases of domestic interference, perhaps even a majority (twelve out of seventeen networks removed by Facebook in December 2020 were domestic, targeting audiences in their own country, for example).[186] Moreover, all those involved in fighting information manipulation have noted that it is not always easy – or sometimes even possible – to differentiate between “external” and “internal” acts of manipulation. This is all the more difficult because our adversaries profit from ambiguity by using non-state proxies and by going through national operatives. What muddy the waters even more is that, at certain times in certain countries the ruling regime itself may spread disinformation, including foreign-originated disinformation – making it more difficult to know if there is external interference. For those many reasons, states under an informational attack lose time trying to first categorize the attacker as “external” or “internal,” as the departments that combat threats from the outside are generally not the same as those who control domestic threats. Yet, the ambiguous nature of these threats means that defense must be cross-sector. It forces us to break down the barriers between departments, to share information more – something that many bureaucracies are not used to doing.

Six: In Europe, the decision to break the EEAS’ response into three unevenly balanced geographic teams, with the “East” (Russia) one getting most of the resources and being in practice the only one really fighting disinformation, is probably not the best configuration for a number of reasons. One is that it actually divides European countries between those comfortable with being labelled “anti-Russia” and those more reluctant, ambivalent, or even actively pro-Russia, that do not support the East task force (and therefore EEAS’ efforts in fighting disinformation) for that reason. Paradoxically, it can even help Russian propaganda: the fact that the website EUvsDisinfo.eu (the name for which seems to indicate that it is about EU efforts to counter *all* disinfo) deals exclusively with Russian disinformation does not help it not to appear “russophobic.” In that respect, it would be much stronger, including against Russian disinformation, if it diversified. Moreover, countries can be (more) preoccupied with threats coming from the South, or feel they are at a crossroads between the East and the South and advocate taking a 360° view. This geographic logic also has its limits because the Russians are very active in Africa, including in actions against European interests[187]: the “East” is also in the “South” and vice versa. In other words, the fact that, often, the West is West-biased can be detrimental not only to the “Global South,” but to the West itself. That is the reason why it can be useful to put in place cooperation actions against disinformation in Africa and Southeast Asia, for example, like the United Kingdom’s Department for International Development has been doing to counter COVID-related disinformation by supporting local journalists, doing outreach actions in direction of local influencers, developing social media monitoring tools, etc.

A second problem is that this Russia-centered configuration may reflect the world of 2014-

[185] | LaForge, *Sweden defends its elections*, 6.

[186] | “December 2020 Coordinated Inauthentic Behavior Report,” Facebook, January 12, 2021, https://about.fb.com/news/2021/01/december-2020-coordinated-inauthentic-behavior-report/.

[187] | Shelby Grossman, Daniel Bush, and Renée DiResta, “Evidence of Russia linked influence operations in Africa,” Working Paper, Stanford Internet Observatory, October 29, 2019, https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/29oct2019_sio_-_russia_linked_influence_operations_in_africa.final_.pdf. In October 2019, Facebook also deleted three networks of accounts that had a “coordinated inauthentic behavior” attributed to Russian players and targeting African countries. See also: “Removing More Coordinated Inauthentic Behavior from Russia,” Facebook, October 30, 2019, https://about.fb.com/news/2019/10/removing-more-coordinated-inauthentic-behavior-from-russia/.

2018 but is now obsolete if it does not take into account other actors, notably China and Turkey's growing assertiveness in information and influence operations, including against Europe. Creating new "China" or "Turkey" Task Forces may not be the solution for the reasons previously mentioned (some states would not support them so as not to appear "anti-China" or "anti-Turkey"), and also because it is not only about Russia, China, and Turkey: there are many other threats for Europe, coming from state and non-state actors. The regional expertise is certainly needed but, rather than multiplying actor-oriented ad hoc teams, it would be more appropriate to develop a broader, bigger, and better-funded actor-agnostic agency combating all forms of foreign information manipulation, wherever they might come from, while retaining internal region-focused teams (like, in the United States, the GEC has Russia, China, and Iran teams among other, transversal teams and as part of a bigger, well-funded structure). In other words, the objective should be to combine an actor-agnostic 360° approach with actor-specific expertise. A stronger, unitary actor would also help improve European coordination and speed, two of the most obvious, and best known, issues in the current configuration.[188]

Seven: Existing cooperation formats (EU, NATO, G7) are useful to coordinate the fight against foreign information manipulation, but they have intrinsic limitations, first and foremost because they were created for other purposes. They are too broad, too narrow, or both (the EU and NATO both include unlike-minded and exclude like-minded states). They are disparate because, in each of those formats, the extent to which countries are preoccupied by foreign information manipulation basically depends on the relations they have or hope to have with those countries generally presented as the main threats in this regard, i.e., Russia and China. A wide range of divergences do exist, sometimes with outright blockages, which not only leads to a slower and less effective decision-making process but may also give rise to an atmosphere of distrust between members. That is why there is a growing number of calls for the more or less formal establishment of a counter-disinformation ad hoc coalition of like-minded actors[189] (with many questions pending, as to what extent it should be limited to an "alliance of democracies" and to what extent it should include not only states, but also civil society organizations, including digital platforms).

Eight: Social media platforms, which have been widely criticized for a lack of transparency, have made significant progress in terms of information sharing over the past few years. They can no longer be accused of doing nothing as was undoubtedly the case at one time. However, their modus operandi is still primarily reactive rather than proactive. They could be expected to steadily do more and, most especially, to demonstrate greater transparency about the measures they do take, and the ones they do not take. For example, in November 2017, Alphabet Chairman Eric Schmidt announced that Google will "derank" stories from Kremlin-owned media *RT* and *Sputnik*,[190] a measure that apparently has never been implemented (or, if it was, has not been publicly acknowledged). That being said, the focus on Facebook and Twitter, and to a lesser extent Google (YouTube), should not eclipse the issue of platform proliferation: the same requirement that any platform over a certain size has to dedicate a certain effort to detecting and dealing with information manipulation should apply to TikTok, Parler, and others, including the small forums involved in the series of operations the DFRLab and Graphika nicknamed Secondary Infektion.[191]

There is also the issue of attribution. "Platforms are increasingly specific in their attributions": they do not hesitate to name and shame not only countries but sometimes specific agencies or

[188] | See, in addition, the recommendations from: James Pamment, *The EU's Role in Fighting Disinformation: Developing Policy Interventions for the 2020s*, Working Paper, Carnegie Endowment for International Peace, September 2020, https://carnegieendowment.org/files/Pamment_-_Regulation.pdf.

[189] | Dan Fried and Alina Polyakova, *Democratic Defense Against Disinformation*, Atlantic Council, March 5, 2018, https://www.atlanticcouncil.org/wp-content/uploads/2018/03/Democratic_Defense_Against_Disinformation_FINAL.pdf; *Democratic Defense Against Disinformation 2.0*, 16; Vilmer, *The "Macron Leaks" Operation*, 46.

[190] | "Google to 'derank' Russia Today and Sputnik," *BBC News*, November 21, 2017, https://www.bbc.com/news/technology-42065644.

actors ("approximately 76% of takedowns were attributed to a specific actor in 2020, compared to 62% in 2019 and 47% in 2018").[192] It should be noted that the EU's General Data Protection Regulation (GDPR) sometimes interferes in the platforms' ability to "name and shame" a third party publicly. In any case, platforms generally do not justify or explain such attributions. For example, when, a few days before a takedown, Facebook gives advance notice to the DFRLab, Graphika, or any other research organization with which they have an information sharing agreement, it informs them of the pages, accounts, groups, or events involved before deleting them on the grounds that they are linked to a specific actor ("Russian military intelligence operators" in the case of the *From Russia with Blogs* Graphika report of February 2020, for example).[193] But Facebook does not say how it arrived at that conclusion. The research community has to either take the company at its word or develop its own means for confirming – or disproving – such attributions. Attribution, which consists of identifying the source of a cyber operation, is notoriously difficult or even impossible in a large number of cases. And even if the search may lead to one computer, that does not tell us who was sitting at it and whether or not that person was acting under their own initiative or under state control.[194] Similarly, on a social media network, the identification of an "account" does not amount to the identification of the "user(s)." Of course, sharing information as sensitive as that which made the attribution possible would amount to exposing detection methods, which would benefit perpetrators, who would find a way to get around them the next time. So, the question is, could digital platforms develop ways to share information that is confidential (and must remain so) with governments but also with a few trusted researchers, by setting up several layers of access? Twitter recently improved its services to researchers by launching a new API (v2).[195] Will it allow researchers to check the platform's methodology for attributing disinformation campaigns before taking them down?

[191] Nika Aleksejeva, Lukas Andriukaitis, Luiza Bandeira, Donara Barojan, Graham Brookie, Eto Buziashvili, Andy Carvin, Kanishk Karan, Ben Nimmo, Iain Robertson, and Michael Sheldon, *Operation "Secondary Infektion,"* DFRLab, June 2019, https://78251c6f-ea33-4b6e-85f0-8ffd60cb35f5.filesusr.com/ugd/9d177c_3e548ca15bb64a85ab936d76c95897c7.pdf; Ben Nimmo, Camille François, C. Shawn Eib, Lea Ronzaud, Rodrigo Ferreira, Chris Hernon, and Tim Kostelancik, *Secondary Infektion*, Graphika, June 16, 2020, https://secondaryinfektion.org/downloads/secondary-infektion-report.pdf.

[192] Josh A. Goldstein and Shelby Grossman, "How disinformation evolved in 2020," *Tech Stream*, Brookings, January 4, 2021, https://www.brookings.edu/techstream/how-disinformation-evolved-in-2020/.

[193] Ben Nimmo, Camille François, C. Shawn Eib, and L. Tamora, *From Russia with Blogs*, Graphika, February 2020, https://public-assets.graphika.com/reports/graphika_report_from_russia_with_blogs.pdf.

[194] To learn more about how Facebook deals with the difficulties of attribution, see: Alex Stamos, "How Much Can Companies Know About Who's Behind Cyber Threats?," Facebook, July 31, 2018, https://about.fb.com/news/2018/11/investigating-threats/#whos-behind-cyber-threats.

[195] Adam Tornes and Leanne Trujillo, "Enabling the future of academic research with the Twitter API," Twitter, January 26, 2021, https://blog.twitter.com/developer/en_us/topics/tools/2021/enabling-the-future-of-academic-research-with-the-twitter-api.html.

Conclusion

This list of limitations is certainly not comprehensive and there would be many more issues to point out. Perhaps a more fundamental problem is that, if states, individually and collectively, NGOs, journalists, researchers, digital platforms and others have to take the measures briefly presented in this report, it is because the broader societal disinterest (whether intentional or not) in prioritizing facts and discarding unverified assertions is forcing the responsibility onto those actors. Such a disinterest is rooted in structural causes,[196] some of them psychological (an intellectual laziness, i.e., a failure to systematically exercise critical thinking, a number of cognitive biases, validation by social interaction, authority argument, illusion of correlation, etc.) and others societal (crisis of confidence in institutions, crisis of the press, and digitalization because, as Ben Nimmo explained, "the spread of digital publishing technologies has made it easier to create false stories. The internet has made it easier to publish fake stories, and social media have made it easier to *spread* false stories"[197]). Ultimately, this causes an epistemological crisis, i.e., a decline of rationalism in an era of so-called "post-truth" politics. And those are not easy issues to fix. The very fact that information manipulation thrives despite all the policy measures taken against it is the symptom of our impotence in that regard. It does not mean we cannot progress, but it does mean that progress is a long-term effort.

[196] | Vilmer et al., *Information Manipulation*, 29-42.
[197] | Ben Nimmo, *Written Representation 26*, Testimony before the Select Comm. on Deliberate Online Falsehoods—Causes, Consequences and Countermeasures, Parliament of Singapore, February 22, 2018, https://www.parliament.gov.sg/docs/default-source/sconlinefalsehoods/written-representation-36.pdf.

About the author

Dr. Jean-Baptiste JEANGÈNE VILMER is director of the Institute for Strategic Research (IRSEM) at the French Ministry of the Armed Forces, and a nonresident Senior Fellow at the Atlantic Council's Europe Center. Trained in philosophy (bachelor, master, PhD), law (bachelor, LLM, post-doctorate), and political science (PhD), he teaches at Sciences Po's Paris School of International Affairs (PSIA). He served as policy officer on security and global affairs at the Policy Planning Staff (CAPS) of the French Ministry of Foreign Affairs, and has held positions at the faculty of law at McGill University, Canada, the department of war studies of King's College London, Yale University, the French Embassy in Turkmenistan, and the University of Montreal. An Honorary Ancien of the NATO Defense College, his research focuses on international relations and new forms of conflictuality. On hybrid threats and information manipulation, he has authored several reports, including *Information Manipulation: A Challenge for Our Democracies* (CAPS/IRSEM, 2018), *The 'Macron Leaks' Operation: A Post-Mortem* (IRSEM/Atlantic Council, 2019), and *Effective State Practices Against Disinformation: Four Country Case Studies* (European Centre of Excellence for Countering Hybrid Threats, 2021).

Atlantic Council

DIGITAL FORENSIC
RESEARCH LAB

Atlantic Council

EUROPE CENTER