VIOLATIONS OF SOVEREIGNTY IN "CYBERSPACE" UNDER THE UNITED NATIONS CHARTER

ASSAF A.

Alaa Assaf — Specialist in International Law, Damascus, Syria (alaa.assaf89@gmail.com). ORCID: 0000-0003-4714-1401

Abstract

Affirming that violating State sovereignty through and against “cyber” infrastructure could be covered by the scope of Art. 2(4) and (7) of the United Nations Charter is one of the most pressing challenges that faces international law today. This article aims to address this issue by expanding on a general taxonomy outlined in the Tallinn Manual 2.0 on violations of sovereignty in “cyberspace”. These violations are categorised as conducts leading to either “infringement upon the target State’s territorial integrity” or “interference or usurpation of inherently governmental functions”. In order to map the taxonomy of the Tallinn Manual 2.0 onto Art. 2(4) and (7), it is necessary to highlight the convergence between territorial sovereignty and “cyberspace” that allows for extending the scope of application of Art. 2. Through recognising data as “assets” that can be subject to a functional sovereignty, that in turn could be subject to unlawful use of force in violation of the general ban codified in Art. 2(4) as an “infringement upon the target State’s territorial integrity”. Extending the scope of Art. 2(7) is contingent upon defining the concept of intervention as a conduct aiming to unlawfully assume an exclusive competence of a State by another State. Under this concept, intervention in “cyberspace” could be envisaged as attempts to gain control over the functionality of certain “cyberspace” infrastructure that is instrumental for the manifestation of State exclusive competences. A process that demands taking control of that entity to an extent impinging the regular functioning of the targeted entity beyond the mere manipulation of data. Under the proposed definition of intervention such conduct of “interference or usurpation of inherently governmental functions” can constitute a violation to the principle of non-intervention as codified by Art. 2(7).

Key words


https://doi.org/10.17323/jil.2023.18848

Introduction

One of the key problems for international lawyers today is to configure political interactions in the dominium of “cyberspace”² with operational lex lata, especially those related to violations of sovereignty. Difficulties range from those related to scarcity of legal instruments on “cyberspace” that could have allowed lawyers to extrapolate norms and rules through interpretation efforts to those related to the lack of adequate knowledge on the technical aspects of “cyberspace” by lawyers.²

“Cyberspace” could be defined as a “domain characterised by the use of electronics and the electromagnetic spectrum to store, modify, and exchange data via networked systems and associated physical infrastructures”.³ “Cyberspace” is a form of Information and Communications Technology (hereinafter — ICT) that facilitates the exchange of data, with data conceived as “given” digital representations by and of actors over “cyberspace”.⁴ The Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations (hereinafter — Tallinn Manual 2.0) — with its lex lata approach — recognised the extension of sovereignty into “cyberspace”, accepted to be applying ratione loci and

ratione personae. Conceiving “cyberspace” as constructed of three layers, all of which are encompassed by the principle of sovereignty: first, the physical layer; second, the logical layer, and, third, the social layer.

With “cyberspace” being recognised as intrinsically physical, the issue of violating the sovereignty of States through “cyberspace” might seem straightforward: any material attack on material subjects or the material manifestations of a State is governed by the corpus of jus ad bellum. However, the issue at hand is far more complex than deciding the locus. Instead, it is to assess how damages can be done through, across, and against “cyberspace” against the sovereignty of a State in a manner that might not constitute any physical consequences. Such modalities of violations are commonly described as “cyber operations”: activities that involve the use of cyber infrastructure or employ cyber means to affect the operation of such infrastructure (cyber act), to achieve certain objects in or through “cyberspace” (a context).

The question now is how to navigate international legislative silence surrounding “cyberspace”? In such situation the most reasonable approach is to conduct empirical inquiries beyond the tautology of positivism. In the case of developing new international legal rules, relying on empirical inquiries might eventually lead up to a behavioural inquiry. Legal behaviouralism could be described as a legal technicality that emphasises on exposing cognitive biases and heuristics to “explain” what actors might reveal rather than to help “understand” the meaning of their actions. With that said, behaviouralism might be accused of over-justifying indeliberate heuristic reasoning and deploying persuasive sampling strategies to construct an apologetic legal argument. Such criticism, while valid, must factor in the initial state of the discourse of legal governance of “cyberspace” that lacks grundnorm to assess derived norms against. In this case, behaviouralism can provide a rudimentary foundation for a prospect development of new international legal rules or principles beyond legal formalism, by transcending governmental bureaucracies into evaluating the social actualities and fluidities of practice of power by selected organs of multiple States, arranged as a discursive analysis on the governmentality of the production of reality, and perhaps also space. A social, empirical practice of legal determination substitutes validity in legal ascertainment with questions of facts of those social practices by actors of a legal system.

While behaviouralism is essentially a social inquiry, that can contribute to formal legal frameworks as that of formal sources of international law. Behaviouralism can be informative of new legal semantics associated with the practice of power, or to contribute to an emerging opinio juris as an aggregation of law-abiding motivations from individual States “intrinsic” to a certain practice of power in a certain political context. And even if behaviouralism might suffer from failing to address less-orthodox constructs of opinio juris in the creation of a new rule of customary international law, as in crediting normative rules through mutual recognition. Yet in the normative circumstances surrounding “cyberspace” governance, criticism of policy-oriented approaches as behaviouralism can be tolerated.

Having decided on a methodology, I shall turn now to sketch a legal framework on sovereignty violations across “cyberspace”. The Tallinn Manual 2.0 provided a general taxonomy on what constitutes a violation of sovereignty ratione loci and ratione personae across the material manifestations of “cyberspace” through “cyber operations”. The first category of sovereignty violations refers to “infringement upon the target State’s territorial integrity”, while the second category refers to “interference or usurpation of inherently governmental functions”. This taxonomy was endorsed by a number of

---

States such as the Netherlands,14 Sweden,15 Switzerland,16 Canada,17 and Norway.18 Moreover, the terminology adopted by the International Group of Experts (hereinafter — IGE) — the working group which drafted the Tallinn Manual 2.0, arguably reflective of an attempt to map “cyber operations” into the framework of United Nations Charter (hereinafter — UN Charter), precisely to map the “infringement upon the target State's territorial integrity” into Art. 2(4) (the general ban on the use of force), and to map “interference or usurpation of inherently governmental functions” into Art. 2(7) (the principle of non-intervention). Accordingly, the aim of this paper is to further the taxonomy of the Tallinn Manual 2.0 along with the biases, heuristics, and vocational factors that contributed to its emergence and endorsement by some States,19 to inquire if international law can accommodate such taxonomy or not.

1. The infringement upon the target state’s territorial integrity

The IGE tried to provide a general roadmap to categorise infringement upon State’s “territorial integrity”. These were: inflicting physical damage, loss of functionality, and infringement upon territorial integrity falling below the threshold of loss of functionality.20 With this roadmap, the IGE fuelled confusion regarding the suggested semantics, despite pushing a case-to-case approach as a defence, but without success.21 This unfortunate conclusion resulted from confusing material and immaterial aspects of “cyberspace” by equating “cyber operations” causing physical damage with those causing only functional, immaterial damages or infringements. All for the sake of forcefully mapping both modalities into the legal framework governing what is referred to as “territorial sovereignty”: the intrinsically-material, spatial manifestation of sovereignty. The better solution is to set aside functional issues for now, and to isolate “cyber operations” causing “physical damages” as the proper representation of violations of a State’s spatial territory.

1.1. Assessing the physical damage of “cyber operations”

The concept of “physical damages” was endorsed to be the legal standard for sovereignty violations across the material layers of “cyberspace”. International law is indecisive on the definition of damage, particularly in relation to the concept of injury, since the issue is heavily contextual ranging from damages to “respected” interests caused by a use of force to those committed against the environment.22 Tort law recognises damnum sine injuria (damage for which there is no remedy in law), and injuria sine damno (legal wrong not causing actual damage),23 and international law seems to accord. The travaux of Draft articles on Responsibility of States for Internationally Wrongful Acts rejected that injury “consists” of damage(s), instead adopting “injury” as an umbrella term that “includes” any damage, material or moral, that give rise to reparation as the central standard. Accordingly, to claim reparation from a physical or material injury it is then required for the damage to be “actionable”:24 a rule known in common law as the legal-injury rule, stating that the damage, or some of it, should already sustained and assessable in

18 UN Doc A/76/136 2021 65–68.
financial terms so a legal claim could rise (de minimis non curat lex). The IGE’s argument of “cyber operations” “infringing” on the spatial territory of a State within the intended meaning of Art. 2(4) of the UN Charter can only be sustained in cases of actual pecuniary, physical damage, regardless of the need to assert a loss of functionality of a related physical-layer object.

In 2010, the Natanz fuel enrichment plant in Iran was hit by a malware referred to as “Stuxnet” that managed to infiltrate industrial control systems of centrifuges used to enrich the uranium, altering their movement speed and causing physical damage. Iran admitted damages without providing estimations. In 2012, Saudi Aramco was hit by malware initially thought to be a wiper virus dubbed later as “Shamoon”. This malware did more than erasure of data logically, since it managed to corrupt Hard Disk Drives’ (hereinafter — HDD) Master Boot Record, hence “destroying” Aramco’s data storage hardware consisting of over 30000 devices, forcing Saudi Arabia to rush physical replacements. The “Stuxnet” and “Shamoon” malwares’ factual causation to the sustained financially assessable physical damages stand as examples of violation of sovereignty through “cyber operations”. In the case of “Stuxnet”, Iran did not accuse any State, but the political narrative surrounding Iran’s nuclear program made it convincing to read “Stuxnet” as a “cyber operation” conducted against Iran by a foe State, notably Israel and/or the United States. Giving its physical remoteness, it would be difficult to ascribe “Stuxnet” as an infringement on the “territorial integrity” of a State, as such description traditionally held a connotation of a physical cross-border activity or belligerent occupation. Against this backdrop, “Stuxnet” was essentially inquired as a violation of Art. 2(4) of the UN Charter, and the first ensuing legal question was whether “Stuxnet” could qualify as a use of force. The Tallinn Manual 2.0 that recognised the Nicaragua definition of an intervention as a low-threshold use of force, agreed that “Stuxnet” was indeed a use of force within the meaning of Art. 2(4), but the IGE did not reach a conclusive answer if “Stuxnet” reached the threshold of an “armed attack” for the purpose of invoking Art. 51 of the UN Charter.

Bearing in mind, arguendo, the likelihood of “Stuxnet” being attributable to a State, the existence of physical damage was fundamental for this assessment, forcing scholars to adopt a holistic approach to circumvent the rigid criteria surrounding the assessment of the use of force in a classic sense, enquiring multiple standards as immediacy, directness, severity, and invasiveness. Those standards hold inherent forensic materiality that are inapplicable to “cyber operations”, hence were abrogated by post hoc analysis of the gravity of the conclusion of the act. Roscini further argued that introducing precise parameters for a gravity factor cannot be feasible, since the text of Art. 2(4) of the UN Charter does not include any such threshold, even Rapporteur R. Ago once mentioned that Art. 2(4) prohibits “any kind of conduct involving any assault whatsoever on the territorial sovereignty of another State, irrespective of its

34 Because of the practical nature of attribution in international law, this article is not concerned with such legal technicalities, but only with grounding doctrinal debates. Hence attribution to a State will be presumed. See on the issue of attribution in relation to “cyber operations” Banks W. Cyber Attribution and State Responsibility // International Law Studies. Vol. 97. 2021. P. 1039, 1046–1054.
magnitude, duration or purposes".36 However, Roscini argues that according to Art. 32 of the Vienna Convention on the Law of Treaties37 a minimal threshold should exist, otherwise a literal interpretation of Art. 2(4) will lead to results that are "manifestly absurd or unreasonable". It would be "absurd or unreasonable" to consider a "cyber operation" causing physical damage to one computer or a server as an act of use of force,38 let alone against the sovereignty of a State.

Applying the same conclusion on the "Shamoon" attack might be difficult, since the damage caused was not physical in the sense of changing the physical characteristics or the form of the target devices through burning, breaking, or melting, as what was caused by "Stuxnet". The "Shamoon" attack is unique in the sense it had effects equal to those caused by physical attack requiring physical remedy through replacing the HDDs, but without causing any tangible physical damage to those devices. Unlike with "Stuxnet", international law scholars were reluctant to address the "Shamoon" attack as a use of force.39 Yet, to avoid confessing a legal vacuum, some scholars attempted the categorise "Shamoon" as unlawful intervention through shifting the attention into the functionality of the operation, depicting such malwares as a "coercion" below the threshold of a use of force as argued in Nicaragua,40 that caused a loss or disturbance of functionality.41 Such approach will bring the doctrinal debate back to square one: the necessity of providing intervention with a conceptual autonomy outside the use of force. Even from a technical point of view, the damages caused by "Shamoon" cannot be equated to attacks that can cause only "logical damages" that could be remedied exclusively through "logical" assistance without permanent data loss.42 Whilst damaged HDDs could be restored to operation through "logical" software remedies, yet it is impossible to restore the "physically" lost data.43 Moreover, it would be economically sound to have them quickly replaced with "clean" new HDDs as Aramco did.

Examples of such "logical" attacks are the Distributed-Denial-of-Service (DDoS)44 attacks in Georgia and Russia during the 2008 South Ossetia war,45 and against Kyrgyzstan in 2009 that crippled the internet across the whole country,46 or even DDoS attacks against Estonia in 2007 that had the effect of shutting down the governmental electronic-based services.47 Despite such financially considerable damages, none could be described as physical, or requiring physical alteration to damaged objects thus changing their identity.48 It appears the "Shamoon" attack stands somewhere in the middle between physically damaging malware such as "Stuxnet", and "logically" incapacitating acts such as DoS attacks. Some scholars suggest that while "Stuxnet" ostensibly constitutes a use of force within the meaning of Art. 2(4) of the UN Charter, "Shamoon" and DDoSs are "logical" interventions causing loss of functionality, "permanent" in the former while "temporary" in the latter.49 Nevertheless, such solution builds on the flawed conception of intervention adopted in Nicaragua as a legislative gap-filler for low-threshold

44 Morley D., Parker C. Understanding Computers: Today and Tomorrow. 16th edition. Cengage Learning, 2016. P. 361. (DDoS attack is an "act of sabotage that attempts to flood a network server or Web server with so many requests for action that it shuts down or simply can no longer handle requests, causing legitimate users to be denied service").
48 I borrowed this argument from the "Ship of Theseus" paradox concerning objects' identity, as whether the object that has replaced each of its components remains essentially the same object. See one of the rare legal applications of the "Ship of Theseus" paradox in the case of Calidad Pty Ltd v Seiko Epson Corporation [2020] High Court of Australia S329/2019.
1.2. The proprietary argument of states’ digital assets

To recognise data as objects capable of being physically targeted through “cyber operations”, data should be recognised as a property. While such idea was generally frowned upon, recently it is becoming more acceptable giving concerns of economic efficiency, civil liberties, and avoidance of unjust interferences through data usage, giving rise to arguments for granting data a legal protection akin to that of private property. The vehicle behind such narratives is indeed technological, however data is now conceived as decentralised, controllable and collectable commodity that can be dematerialised, moved, and then re-materialised or vice versa, all while maintaining data’s contextual integrity and identity. Moreover, individual “givers” of data can aggregate their collection and processing by third parties. Against this backdrop, few legislators reflected through user or individual-centric instruments recognising “personally identifiable information” and “personal data” as interests subject to statutory protection, notably the EU’s General Data Protection Regulation (hereinafter — GDPR) in 2016 and California Consumer Privacy Act (hereinafter — CCPA) in 2018. Both GDPR and CCPA relied on civil law analogy to define “personal data” of natural and artificial persons as possible object of ownership since data, in general, is now thought of as enjoying clearly delineable boundaries as the case with user-held provides, data has economic value and, and data can-in principle- always be disposed of.

In the same vein, the UK Law Commission conditioned the proprietary approach to data referred to as “digital assets”, arguing that even if digital assets can neither fall under the existing trite law categories of chose in possession since data does not have a tangible form stricto sensu that define their very being through physical form (gold, house, car…), nor they can be considered as chose in action since they are not claimable or enforceable by legal action. However, current technical and legal developments in relevant treaty-law, case-law, and literature have started to recognise digital assets with the same “function” of proprietary objects, as a third category tertium quid. Data represented in an electronic medium allows for its definability and retrievability without the very data becoming chose in possession, hence granting the foundation for the controllability of data. Data as objects exist independently of persons, they are relational to an “owner”, and not part of the “owner”, allowing for changes in owner’s identity and even being abandoned. Accordingly, data rights are similar to property rights related to “things” asserted against persons generally, unlike personal rights that are asserted only against the

particular person to whom they relate.\textsuperscript{59} Furthermore, data are independent of the legal system, they are neither a creation of law nor recoverable only by legal action such as the case of intellectual property rights.\textsuperscript{60} Finally, data rights are also rivalrous, meaning that property law should allocate rivalrous objects between persons, and to protect their liberty to use those objects free from the interference of others, this feature has a declaratory aspect requiring practical, legal and moral considerations to recognise this characteristic as a content of a prospect law.\textsuperscript{61} The examples of GDPR and CCPA and their focus on the concept of data privacy serve as an indication of the excludability of data whose protection serves as the object and purpose of those instruments.\textsuperscript{62} Technical examples using Blockchains protocol such as crypto-currencies and Non-Fungible Tokens further support the rivalrousness of “data assets” as an inherent feature in the data-object itself.\textsuperscript{63}

International law seems also to be heading in the same path of recognising data with proprietary legal status. The 2023 International Institute for the Unification of Private Law (hereinafter — UNIDROIT) Draft Principles on Digital Assets and Private Law Art. 2(2) recognised that “digital asset” means an electronic record which is capable of being subject to control.\textsuperscript{64} Art. 3(1) added that “digital asset” can be the subject of proprietary rights.\textsuperscript{65} The emphasis on the control is central for the UNIDROIT and is reflective of the core conclusions of the UK Law Commission mentioned above (data represented in an electronic medium, independent existence, and rivalrousness), Art. 6(1)(a) defined, that if a person has “control” of a digital asset, this person possess, firstly, the exclusive ability to prevent others from obtaining substantially all of the benefit from the digital asset; secondly, the ability to obtain substantially all the benefit from the digital asset, and, thirdly, the exclusive ability to transfer the abilities in subparagraphs (a)(i), (a)(ii) and (a)(iii) to another person (a “change of control”).\textsuperscript{66}

Data as res provides that it can be subject to political rule, thus territorialised vis-à-vis the national legal order. The issue now is how to frame this territorialisation under international law, or as a question of in rem or competence. Literature on the issue of property in international law concerns individual rights, and very little attention was paid to property rights of States under international law.\textsuperscript{67} The International Court of Justice (hereinafter — ICJ) had a unique opportunity to tackle the issue of immunities 	extit{ratione materiae} against State assets while addressing a claim concerning the seizure of Timorese governmental documents by Australia. Timor-Leste contended that the seized documents and data are protected properties under international law, arguing that the inviolability of State property and State immunity is a well-established rule of customary international law.\textsuperscript{68} Against that particular point Australia replied that there is no general inviolability of State property, indicating that inviolability and foreign State immunity are different concepts and should not be confused, as the law of inviolability applies to specific subjects under specific legal regimes, which do not apply in this case.\textsuperscript{69} Despite the unfortunate discontinuance by Timor-Leste, the Australian counter-arguments are worth noting. International treaty-law supports the Australian stance, ranging from instruments regulating diplomatic immunity of subjects located within the spatial territory of a State,\textsuperscript{70} to those regulating State-owned vessels outside that spatial territory.\textsuperscript{71}

\textsuperscript{59} Ibid. P. 82–84.
\textsuperscript{61} Digital Assets: Consultation Paper 256. P. 87–94.
\textsuperscript{63} Digital Assets: Consultation Paper 256. P. 90.
\textsuperscript{65} Ibid. Art. 3(1).
\textsuperscript{66} Ibid. Art. 6(1)(a).
While it is plausible to argue that inviolability of State properties has been exclusively governed through lex specialis regimes through a reactive piecemeal approach rather than a general proactive approach. Still, States can own property in their sovereign capacity, and the sovereignty that applies on such property could be described as essentially sovereignty ratione materiae or functional sovereignty. And if we take into consideration that data does qualify as a property, hence nothing prevents the extension of sovereignty on State-owned data. The key example that describes the attempt to merge State-owned property (including public property) with functional sovereign immunity is Estonia’s “data embassy” that was inaugurated in Luxembourg following a bilateral agreement in 2017, that referred in its preamble to the insufficiency of current diplomatic relation lex lata to set a general legal framework for hosting of data and information systems. Still the agreement was admitted as being concluded in the spirit of the Vienna Convention on Diplomatic Relations. The “data embassy” is a “data centre” to host Estonian data storage systems or “assets” located in “premises” provided by Luxembourg. Art. 3(1) of the agreement explicitly provides that the “premises shall be inviolable and thus exempt from search, requisition, attachment or execution”.

Within the proposed framework of “functional sovereignty” over State-owned data a question might rise what if an attributable State-sponsored “cyber operation” like “Shamoon” managed to breach encryption defences of State-owned digital assets and caused “physical” data loss beyond “logical” restoration. In such a case will such conduct amount to an unlawful use of force? M. Schmitt suggested that attacks on data designed to be immediately convertible into tangible objects, like banking data, could be “reasonably” described as a use of force based on qualitative assessments of the damages. In this reasoning the centrality of the property element is undeniable. Norway categorises “cyber operations” causing total loss of data among sovereign violations amounting for infringements upon the target State’s territorial integrity just as physical attacks, and not as a usurpation of “inherently governmental functions”. It can be argued that an attack as “Shamoon” violates the sovereignty of the State ratione loci as unlawful use of force contrary to Art. 2(4) of the UN Charter. And also violates the “functional sovereignty” of the State(s) owning the damaged digital assets on the same basis, but only if the digital assets were inviolable according to lex specialis regimes, not exclusive to diplomatic “digital embassies”.

2. Interference or usurpation of inherently governmental functions

The second category suggested by the Tallinn Manual 2.0 describing sovereignty violations in “cyberspace” is “interference or usurpation of inherently governmental functions”. Clearly, the first element to be inquired in this context is what exactly are inherently governmental functions? The IGE could not define inherently governmental functions, nor could reach a consensus on whether such conducts need to (physically) manifest on “cyber” infrastructure of the victim State. The only indicative trace they left was a footnote referring to acta jure imperii used in the context of State immunity to assess the inherently governmental nature of targeted entity. Furthermore, the academic literature is not as elaborate on the matter of interference or usurpation of inherently governmental functions, as it is regarding “territorial inviolability”,

---

72 Confotti B. The Theory of Competence in Verdross // European Journal of International Law. 1995. Vol. 6. P. 70. (Functional sovereignty, or sovereignty ratione materiae can be defined as an exclusive competence of a State exercisable only, and within the limits of which it is necessary, in order to reach a definite object, to satisfy a definite interest, and this competence cannot be presumed unlike for example the spatial sovereignty of a State inside its spatial territory).
73 UN Doc A/76/136. P. 67–68.
75 Ibid. Art. 1–2.
76 Ibid. Art. 3(1) [emphasis added].
or to criticise the very category of inherently governmental function for the lack of granularity and the difficulty in applying the concept of coercion in “cyberspace” within the framework of non-intervention to begin with. 83

2.1. Conceptualising intervention in “cyberspace”

The objection concerning the relation between coercion and non-intervention raised by Nicaragua can be answered by suggesting *arguendo* that intervention should be understood as usurpation or dictatorial interference with sovereign prerogatives and functions, a much broader scope than cross-border physical coercion. 84 However, the objection concerning the vagueness surrounding the concept of inherently governmental functions is more legitimate. In order to eliminate this vagueness, it is necessary to trace back its origins. The earliest verbatim usages of this expression can be found in the US Federal Activities Inventory Reform Act (hereinafter — FAIR) of 1998. 85 Art. 5 of the Act defined inherently governmental functions as those activities “so intimately related to the public interest as to require performance by Federal Government employees”. 86 Those activities are “among other things, the interpretation and execution of the laws of the United States”. 87

The American legislator was reluctant to use an umbrella term to describe the subject matter of the provision, unlike the IGE who were more relaxed — even if with a footnote — to use the description of *acta jure imperii*, and this understandable giving that the FAIR Act was intended to be a regulation rather than a law. That said, it is regrettable the IGE did not draw any interpretive reference to the arguably borrowed expression. Surely reference to *acta jure imperii* provides some insights on the content of expression under international law. But bridging between national laws and regulations and international law is not straightforward, giving the normative discordant that plagues the relation between the two orders. In policy terms, inherently governmental functions could be simply descriptive of public governance areas that require officials to exercise discretion. 88 Hence national laws related to these functions will be presumed teleologically under the disguise of *acta jure imperii* as acting virtue of a rule of competence under international law. 89 The maxims of *civitas quae superiorem non recogno et par in parem non habet imperium* do lend support to this hypothesis. This time the issue here concerns the content of those competences within the context of “cyber operations”. On this matter governmental “cyber” policy instruments might give some insight.

Australia links directly between prohibited intervention as encapsulated by Art. 2(7) of the UN Charter and “[C]oercive means are those that effectively deprive or are intended to deprive the State of the ability to control, decide upon or govern matters of an inherently sovereign nature”. 90 Switzerland’s position is more nuanced in describing what could constitute sovereignty violations through interference with or usurpation of inherently governmental functions, by focusing on the concept of “control” where related data has been altered interfering with the operation and control public infrastructure, public services (social services, conducting elections and referendums, taxes…), or public decision-making processes. 91

90 UN Doc A/76/136. P. 5, 16.
Norway also went into the same direction interpreting functionality loss or disturbance as the loss of control over public infrastructure and services regardless of the need to prove physical damages.\textsuperscript{92}

Out of the above, four commonalities could be extracted: first, the element of “control” informs the notions of “interference” and “usurpation”; second, the target of the “cyber operation” must qualify as an inherently governmental function; third, actual manifestations should already be sustained and noted, and causally linked to a State-sponsored “cyber operation”. The mere loss or interference of control does not suffice, and; fourth, physical damage is not a requirement; its occurrence will instead trigger the use of force framework as encapsulated by Art. 2(4) of the UN Charter.

2.2. Mapping “cyber” intervention against inherently governmental functions

In 2015, the national power grid of Ukraine fell victim to a malware dubbed “BlackEnergy”, allegedly attributed to Russia.\textsuperscript{93} “BlackEnergy” caused widespread electricity outages across Ukraine, as it managed to “take control” of the control systems of certain power stations and executed shutdown operations at substations. Failures of infrastructure were also reported alongside the “destruction” of data on servers and denied support services to subscribers. The “BlackEnergy” was described as “the world’s first case of a successful ‘cyberattack’ on energy facilities.”\textsuperscript{94}

Whether “BlackEnergy”, and indeed similar “cyber operations”, could amount as an “interference or usurpation” of inherently governmental functions in violation of Art. 2(7) of the UN Charter required the first criterion to determine if the element of “control” was present. “Control” as a word of law refers to the direct or indirect power to govern the management and policies of a person or entity, whether through ownership of voting securities, by contract, or otherwise.\textsuperscript{95} “Control” in the context of Information and Communication Technology is only different on the level of subjects, as it means to regulate, direct, command, or govern a “system”, and a “system” in turn is a collection or arrangement of elements (subsystems). A “control system” is hence an arrangement of physical components connected or related in such a manner as to command, regulate, direct, or govern itself or another system. A control system is delimited through identifying the process(s) of input(s) and output(s) across that system, the input(s) represents the stimulus reflecting intended outcome(s), while the output(s) represents a response reflecting the actual outcome(s).\textsuperscript{96} The process of input(s) and output(s) of a system is managed through a System Control Software or a “control program”.\textsuperscript{97} In “cyberspace” terms, control refers to a “logical” arrangement that governs the input(s) and output(s) of a “physical” system. Consequently, “control” as a technical arrangement is better conceived as a question of fact, not as a question of law\textsuperscript{98} informing a \textit{prima facie} legal trigger, since such “logical” “cognitive” formulation is inscrutable by a behaviourist international law. In the case of “BlackEnergy”, the criterion of control was arguably fulfilled even if only at a “logical” level, since “BlackEnergy” managed to “take control” of the functions of inputs and outputs of the affected power systems.

The second criterion requires answering whether electricity grids could qualify as an inherently governmental function. Strikingly, outside the US and the Tallinn Manual 2.0 the concept of cannot be found. This concept even proved troublesome for the American legislator when applied to contractors, with expandable duties closely associated with inherently governmental functions without sufficient supervision or control on the part of the government. Regulative suggestions aimed to alter the focus for the assessment of an inherently governmental function on the nature of the function instead of checking an exhaustive list of what constitutes these functions.\textsuperscript{99} Moreover, the concept of “critical functions” was

\begin{footnotesize}
\textsuperscript{92} UN Doc A/76/136. P. 68.
\end{footnotesize}
suggested to assist in identifying inherently governmental functions, referring to those functions that are necessary to the agency to effectively perform and maintain control of its mission and operations.\textsuperscript{102} The American law-maker tried to establish a subcategory of inherently governmental functions that are critical for governmental functions but could be nonetheless delegated to a contractor, unlike the original approach of inherently governmental function in the FAIR Act that distinguished these functions through deciding if the function could be delegated to a contractor or not.\textsuperscript{101} By introducing the dichotomy inherently governmental functions and “critical functions” the attention of security policies was shifted into the latter, as demonstrated by the US Critical Infrastructure Security and Resilience Presidential Policy Directive (hereinafter — PPD-21) that designated 16 specific critical infrastructure sectors (communications, defence, energy, food and agriculture, healthcare...) the protection of which now qualifies as a national security interest under the auspice of US Department of Homeland Security.\textsuperscript{102}

The translation of the development of the American approach to the protection of critical infrastructure and functions begs inquiring national policies to determine what constitutes an inherently governmental function or a “critical function”, yet it is difficult to envisage any normative string between the American national practice and that of other States that needs to bounce back at the conclusions of the Tallinn Manuals. The only option to extrapolate the American practice into international law is to dwell on common semantics of States who might also adopt a contextually identical concept of critical function, even at the expense of playing down some of the semantics concluded by the Tallinn Manuals. Put differently, the criterion of inherently governmental function will be replaced by the “critical functions” of States for the purpose of the assessment of interventionist “cyber operations” contrary to Art. 2(7) of the UN Charter, as this term reflects the acculturated behaviour of States that could yield normativity if mapped within a constructivist framework of compliance, created through an endorsement of a desired mutual obligation\textsuperscript{103} that could be later captured by one of the tools of sources of international law.

Surely, inquiring State practice for shared semantics requires laborious data compiling.\textsuperscript{104} Therefore, one should rely only on relevant regional or international instruments to check the fish after being captured by the net. The key fishnets are an EU — hard law — directive and a UN — soft law — report prepared by United Nations Office for Disaster Risk Reduction (hereinafter — UNDRR). The UNDRR report provided the most elaborate disaster risk reduction terminology on a global level. The terminology identified “critical infrastructure” as the “physical structures, facilities, networks and other assets which provide services that are essential to the social and economic functioning of a community or society”.\textsuperscript{105} By prioritising “infrastructure” the definition shifts the attention away from the functionality into the materiality of the subject-matter of protection to be fixed within a spatial network of jurisdiction, however, the functionality remains determinative for designating certain infrastructure as “critical”. The 2022 EU Directive 2022/2557 concerning the resilience of “critical entities” also used the same functionality-based general criteria to identify “entities” that are deemed critical by a State. The directive followed the PPD-21 in suggesting specific 11 categories of what constitute a “critical entity”,\textsuperscript{106} States are only guided by a certain objective checklist that constrains their subjective qualification of an entity to be classified as “critical”. The objective constraints demands the entity concerned to “provides one or more essential services”\textsuperscript{107} the disruptive of which can be significant (the “significant disruptive effect”), that is measurable in geographical, social, and economical terms.\textsuperscript{108}

\textsuperscript{102}Ibid. Federal Register 56233, 56236.
\textsuperscript{103}Maurer T. Op. cit. P. 142.
\textsuperscript{107}UN Doc A/71/644 (Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction), 2015. P. 12.
\textsuperscript{109}Ibid. Art. 7.
The significance of Directive 2022/2557 is that it recognised “critical entities”, whether public or private, governmental, or non-governmental, as interests worthy of legal protection on an international EU level. While the directive identified critical entities in terms of jurisdiction ratione loci following their infrastructure, yet interestingly, the directive added in Art. 17 extra measures for protection to critical entities of “Particular European Significance” that provide essential services for “six or more” EU States. By introducing such legal framework, the directive effectively recognised the possibility to externalise the protection of such entities in application of the concept of domaine réservé the subject-matter of which is particularly susceptible to violations resulting from control alternation. But more crucially, the possibility of applying the concept of domaine réservé informs matters that should be treated by international law as sovereign functions of a State; sovereignty ratione materiae.

That said, if designating “critical entities” as an internal affair of a State, then States who are not party to a legally binding multilateral cooperation framework -as with Directive 2022/2557- cannot be bound by unilateral acts of a third-party State. This argument is certainly true; however, the aim of this analysis is not to scrutinise lex lata, but rather to contemplate de lege ferenda. The rising importance of “critical entities” and the accelerated regional and global efforts to ascribe certain legal protection to those “critical entities” will make State-attributed disruptive effects against their functionality very plausible candidates as sovereignty violations, since wrongful acts against “critical entities” are no longer mere questions of wrongful acts against res, rather a subject-matter of violation of public rights in rem. Accordingly, a disruptive effect through a State-attributed “cyber operation” against electric energy networks, causing total loss of functionally even if temporarily as the case with “BlackEnergy”, is arguably describable as a violation of an inherent internal affair of a State. The same logic could extend to disruptive DDoS attacks against the “public good” of the internet.

The third and fourth requirements are interrelated, they both concern the assessment of the consequences of the wrongful acts. The notion of “damage” is already exhausted by the assessment of “infringement upon the target State’s territorial integrity”. Therefore, a different standard should be adopted to avoid the confusion associated with depicting intervention as a low-threshold use of force. Here the issue concerns the assessment of the consequences resulting from the loss of control over the functionality of a critical entity, without directly leading to a physical damage. Without a doubt this is a very challenging task for a concrete legal order as international law. The Tallinn Manual 2.0 confronted this issue as the treatment by international law of “cyber operations” that results in neither physical damage nor a permanent loss of functionality but only temporarily. Only a minority of the IGE recognised such “cyber operations” as violations of sovereignty, based on a teleological interpretation of the concept of sovereignty “that affords States the full control over access to and activities on their territory”. Since in a different physical scenario as that concerning the function of use of force, the loss of control could be measured applying the “territorial integrity” test. Hence with lack of spatial delineation of function, the attention can be shifted into the very function of critical entities, such as that of electricity grids to produce and provide electricity, thus any reported disruption of that functionality informs a violation.

While this logic might seem entreatable in the case of the energy sector, identifying the functionality of different types of critical entities can get far more complicated than starring at some light bulbs. For instance, what is the critical functionality of an administrative entity conducting an election or a referendum process? Since participation in elections could be disturbed by much more than blocking or spoofing access to voting platforms, as in inserting, manipulating, or deleting data related to the voting process. M. Schmitt suggested that a “rule of reason” should be applied to adhere with assessment to violations against States’ domaine réservé that centres the focus on acts that deprive States from acting vis-à-vis the targeted domaine réservé, instead of the acts or attitudes of the beneficiaries of those critical entities, since the very depriving of a State from control constitutes the requirement of coercion necessary to invoke the non-intervention framework. This rationale suggests that since the concept of physical
damage could not be utilised in cases of *prima facie* intervention, then a general notion of “harm” will substitute “damage” to assess the coercion through the consequences attendant upon the “harm” caused by the deprivation of control. Harm will be measured in relation to a *de minimis* approach to the assessment of the effect of a “cyber operation”, since the concept of sovereignty is informative of a relative “strict inviolability” rule that prohibits all sovereignty degradation attempts as long as they exceed a specific *de minimis* threshold, then, control-related, interventionist, sovereignty degradations should be scaled according to the degree of their interference with the functionality of a targeted “cyber” entity. For example, a “cyber operation” causing only control takeover, or leading to access, stealing, and leaking data, such harm is below the *de minimis* threshold for a prohibited intervention, since such conduct is equitable in its consequences to espionage; a widely permitted conduct under international law during both wartime and peacetime, at least in the case of former when targeting acts of the instrumentalities of States or during the exercise of governmental functions, acting in sovereign or public capacity (*acta jure imperii*). State practice as digested by the IGE concurred with this conclusion.

However, if this control-related, interventionist, “cyber operation” managed to not merely breach the borders of encryption, but also to manipulate data through inserting, deleting or (re)encrypting without causing any direct or indirect damage. Such cases include attacks of power grid systems as in “BlackEnergy” and a similar “NotPetya” ransomware attack in 2017. Both attacks manipulated data in a manner that caused the targeted critical entities to lose functionality without causing any physical damages. “NotPetya” was deemed by the UK a “disregard for Ukrainian sovereignty” by Russia. The US further accused the “Russian military” of conducting the “most destructive and costly cyber-attack in history” that “will be met with international consequences”.

A less pronounced harm caused by “cyber operations” as that caused by “BlackEnergy” and “NotPetya” could also be classified as violation of sovereignty. Such cases include data manipulation without causing any loss of functionality of the targeted critical entity. The 2014 Sony Pictures Entertainment “cyber operation” caused deleting and leaking data related to media production from its database located in New York, notably against a then-upcoming satirical movie “The Interview” that depicted an assassination attempt against North Korean Supreme Leader Kim Jong-Un. The US considered the act as “cybervandalism, not war” against “modern business landscape”, hinting at target’s critical functionality being a commercial facility, and promising to take proportionate measures against North Korea.

Two interesting cases of “cyber operations” targeting the election process are worth noting. The 2017 “MacronLeaks” “cyber operation” caused leaking of thousands of confidential emails belonging to the presidential campaign of Emmanuel Macron. The attack was interestingly described by the then-president Macron as “*immixtion*”/*interference*, a francophone term usually associated with low-intensity

---


interventions against State sovereignty not including any use of force or physical coercion, nonetheless a wrongful act under international law.\textsuperscript{126} Similarly in 2016, Russia was “accused” of hacking into the Democratic National Committee (hereinafter — DNC) of the US Democratic Party, and leaking thousands of confidential emails disclosing the DNCs policies for the then-upcoming 2016 presidential election.\textsuperscript{127} The official response came almost identical to that of the “MacronLeaks” by describing the incident as an “interference” that “undermine established international norms of behaviour”, promising to “response to Russia’s aggressive activities” by “take[ing] a variety of actions at a time and place of our choosing”, and “to holding Russia accountable for what it has done”.\textsuperscript{128}

State practice described above indicates “cyber operations” leading only to a manipulation of data whether overtly or covertly, are increasingly dealt with by States as violations of international law informing acts of retorsion, without necessarily constituting a violation of sovereignty. On the other hand, States are willing to accept control-based “cyber operations” as a violation of the sovereignty of the State whose domaine réservé was targeted to the degree of losing control, without the need to invoke the framework of the use of force. In this regard Australia’s stance is of particular importance giving the direct reference to Art. 2(7) of the UN Charter as the governing frameworks of such violations to sovereignty ratione materiae. I will conclude the results of this Part regarding “cyber-based” sovereignty violations that qualify as interference or usurpation of States’ critical functions, governed by the principle of non-intervention encapsulated by Art. 2(7) of the UN Charter in the following table (Table 1):

<table>
<thead>
<tr>
<th>Control</th>
<th>Data manipulation</th>
<th>Impact on the functionality of a critical entity</th>
<th>International legal consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access</td>
<td>No</td>
<td>No</td>
<td>Interference with the domaine réservé of a State below the threshold of triggering UN Charter Art. 2(7)</td>
</tr>
<tr>
<td>Access</td>
<td>Yes</td>
<td>No</td>
<td>Violation of sovereignty ratione materiae of a State under UN Charter Art. 2(7)</td>
</tr>
<tr>
<td>Access</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

### Conclusion

This article illustrated that the taxonomy suggested by the Tallinn Manual 2.0 is defendable under international law. Hence the behaviour of endorsements and reflection by States to the IGE taxonomy is not void of legal value that in turn could accumulate for a normative basis for future assessments.

Accordingly, it can be argued that Art. 2(4) and (7) do apply to “cyberspace”. Art. 2(4) in the context of “cyberspace” describes “infringement upon the target State’s territorial integrity” manifested causing physcially tangible damages or logically irreversible losses to data that are “actionable” under a


proprietary approach to "digital assets". Such “cyber operation” will qualify as a use of force in violation of Art. 2(4) if the criterion of attribution could be satisfied.

As for the application of Art. 2(7) a key point should be clarified regarding the concept of intervention in international law. Intervention does not refer to the unauthorised cross-border conduct between States, rather it is a reference to attempts to usurp an exclusive competence of a State, regardless of the question of *locus*. Accordingly, an inherently governmental function could be usurped by a State through “cyber operation” even if conducted through the “logical layer” with no physical manifestations. The key here is to confirm a loss of control over and inherently governmental functions in a manner that impinges the very functionality of those functions. In such cases the interventionist conduct of States — if attributable — can qualify as a violation of the principle of non-intervention as codified by Art. 2(7) of the UN Charter. Under the current status of international law, the mere access of control over an inherently governmental function does not qualify as an unlawful intervention or give rise to international legal responsibility. Although in cases of data manipulation, the corpus of secondary rules could be triggered.

Нарушение государственного суверенитета в «киберпространстве»: взгляда через призму Устава ООН

Ассаф А.

Алла Ассаф — специалист по международному праву. Дамаск, Сирия (alaa.assaf9@gmail.com). ORCID: 0000-0003-4714-1401

Аннотация

Вопрос о том, может ли нарушение государственного суверенитета посредством и против кибернетической инфраструктуры подпадать под действие п. 4 и 7 ст. 2 Устава Организации Объединенных Наций, является одним из наиболее насущных вопросов современного международного права. В настоящей статье предприимается попытка ответить на него путем развития общей классификации, предусмотренной Таллинским руководством 2.0 в отношении нарушений суверенитета в «киберпространстве», которое классифицирует эти нарушения как действия, ведущие к «посещению территории» «вмешательство» или к «вмешательству в осуществление функций, присущих государству, или узурпации таких функций». Сближение концепций территориального суверенитета и «киберпространства» позволяет расширить сферу применения ст. 2 и, таким образом, установить соответствие классификации Таллинского руководства 2.0 п. 4 и 7 ст. 2 Устава ООН. Признание данных в качестве актива, на который может распространяться функциональный суверенитет государства, и который может стать объектом незаконного применения силы в нарушение общего запрета, закрепленного в п. 4 ст. 2, позволяет признать атаку на данные «посещению на территориальную целостность государства-мишени». Расширение сферы действия п. 7 ст. 2 зависит от определения понятия «вмешательства» как поведения, направленного на неправомерное присвоение внутренней компетенции одного государства другим. В рамках такой концепции «вмешательство» в «киберпространстве» можно рассматривать как попытку получить контроль над функциональностью определенной кибернетической инфраструктуры, которая используется государством для осуществления им своих суверенных функций. Речь идет о получении контроля над объектом инфраструктуры в такой степени, что это нарушает его нормальное функционирование, то есть вмешательство выходит за рамки простого манипулирования данными. Автор полагает, что в таком случае «вмешательство в осуществление функций, присущих государству, или узурпации таких функций» может представлять собой нарушение принципа невмешательства, закрепленного в п. 7 ст. 2 Устава ООН.

Ключевые слова

«киберпространство», Таллинское руководство 2.0, Устав ООН, функциональный суверенитет, применение силы, принцип невмешательства, внутренние дела государства, кибернетическая инфраструктура


https://doi.org/10.17323/jil.2023.18848

References / Список источников


