University of Reading

An international attribution mechanism for hostile cyber operations

Article

Published Version

Shany, Y. and Schmitt, M. N. ORCID: https://orcid.org/0000-0002-7373-9557 (2020) An international attribution mechanism for hostile cyber operations. International Law Studies, 96. pp. 196-222. ISSN 2375-2831 Available at https://centaur.reading.ac.uk/91877/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: https://digital-commons.usnwc.edu/ils/vol96/iss1/8/

Publisher: U.S. Naval War College

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

University of Reading

Reading's research outputs online

INTERNATIONAL LAW STUDIES

Published Since 1895

An International Attribution Mechanism for Hostile Cyber Operations

Yuval Shany and Michael N. Schmitt

96 INT'L L. STUD. 196 (2020)

Volume 96 2020

Published by the Stockton Center for International Law

ISSN 2375-2831

An International Attribution Mechanism for Hostile Cyber Operations

Yuval Shany[*] *and Michael N. Schmitt*[**]

CONTENTS

I. Introduction 197
II. The Role of Fact-Finding Mechanisms in International Law 202
III. The Present State of Cyber Attribution 211
IV. Proposals for International Attribution Mechanisms 215
V. The Way Forward 218
VI. Concluding Thoughts 221

* Hersch Lauterpacht Chair in Public International Law and Director of CyberLaw Program at the Federmann Cyber Security Research Center, Hebrew University; Vice-President for Research, Israel Democracy Institute.

** Professor of International Law, University of Reading; Research Associate, Federmann Cyber Security Research Center, Hebrew University; Charles H. Stockton Distinguished Scholar-in-Residence, U.S. Naval War College; Strauss Center Distinguished Scholar, University of Texas; Francis Lieber Distinguished Scholar, U.S. Military Academy at West Point; Senior Fellow, NATO Cooperative Cyber Defence Centre of Excellence.

The thoughts and opinions expressed are those of the authors and not necessarily those of the U.S. government, the U.S. Department of the Navy, or the U.S. Naval War College.

I. INTRODUCTION

Cyberspace has long been characterized as anarchic, a domain devoid of normative constraint. Indeed, in remarks to the U.N. General Assembly on his priorities for 2020, U.N. Secretary-General António Guterres urged the international community to "usher in order to the Wild West of cyberspace,"[1] an echo of President Barack Obama's 2015 remark at Stanford University that "[t]he cyber world is sort of the Wild West."[2] Obviously, neither leader meant that international law did not reach cyber activities, but they were signaling the uncertainly pervading the precise application of that body of law in the cyber context.

Scholars and practitioners have been laboring to address that uncertainty. Most significant in this regard have been the two Tallinn Manual projects sponsored by the NATO Cooperative Cyber Defence Center of Excellence.[3] However, both manuals avoided claims of certainty where reasonable differences of opinion existed as to the interpretation of a rule of law in the cyber context. Indeed, their major contribution may have been in pointing out where views diverged, thereby allowing States to focus interpretive efforts where they were most needed.

States are only now beginning to set forth their views on how international law governs cyberspace. Notable examples are statements by, *inter alia*, the United Kingdom, France, the Netherlands, and Australia,[4] but they are

1. U.N. Secretary-General, Remarks to the General Assembly: The Secretary-General's Priorities for 2020 (Jan. 22, 2020), https://www.un.org/sg/en/content/sg/speeches/2020-01-22/remarks-general-assembly-priorities-for-2020.

2. President Barack Obama, Remarks by the President at the Cybersecurity and Consumer Summit, Stanford University (Feb. 13, 2015), https://obamawhitehouse.archives.gov/the-press-office/2015/02/13/remarks-president-cybersecurity-and-consumer-protection-summit.

3. TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS (Michael N. Schmitt ed., 2017) [hereinafter TALLINN MANUAL 2.0]; TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE (Michael N. Schmitt ed., 2013).

4. Jeremy Wright, Attorney General, United Kingdom, Cyber and International Law in the 21st Century, Chatham House (May 23, 2018), https://www.chathamhouse.org/event/cyber-and-international-law-21st-century; MINISTRY OF THE ARMIES, REPUBLIC OF FRANCE, INTERNATIONAL LAW APPLIED TO OPERATIONS IN CYBERSPACE (2019), https://www.defense.gouv.fr/content/download/567648/9770527/file/international+law+applied+to+operations+in+cyberspace.pdf; Letter from the Netherlands Minister of Foreign Affairs to the President of the House of Representatives on the International Legal

overshadowed by the number of States that remain reticent to do so. Furthermore, a review of State practice demonstrates that even States targeted by a hostile cyber operation often refrain from invoking the language of international law when denouncing the attack, when attributing it to a malicious State or non-State actor, or when resorting to cyber or non-cyber measures in response.[5]

Numerous explanations have been proffered for the reluctance of States to invoke international law in relation to hostile cyber operations. One is the assertion that norms developed in a kinetic or offline context are inadequate to address the unique features of cyber operations, which can be perpetrated by proxies, are often designed to spoof the originator, and can cause non-physical effects that are nevertheless very severe. For instance, the concepts of use of force and armed attack in the *jus ad bellum* seem ill-suited when applied to cyber operations that dramatically disrupt civilian life without causing physical damage or that target data. In the same vein, international humanitarian law's rules on targeting cannot easily be interpreted and applied in the cyber context in a manner that reasonably balances military necessity and humanitarian considerations given the integrated nature of many military and civilian infrastructures.[6]

The political or operational interests of States that find themselves on the virtual frontline, either facing hostile operations or launching them, is

---

Order in Cyberspace, app.: International Law in Cyberspace (July 5, 2019), https://www.government.nl/documents/parliamentary-documents/2019/09/26/letter-to-the-parliament-on-the-international-legal-order-in-cyberspace; GOVERNMENT OF AUSTRALIA, AUSTRALIA'S INTERNATIONAL CYBER ENGAGEMENT STRATEGY, ANNEX A: SUPPLEMENT TO AUSTRALIA'S POSITION ON THE APPLICATION OF INTERNATIONAL LAW TO STATE CONDUCT IN CYBERSPACE (2017), https://www.dfat.gov.au/publications/international-relations/international-cyber-engagement-strategy/aices/chapters/annexes.html#Annex-A; *see also* 2019 INTERNATIONAL LAW SUPPLEMENT to *id.*, https://www.dfat.gov.au/publications/international-relations/international-cyber-engagement-strategy/aices/chapters/2019_international_law_supplement.html.

See also papers by States submitted to the 2019–2021 Advancing Responsible State Behaviour in Cyberspace in the Context of International Security. *Group of Government Experts*, OFFICE OF DISARMAMENT AFFAIRS, UNITED NATIONS, https://www.un.org/disarmament/group-of-governmental-experts/ (last visited July 14, 2020); *Open-Ended Working Group*, OFFICE OF DISARMAMENT AFFAIRS, UNITED NATIONS, https://www.un.org/disarmament/open-ended-working-group/ (last visited July 14, 2020).

5. Martha Finnemore & Duncan Hollis, *Beyond Naming and Shaming: Accusations and International Law in Cybersecurity*, EUROPEAN JOURNAL OF INTERNATIONAL LAW (forthcoming 2020).

6. *See generally* Michael N. Schmitt, *Wired Warfare 3.0: Protecting the Civilian Population during Cyber Operations*, 101 INTERNATIONAL REVIEW OF THE RED CROSS 333 (2019).

another common explanation for their hesitancy to invoke international law.[7] They may be reluctant to draw attention to a hostile cyber operation out of embarrassment, because doing so could reveal technical cyber security capabilities and vulnerabilities, or due to concern that it would create unwelcome domestic pressure to respond. States on both sides of the equation might be apprehensive about committing to a specific interpretation of an international law rule, such as on where the threshold for a violation of sovereignty or wrongful use of force lies. And some States certainly believe that a policy of deterrence unfettered by normative strictures offers a more promising way to prevent harmful cyber operations than reliance on international law norms and institutions.

There may be, however, a further explanation for such reluctance—the lack of a credible attribution mechanism capable of validating the facts underlying State legal claims regarding cyber operations. Such information is necessary not only to understand what actually happened, and to identify the culprit, but also to mobilize third-party support for the victim State's assertions, including through collective attribution of (and response to) the operation. More to the point, as a matter of law, "internationally wrongful acts" (violations of international law) require both breach of a State's international law obligation and attribution of the act (which may consist of either an action or an omission) to a "responsible" State.[8]

The International Law Commission's Articles on State Responsibility are generally accepted as restating the customary international law standards for attribution to States. In the cyber context, the two most likely, albeit not only, bases for attribution are that an organ of a State, such as the armed forces, launched the cyber operation that breached the obligation in question,[9] or that a non-State actor, like a hacktivist group or a private company, did so upon the instructions or under the effective control of a State.[10] There are different standards of proof depending on the purpose of the assertion

7. Finnemore & Hollis, *supra* note 5.

8. *Report of the International Law Commission to the General Assembly*, 53 U.N. GAOR Supp. No. 10, art. 22, at 75–76, U.N. Doc. A/56/10 (2001), *reprinted in* [2001] 2 YEARBOOK OF THE INTERNATIONAL LAW COMMISSION 75–76, U.N. Doc. A/CN.4/SER.A/2001/Add.1 (Part 2) [hereinafter Articles on State Responsibility]. For instance, the internationally wrongful act of violating another State's sovereignty requires that the cyber operation qualify as a breach of sovereignty, for instance, because it causes permanent loss of functionality of the targeted cyberinfrastructure, and establishing that the operation had been conducted by a foreign State. *See* TALLINN MANUAL 2.0, *supra* note 3, r. 4, at 17–27.

9. Articles on State Responsibility, *supra* note 8, art. 4.

10. *Id.* art. 8.

of unlawfulness,[11] but legal rights under international law can only be effectively invoked and relied upon in practice on the basis of an adequate factual foundation.

The linkage between the need for factual information about those behind hostile cyber operations and giving effect to the rules of international law has been the driving force behind initiatives by the Atlantic Council,[12] Microsoft,[13] RAND,[14] and academics in support of an international attribution mechanism.[15] The goal of these initiatives is to promote accountability for past operations that violate international law and raise deterrence against future ones. A potentially important step towards realizing this vision was the 2019 establishment of the CyberPeace Institute, which is co-sponsored by the Hewlett Foundation, Mastercard, and Microsoft.[16] Among its goals, the institute intends to perform, facilitate, and coordinate "collective analysis, research, and investigations of sophisticated cyberattacks" in order to "close the accountability gap."[17] Yet, to date, no concrete effort is underway to establish a full-fledged international mechanism charged with engaging in technical attribution of hostile cyber operations, and in linking such operations to States.

It is against this backdrop that we organized an international research project funded by the Dutch Ministry of Foreign Affairs and the Federmann Cyber Security Research Center of the Hebrew University of Jerusalem to consider the feasibility of establishing an international attribution mechanism, as well as the usefulness of such a body.[18] By international attribution

---

11. TALLINN MANUAL 2.0, *supra* note 3, at 81–83.

12. JASON HEALEY ET AL., ATLANTIC COUNCIL, CONFIDENCE-BUILDING MEASURES IN CYBERSPACE: A MULTISTAKEHOLDER APPROACH FOR STABILITY AND SECURITY (2014).

13. SCOTT CHARNEY ET AL., MICROSOFT, FROM ARTICULATION TO IMPLEMENTATION: ENABLING PROGRESS ON CYBERSECURITY NORMS (2016).

14. JOHN S. DAVIS II ET AL., RAND CORPORATION, STATELESS ATTRIBUTION (2017).

15. *See, e.g.*, Elena Chernenko, Oleg Demidov & Fyodor Lukyanov, *Increasing International Cooperation in Cybersecurity and Adapting Cyber Norms*, COUNCIL ON FOREIGN RELATIONS (Feb. 23, 2018), https://www.cfr.org/report/increasing-international-cooperation-cybersecurity-and-adapting-cyber-norms; Milton Mueller et al., *Cyber Attribution: Can a New Institution Achieve Transnational Credibility?*, CYBER DEFENCE REVIEW, Spring 2019, at 107, 107.

16. *Who We Are*, CYBERPEACE INSTITUTE, https://cyberpeaceinstitute.org/who-we-are (last visited July 14, 2020). In full disclosure, Professor Michael Schmitt is a member of the CyberPeace Institute Advisory Board.

17. *Accountability*, CYBER PEACE INSTITUTE, https://cyberpeaceinstitute.org/about-us/accountability (last visited July 14, 2020).

18. Details about the project and workshops it sponsored can be found on Federmann Cyber Security Research Center's website. *See Attribution of Cyber Attacks: Technological and*

mechanism, we mean an entity that is multinational and diverse in its make-up, whether composed of private individuals or public officials, with responsibility for identifying the State or non-State actors who have either conducted a hostile cyber operation or are otherwise involved in that operation.

Workshops organized by the project brought together academics and policymakers, including experts with relevant legal, political, and technological backgrounds, to discuss papers prepared by the project's researchers on such topics as standards of proof for attributing cyber operations under the law of State responsibility, the use of private cybersecurity companies to investigate cyber incidents, investigative models drawn from other technology-intensive fields like weapons control regimes, and the collective attribution practices of the European Union (EU) and North Atlantic Treaty Organization (NATO). This process enabled us to identify certain contexts in which an international attribution mechanism could prove useful and a number of constituencies that might be interested in turning to it in appropriate cases.

The discussions, which were subject to the Chatham House Rule, have led us to conclude that, for the time being, States wielding significant cyber capability have little interest in creating an international attribution mechanism for cyber incidents. Such States appear to be of the view that they can generate sufficient accountability and deterrence based on their independent technological capacity, access to expertise and to offensive (active defense) cyber tools, political clout, security alliances, and other policy tools, such as sanctions.

At the same time, our discussions suggested that countries with limited technological capacity and less ability to mobilize international support for collective attribution are more amenable to the prospect, especially as a tool for "naming and shaming" States conducting unlawful cyber operations against private and public infrastructure in their territory. Furthermore, we are of the view that international or regional organizations that have the ability to facilitate collective sanctions in relation to unlawful cyber operations directed against member States or their partners, like the EU with its sanctions regime,[19] might appreciate in certain situations an official finding of legal responsibility for a number of political and legal reasons.

Legal Dimensions (CSRCI), https://csrcl.huji.ac.il/event/attributing-cyber-attacks (last visited July 14, 2020); *International Accountability Mechanisms: Political and Legal Feasibility* (CRCSI), https://csrcl.huji.ac.il/event/attribution-workshop (last visited July 14, 2020).

19. Council Decision (CFSP) 2019/797 of 17 May 2019, Concerning Restrictive Measures against Cyber-Attacks Threatening the Union or its Member States, 2019 O.J. (L 1291) 13 (EC); Council Regulation (EU) 2019/796, Concerning Restrictive Measures against

This article examines several possible justifications for establishing an international attribution mechanism and its principal constituencies. Part II sets forth the general case for international fact-finding mechanisms in international law, particularly when there is a need to rely on scientific and technological expertise. Part III details the shortcomings of the existing processes for attributing responsibility for hostile cyber operations. Part IV reviews proposals to establish an international attribution mechanism, while Part V examines the constituency for a new mechanism. Our concluding thoughts are set forth in Part VI.

II. THE ROLE OF FACT-FINDING MECHANISMS IN INTERNATIONAL LAW

The linkage between international peace and security and international fact-finding has a long pedigree in international law. In the aftermath of the February 1898 sinking of *USS Maine* in Havana harbor and the ensuing Spanish-American War,[20] a Permanent Court of Arbitration was created pursuant to the 1899 Hague (I) Convention for the Pacific Settlement of International Disputes, a body that continues to operate today.[21] The agreement also provided for international commissions of inquiry "to facilitate a solution of ['differences of opinion'] by elucidating the facts by means of an impartial and conscientious investigation."[22] Fyodor Martens, the prominent Russian diplomat who introduced the commissions of inquiry provisions during the 1899 Peace Conference, described them as establishing a voluntary fact-finding mechanism that could avoid inflammatory disinformation about the causes of an international dispute while gaining time to resolve the matter.[23]

Interestingly, the Russian delegation to the 1899 Conference proposed adding the investigation of questions of responsibility to the duties of the commissions of inquiry, although Martens himself expressed limited enthu-

---

Cyber-Attacks Threatening the Union or its Member States, 2019 O.J. (L 1291) 1. For a summary, see Adam Botek, *European Union Establishes a Sanction Regime for Cyber-attacks*, CCDCOE, https://ccdcoe.org/library/publications/european-union-establishes-a-sanction-regime-for-cyber-attacks/ (last visited July 14, 2020).

20. J. G. MERRILLS, INTERNATIONAL DISPUTE SETTLEMENT 43–44 (4th ed. 2005).

21. Convention (I) for the Pacific Settlement of International Disputes art. 20, July 29, 1899, 32 Stat. 1779, 187 Consol. T.S. 410, as amended, Oct. 18, 1907, 32 Stat. 2199, 205 Consol. T.S. 536.

22. *Id.* art. 9.

23. THE REPORTS TO THE HAGUE CONFERENCES OF 1899 AND 1907, at 50 (James Brown Scott ed., 1917).

siasm for the proposal, noting that it could trespass into the realm of arbitration. He conceded, however, that States were free to conclude special agreements to that effect.[24] Indeed, in 1904, a commission of inquiry established to investigate a maritime incident in the North Sea was authorized to establish responsibility for the incident and to allocate the degree of blame of those involved.[25]

In recent decades, fact-finding and inquiry mechanisms often have been employed in the field of international human rights law to establish accountability for violations and to facilitate subsequent action by political bodies entrusted with promoting respect for human rights. The frequent utilization of ad hoc bodies, such as commissions of inquiry and fact-finding missions,[26] more permanent bodies with renewable mandates, such as thematic and country rapporteurs and working groups,[27] and international treaty bodies invested with the power to conduct inquiries,[28] reflects the international human rights system's post-1967 shift from standard-setting to implementation activities.[29]

---

24. *Id.* at 313.

25. Declaration between the United Kingdom and Russia Relating to the Constitution of an International Commission of Inquiry on the Subject of the North Sea Incident art. 2, Nov. 25, 1904, *reprinted in* HAGUE COURT REPORTS 411 (James Brown Scott ed., 1916). For a full discussion of this Commission, see Jan Martin Lemnitzer, *International Commissions of Inquiry and the North Sea Incident: A Model for a MH17 Tribunal?*, 27 EUROPEAN JOURNAL OF INTERNATIONAL LAW 923 (2016).

26. For a list of the thirty-one ad hoc bodies established by the U.N. Human Rights Council, see *List of HRC-Mandated Commissions of Inquiries/Fact-Finding Missions & Other Bodies (As of October 2019)*, UNITED NATIONS HUMAN RIGHTS COUNCIL, https://www.ohchr.org/EN/HRBodies/HRC/Pages/ListHRCMandat.aspx (last visited July 14, 2010).

27. For a list of thematic and country specific mandates created by the U.N. Human Rights Council, see *Country Mandates*, UNITED NATIONS HUMAN RIGHTS, OFFICE OF THE HIGH COMMISSIONER, https://spinternet.ohchr.org/ViewAllCountryMandates.aspx (last visited July 14, 2020).

28. For a list of U.N. human rights treaty bodies authorized to conduct inquiries, see *Human Rights Bodies – Complaints Procedures*, UNITED NATIONS HUMAN RIGHTS, OFFICE OF THE HIGH COMMISSIONER, https://www.ohchr.org/EN/HRBodies/TBPetitions/Pages/HRTBPetitions.aspx (last visited July 14, 2020).

29. Frans Viljoen, *Fact-Finding by UN Human Rights Complaints Bodies – Analysis and Suggested Reforms*, 8 MAX PLANCK YEARBOOK OF UNITED NATIONS LAW 49, 54 (2004). It has also been suggested that human rights fact-finding, unlike traditional fact-finding, aims to identify and blame those responsible for violations. *See* Larissa J. van den Herik, *An Inquiry into the Role of Commissions of Inquiry in International Law: Navigating the Tensions between Fact-Finding and Application of International Law*, 13 CHINESE JOURNAL OF INTERNATIONAL LAW 507, 536–37 (2014).

Fact-finding mechanisms have also been used to support action by political bodies in other fields of international law, such as international civil aviation law[30] and international labor law.[31] By contrast, international humanitarian law fact-finding mechanisms, most notably the fact-finding commission provided for in the 1977 Additional Protocol I to the 1949 Geneva Conventions,[32] have not been operationalized because of limited interest in resorting to them by States involved in armed conflicts.[33]

In many cases where fact-finding mechanisms have been established and utilized, the resort to independent experts links, directly or indirectly, to the goal of advancing legal and political accountability.[34] By elucidating facts

30. *See, e.g.*, John T. Phelps II, *Aerial Intrusions by Civil and Military Aircraft in Time of Peace*, 107 MILITARY LAW REVIEW 255, 265 (1985).

31. *See, e.g.*, DAVID TAJGMAN & KAREN CURTIS, FREEDOM OF ASSOCIATION: A USER'S GUIDE – STANDARDS, PRINCIPLES AND PROCEDURES OF THE INTERNATIONAL LABOUR ORGANIZATION 72 (2000).

32. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts art. 90, June 8, 1977, 1125 U.N.T.S. 3; *see also* Convention (I) for the Amelioration of the Condition of the Wounded and Sick in the Armed Forces in the Field art. 52, Aug. 12, 1949, 6 U.S.T. 3114, 75 U.N.T.S. 31; Convention (III) Relative to the Treatment of Prisoners of War art. 132, Aug. 12, 1949, 6 U.S.T. 3316, 75 U.N.T.S. 135; Convention (IV) Relative to the Protection of Civilian Persons in Time of War art. 149, Aug. 12, 1949, 6 U.S.T. 3516, 75 U.N.T.S. 287.

33. Heike Spieker, *International (Humanitarian) Fact-Finding Commission*, *in* THE MAX PLANCK ENCYCLOPEDIA OF PUBLIC INTERNATIONAL LAW 505, 513-514 (Frauke Lachenmann & Rüdiger Wolfrum eds., 2017); Robert Heinsch, *The Future of the International Humanitarian Fact-Finding Commission: A Possibly to Overcome the Weakness of IHL Compliance Mechanism*, *in* THE COMPANION TO INTERNATIONAL HUMANITARIAN LAW 79, 81–82 (Dražan Djukić & Niccolò Pons eds., 2018); van den Herik, *supra* note 29, at 529–31.

34. Larissa J. van den Herick & Catherine Harwood, *Commissions of Inquiry and the Charm of International Criminal Law: Between Transactional and Authoritative Approaches*, *in* THE TRANSFORMATION OF HUMAN RIGHTS FACT-FINDING 233, 238–39 (Philip Alston & Sarah Knuckey eds., 2016); *see also* U.N. Human Rights Council Res. S-17/1, Situation of Human Rights in the Syrian Arab Republic, ¶ 13, U.N. Doc. A/HRC/S-17/1 (Nov. 23, 2011)

> Decides to dispatch urgently an independent international commission of inquiry, to be appointed by the President of the Human Rights Council, to investigate all alleged violations of international human rights law since March 2011 in the Syrian Arab Republic, to establish the facts and circumstances that may amount to such violations and of the crimes perpetrated and, where possible, to identify those responsible with a view to ensuring that perpetrators of violations, including those that may constitute crimes against humanity, are held accountable.

U.N. Human Rights Council Res. 28/30, Technical Assistance and Capacity-Building to Improve Human Rights in Libya, ¶ 18, U.N. Doc. A/HRC/RES/28/30 (Apr. 7, 2015) ("Requests the High Commissioner urgently to dispatch a mission to investigate violations and abuses of international human rights law committed in Libya since the beginning of

through a process of investigation that is legitimate and credible, States that violate international law are less able to avoid legal responsibility by denying the underlying facts or their involvement in the situation. Furthermore, publicizing facts with self-evident legal or moral implications can "shame" State and non-State actors, as well as alert them and other actors to the possibility of being held accountable should they continue to engage in unlawful conduct.[35] Determining the facts can also provide a degree of satisfaction for victims.[36] And of course, establishing the factual record can enable follow-up action to assign legal or political responsibility and the taking of appropriate measures against those responsible.[37] In other words, fact-finding fosters accountability by exposing facts and facilitating accountability-generating processes, thereby also enhancing deterrence. In doing so, fact-finding mechanisms contribute to the rule of law in international relations.[38]

A specific area of international relations in which fact-finding mechanisms have proven effective is arms control. Fact-finding in this field, especially with respect to the development, use, and proliferation of unconventional weapons, is especially relevant to cyber activity because States tend to operate behind a shroud of national security secrecy and investigations require sophisticated scientific expertise to collect and analyze data. Examples include the technical verification mechanisms developed under the auspices

---

2014, and to establish the facts and circumstances of such abuses and violations with a view to avoiding impunity and ensuring full accountability."); U.N. Human Rights Council Res. 22/13, Situation of Human Rights in the Democratic People's Republic of Korea, ¶ 5, U.N. Doc. A/HRC/RES/22/13 (Apr. 9, 2013)

> Further decides that the commission of inquiry will investigate the systematic, widespread and grave violations of human rights in the Democratic People's Republic of Korea as outlined in paragraph 31 of the report of the Special Rapporteur, including the violation of the right to food, the violations associated with prison camps, torture and inhuman treatment, arbitrary detention, discrimination, violations of freedom of expression, violations of the right to life, violations of freedom of movement, and enforced disappearances, including in the form of abductions of nationals of other States, with a view to ensuring full accountability, in particular where these violations may amount to crimes against humanity.

35. van den Herick & Harwood, *supra* note 34, 237–38.

36. Cecilie Hellestveit, *International Fact-Finding Mechanisms: Lighting Candles or Cursing Darkness?*, *in* PROMOTING PEACE THROUGH INTERNATIONAL LAW 368, 369 (Cecilia Marcela Bailliet & Kjetil Mujezinovic Larsen eds., 2015).

37. Michael A. Becker & Sarah Nouwen, *International Commissions of Inquiry: What Difference Do They Make? Taking an Empirical Approach*, 30 EUROPEAN JOURNAL OF INTERNATIONAL LAW 819, 834 (2019). This is not always the case, for some fact-finding missions do not result in follow-up. *See* Romana Schweiger, *Late Justice for Burundi*, 55 INTERNATIONAL AND COMPARATIVE LAW QUARTERLY 653, 656 n.22 and accompanying text (2006).

38. Dan Saxon, *Purpose and Legitimacy in International Fact-Finding Bodies*, *in* QUALITY CONTROL IN FACT-FINDING 211, 219 (Morten Bergsmo ed., 2013).

of the International Atomic Energy Agency (IAEA) (through an Additional Protocol to the Comprehensive Safeguards Agreement),[39] the Organization for the Prohibition of Chemical Weapons (OPCW) (through the inspection activity of the Technical Secretariat),[40] and the Comprehensive Test Ban Treaty (CTBT) (through the verification and monitoring activities of the Technical Secretariat).[41]

Verification mechanisms typically have ongoing monitoring responsibilities, such as routine inspections,[42] and the collection of data from monitoring stations[43] and on-site instruments.[44] They can also respond to specific concerns about compliance, as with "challenge inspections" under the Chemical Weapons Convention,[45] investigation of alleged use of chemical weapons pursuant to the U.N. Secretary-General Mechanism,[46] inspections upon request by a State party to clarify whether another State has conducted a nuclear test in violation of the CTBT,[47] and special inspections undertaken

39. International Atomic Energy Agency [IAEA], *Model Protocol Additional to the Agreement(s) Between State(s) and the International Atomic Energy Agency for the Application of Safeguards*, IAEA Doc. INFCIRC/540.Corr (Sept. 1997).

40. Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction art. VIII, Jan. 13, 1993, 1974 U.N.T.S 45 [hereinafter CWC]; *id.*, Verification Annex.

41. Comprehensive Nuclear Test Ban Treaty, Sept. 10, 1996, arts. II, IV, 35 INTERNATIONAL LEGAL MATERIALS 1439 (not yet in force) [hereinafter CTBT]; *id,* Protocol [hereinafter CTBT Protocol].

42. Statute of the International Atomic Energy Agency art. XII(6), Oct. 23, 1956, 276 U.N.T.S 3.

43. CTBT Protocol, *supra* note 41, pt. I.

44. CWC, *supra* note 40, Verification Annex, pt. III(B)(10).

45. *Id.* art. X.

46. G.A. Res. 42/37C, Chemical and Bacteriological (Biological) Weapons (Nov. 30, 1987). The Secretary-General resorted to this mechanism in 2013 in connection with allegations that chemical weapons were used in Syria. The inspections, conducted by experts from the OPCW and World Health Organization, were authorized only to determine whether chemical weapons were used, without determining who used them. David Martin, *The Chemical Weapons Convention: Hollow Idealism or Capable Mechanism? The Syrian Intervention as a Test Case*, 37 LOYOLA OF LOS ANGELES INTERNATIONAL AND COMPARATIVE LAW REVIEW 31, 50 (2015).

47. CTBT, *supra* note 41, art. IV(D). See in particular *id.* ¶ 35 ("The sole Purpose of an on-site inspection shall be to clarify whether a nuclear weapon test explosion or any other nuclear explosion has been carried out in violation of Article I and, to the extent possible, to gather any facts which might assist in identifying any possible violator.").

by the IAEA.[48] In such cases, information collected and analyzed by verification mechanisms can include, for instance, documents, technical data, samples, and interviews.[49]

In that verification mechanisms have the potential to embarrass violators[50] and set international legal and political processes of condemnation and sanction in motion,[51] they are integral to the stability of arms control regimes. Such mechanisms not only deter would-be violators by enhancing the likelihood of detection,[52] but also provide a means for States to refute unfounded allegations.

The recent resort by the United Kingdom to the OPCW inspection machinery in connection with the Salisbury incident illustrates how a technical fact-finding apparatus can be used to shame a violating State and support accountability. It also casts new light on some of the principled arguments discussed below, which question the necessity of developing an international accountability mechanism in the field of cyber security.

On March 4, 2018, two Russian nationals were poisoned in Salisbury through exposure to a rare nerve agent. Eight days later, U.K. Prime Minister Theresa May announced in the House of Commons that the poison was a

---

48. *See, e.g.*, International Atomic Energy Agency [IAEA], *The Text of the Agreement between Iran and the Agency for the Application of Safeguards in Connection with the Treaty on the Non-Proliferation of Nuclear Weapons*, art. 73, at 17, IAEA Doc. INFCIRC/214 (Dec. 13, 1974); *see also* Wolfgang Fischer & Gotthard Stein, *On-Site Inspections: Experiences from Nuclear Safeguarding*, 3 DISARMAMENT FORUM 45, 49–50 (1999).

49. *See, e.g.*, CTBT Protocol, *supra* note 41, at IV(D), ¶35; CWC, *supra* note 40, Verification Annex, pt. XI(D), at 25–26

> The final report shall summarize the factual findings of the inspection, particularly with regard to the alleged use cited in the request. . . . If the inspection team collects through, inter alia, identification of any impurities or other substances during laboratory analysis of samples taken, any information in the course of its investigation that might serve to identify the origin of any chemical weapons used, that information shall be included in the report.

But see Martin, *supra* note 46, at 50.

50. TREVOR FINDLAY, PROLIFERATION ALERT! THE IAEA AND NON-COMPLIANCE REPORTING 73 (2015) (discussing the special inspection of the Republic of Korea).

51. Thilo Marauhn, *Consultations, Cooperation and Fact-Finding*, *in* THE CHEMICAL WEAPONS CONVENTION: A COMMENTARY 297, 325 (Walter Krutzsch, Eric Myjer & Ralf Trapp eds., 2014); David Cortright, Linda Gerber & George A. Lopez, *Implementing Targeted Sanctions: The Role of International Agencies and Regional Organizations*, *in* INTERNATIONAL SANCTIONS: BETWEEN WARS AND WORDS 144, 146 (Peter Wallensteen & Carina Staibano eds., 2005).

52. Jenifer Mackby, *Nonproliferation Verification and the Nuclear Test Ban Treaty*, 34 FORDHAM INTERNATIONAL LAW JOURNAL 697, 707 (2011); SARAH J. DIEHL & JAMES CLAY MOLTZ, NUCLEAR WEAPONS AND NON-PROLIFERATION 50 (2002).

military-grade nerve agent of a type developed by Russia known as "Novichok."[53] The United Kingdom then took retaliatory measures against Russia, including the expulsion of twenty-three Russian diplomats.[54] The United Kingdom also looked to the OPCW for help in investigating the incident, a decision welcomed by the EU.[55] Following a technical assistance visit, the organization released a public summary of its findings on April 12.[56] The full confidential report, which reportedly corroborated the U.K. law enforcement findings concerning the nerve agent, was circulated to OPCW member States.[57] On July 13, the United Kingdom asked again for OPCW assistance in investigating a poisoning incident, this time in Amesbury; the resulting report matched the Amesbury toxic chemicals with those found in Salisbury.[58]

A central issue for the purposes of this article is the United Kingdom's motivation in seeking OPCW assistance. That the OPCW merely confirmed the United Kingdom's findings demonstrates that the organization was not filling a capability gap. Indeed, on March 27, even before completion of the OPCW investigation, twenty of the United Kingdom's Western partners moved to expel Russian diplomats (more than one hundred in total) based

53. Prime Minister Theresa May, Oral Statement to Parliament: PM Commons Statement on Salisbury Incident: 12 March 2018 (Mar. 12, 2018), https://www.gov.uk/government/speeches/pm-commons-statement-on-salisbury-incident-12-march-2018.

54. *Russian Spy: UK to Expel 23 Russian Diplomats*, BBC (Mar. 14, 2018), https://www.bbc.com/news/uk-43402506.

55. Press Release, Council of the European Union, Statement by the Foreign Affairs Council on the Salisbury Attack (Mar. 19, 2018), https://www.consilium.europa.eu/en/press/press-releases/2018/03/19/statement-by-the-foreign-affairs-council-on-the-salisbury-attack/.

56. Technical Secretariat, OPCW, *Note by the Technical Secretariat: Summary of the Report on Activities Carried Out in Support of a Request for Technical Assistance by the United Kingdom of Great Britain and Northern Ireland (Technical Assistance Visit TAV/02/18)*, S/1612/2018, Apr. 12, 2018, https://www.opcw.org/sites/default/files/documents/S_series/2018/en/s-1612-2018_e_.pdf.

57. *Chemical Watchdog Confirms UK Findings on Salisbury Nerve Agent*, UN NEWS (Apr. 18, 2018), https://news.un.org/en/story/2018/04/1007642.

58. Technical Secretariat, OPCW, *Note by the Technical Secretariat: Summary of the Report on Activities Carried Out in Support of a Request for Technical Assistance by the United Kingdom of Great Britain and Northern Ireland (Technical Assistance Visit TAV/03/18 and TAV/03B/18 "Amesbury Incident")*, S/1671/2018, Sept. 4, 2018, https://www.opcw.org/sites/default/files/documents/2018/09/s-1671-2018%28e%29.pdf.

on those findings.[59] Rather, it appears the United Kingdom leveraged the organization to enhance the credibility of its findings regarding the poison, thereby enabling it and other States to name, blame, and shame Russia more effectively.[60]

Similarly, in a speech at the OPCW on June 26, 2018, U.K. Foreign Minister Boris Johnson stressed the need to further enhance accountability by authorizing the Technical Secretariat to attribute responsibility for chemical attacks in the context of Syria.

> Our aim . . . is to reinforce the OPCW as an institution. Last November, the Security Council was prevented from renewing the Joint Investigative Mechanism, meaning that no international body is working to attribute responsibility for chemical weapons attacks in Syria. At present, the OPCW's experts will say where and when an attack happened, but not who was responsible. If we are serious about upholding the ban on chemical weapons, that gap must be filled. Attributing responsibility for an attack is clearly part of the OPCW's technical remit, requiring no change to the Chemical Weapons Convention. The Director General has confirmed that the OPCW is able and willing to perform this essential task.[61]

In a decision adopted by the Conference of State Parties the next day, authorization was granted. Reaffirming that "those responsible for the use of chemical weapons should be held accountable," the Conference decided that:

---

59. Julian Borger, Patrick Wintour & Heather Stewart, *Western Allies Expel Scores of Russian Diplomats over Skripal Attack*, GUARDIAN, Mar. 27, 2018, https://www.theguardian.com/uk-news/2018/mar/26/four-eu-states-set-to-expel-russian-diplomats-over-skripal-attack.

60. Interestingly enough, on March 13, 2018, the Russians called for the United Kingdom to involve the OPCW and to request their technical assistance in response to accusations by the United Kingdom, on the basis that they were not involved in the Salisbury incident. Executive Council, OPCW, *Russian Federation: Statement by H.E. Ambassador A. V. Shulgin Permanent Representative of the Russian Federation to the OPCW at the Eighty-Seventh Session of the Executive Council (On the Chemical Incident in Salisbury)*, EC-87/NAT.9, Mar. 13, 2018, https://www.opcw.org/sites/default/files/documents/EC/87/en/ec87nat09_e_.pdf.

61. Boris Johnson, U.K. Foreign Secretary, Foreign Secretary's Speech at the OPCW Special Conference of the States Parties (June 26, 2018), https://www.gov.uk/government/speeches/foreign-secretarys-speech-at-the-opcw-special-conference-of-the-states-parties. For criticism of the initiative, *see* Oliver Meier & Ralf Trapp, *Playing Politics with Chemical Weapons? The UK's Initiative on Chemical Weapons Accountability*, BULLETIN OF THE ATOMIC SCIENTISTS (June 20, 2018), https://thebulletin.org/2018/06/playing-politics-with-chemical-weapons-the-uks-initiative-on-chemical-weapons-accountability/.

> the Director-General, if requested by a State Party investigating a possible chemical weapons use on its territory, can provide technical expertise to identify those who were perpetrators, organizers, sponsors or otherwise involved in the use of chemicals as weapons, and further decide[d] that, in this context, the Director-General may enlist support as appropriate from outside experts with relevant qualifications and professional experience, and invites the Director-General to submit to the Conference at its next regular session specific proposals to establish such independent, impartial, expert arrangements.[62]

The day after the Conference acted, the EU Council expressed support for the implementation of the decision and for the development of a chemical weapons sanctions regime.[63] Such a regime subsequently was characterized by the United Kingdom as an effective means of holding individuals and entities responsible for the proliferation and use of chemical weapons.[64] Further Council support for "international initiatives aimed at tackling the threat of chemical weapons" came in October with the implementation of a sanctions regime.[65]

To summarize, flagrant violations of international law norms prohibiting the employment of chemical weapons initiated an effort to impose accountability through the use of an international fact-finding mechanism and by authorizing it to attribute legal responsibility. These moves were accompanied by ad hoc sanctions against suspected perpetrators and the development of a standing sanctions mechanism. The question is whether a similar dynamic involving international fact-finding, attribution of international responsibility, and the imposition of collective sanctions would be viable in

62. Conference of the States Parties, OPCW, *Decision: Addressing the Threat from Chemical Weapons Use*, at 4, ¶ 20, C-SS-4/DEC.3, June 27, 2018, https://www.opcw.org/sites/default/files/documents/CSP/C-SS-4/en/css4dec3_e_.doc.pdf.

63. Press Release, European Council, European Council Conclusions, 28 June 2018 (June 29, 2018), https://www.consilium.europa.eu/en/press/press-releases/2018/06/29/20180628-euco-conclusions-final/.

64. Press Release, Foreign and Commonwealth Office, United Kingdom, Foreign Secretary Urges EU to Press Ahead with Listings Under New Chemical Weapons Sanctions Regime (Oct. 14, 2018), https://www.gov.uk/government/news/foreign-secretary-urges-eu-to-press-ahead-with-listings-under-new-chemical-weapons-sanctions-regime. The statement also alluded to the need for a cyber-related sanctions regime.

65. Council Decision 2018/1544 of 15 October 2018, Concerning Restrictive Measures against the Proliferation and Use of Chemical Weapons, 2018 O.J. (L 259) 25.

other fields of activity involving national security that pose technical attribution challenges and face denials by responsible States, specifically cyber operations.

III. THE PRESENT STATE OF CYBER ATTRIBUTION

States that have fallen victim to hostile cyber operations are increasingly willing to attribute them to other States.[66] Moreover, attribution is often collective in the sense that it involves the issuance of a common statement or endorsement of another State's assertion of responsibility. The collective attributions of the WannaCry[67] and NotPetya[68] cyber operations, as well as hostile cyber operations targeting the OPCW[69] and Georgia,[70] are illustrative of this developing practice.

---

66. For a survey of eleven prominent cyber operations that occurred between 2012 and 2017 in which State involvement was suspected, see Dan Efrony & Yuval Shany, *A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyberoperations and Subsequent State Practice,* 112 AMERICAN JOURNAL OF INTERNATIONAL LAW 583, 655–57 (2018). According to the survey, five cases did not result in an official statement by the victim State pointing the finger at the suspected attacking State, and in one case (the attack on the Bundestag) the statements made were somewhat uncertain. Furthermore, the survey suggests an increase in the rate of attributions during the surveyed period. *See id.*; *see also* Kristen E. Eichensehr, *The Law & Politics of Cyberattack Attribution*, 67 UCLA LAW REVIEW (forthcoming 2020) (including under the definition of attribution not only official statements, but also indictments and technical alerts).

67. Press Briefing, The White House, The Attribution of the WannaCry Malware Attack to North Korea (Dec. 19, 2017), https://www.whitehouse.gov/briefings-statements/press-briefing-on-the-attribution-of-the-wannacry-malware-attack-to-north-korea-121917/ (attributing the malware attack to North Korea and indicating that the United Kingdom, Canada, Australia, New Zealand, and Japan join the statement).

68. Stilgherrian, *Blaming Russia for NotPetya was Coordinated Diplomatic Action*, ZDNET (Apr. 12, 2018), https://www.zdnet.com/article/blaming-russia-for-notpetya-was-coordinated-diplomatic-action/ (discussing the coordination of collective attribution by the United Kingdom, United States, Denmark, Lithuania, Estonia, Canada, Australia, New Zealand, Norway, Latvia, Sweden, and Finland).

69. *Netherlands Defence Intelligence and Security Service Disrupts Russian Cyber Operation Targeting OPCW*, GOVERNMENT OF THE NETHERLANDS (Oct. 4, 2018), https://www.government.nl/latest/news/2018/10/04/netherlands-defence-intelligence-and-security-service-disrupts-russian-cyber-operation-targeting-opcw (noting that the attribution is supported by the United Kingdom and that the United States has opened a criminal investigation against the implicated GRU officers).

70. Przemyslaw Roguski, *Russian Cyber Attacks against Georgia, Public Attributions and Sovereignty in Cyberspace,* JUST SECURITY (Mar. 6, 2020), https://www.justsecurity.org/69019/

There has also been some progress in developing a structure for collective attribution. Of particular note, the EU Cyber Diplomacy Toolbox,[71] which was adopted in 2017, provides for a joint EU diplomatic response to hostile cyber operations. It is premised on the belief that "clearly signaling the likely consequences of a joint EU diplomatic response to . . . malicious cyber activities influences the behavior of potential aggressors in cyberspace thus reinforcing the security of the EU and its Member States."[72] As part of the broader EU Cyber Diplomacy Toolbox, the EU established a cyber-related sanctions regime in 2019 that provides for the imposition of "targeted restrictive measures" on natural and legal persons.[73] Both the Toolbox and the sanctions regime clarify, however, that a joint diplomatic response or the imposition of sanctions should be distinguished from a decision to attribute responsibility to a foreign State, which is described as "a sovereign political decision taken on a case-by-case basis,"[74] one to be "based on all-source intelligence and . . . in accordance with [the] international law of State responsibility."[75]

Other States and international and regional organizations are also crafting collective responses to hostile cyber operations. In 2018, the United States announced an International Cyber Deterrence Initiative aimed at building a coalition of like-minded States that can act "in concert" to impose "consequences" on adversaries, so as to ensure that they "understand the consequences of their malicious cyber behavior."[76] The 2018 U.S. National Cyber Strategy also envisions intelligence sharing with key partners to identify hostile State and non-State cyber activities.[77] The same year, NATO leaders adopted the Brussels Summit Declaration, which confirmed that

---

russian-cyber-attacks-against-georgia-public-attributions-and-sovereignty-in-cyberspace/ (noting that more than twenty countries attributed the attacks to Russia).

71. General Secretariat of the Council, Council of the European Union, Draft Council Conclusions on a Framework for a Joint EU Diplomatic Response to Malicious Cyber Activities ("Cyber Diplomacy Toolbox") – Adoption, 9916/17, June 7, 2017 [hereinafter Cyber Diplomacy Toolbox].

72. *Id*., annex, at 5, ¶ 4.

73. Council Decision (CFSP) 2019/797 of 17 May 2019, Concerning Restrictive Measures against Cyber-Attacks Threatening the Union or its Member States, 2019 O.J. (L 129) 13 (EC); Cyber Diplomacy Toolbox, *supra* note 71.

74. Council Decision (CFSP) 2019/797, pmbl., ¶ 9.

75. *See* Cyber Diplomacy Toolbox, *supra* note 71, annex, at 5, ¶ 4.

76. THE WHITE HOUSE, NATIONAL CYBER STRATEGY OF THE UNITED STATES OF AMERICA 21 (2018).

77. *See id.* The Strategy also stipulates that the U.S. intelligence community should "continue to lead the world in the use of all-source cyber intelligence to drive the identification

NATO's collective defense policies applied to cyberspace and called on its members to consider responding to malicious cyber activity in a coordinated manner.[78]

It should be noted that attribution is seldom accompanied in practice by the release of the underlying evidence, despite broad international support for the principle that States should, where possible, provide the basis for that attribution. For instance, the 2015 U.N. Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, which included all permanent members of the Security Council and whose consensus report was endorsed by the General Assembly, noted that "accusations of organizing and implementing wrongful acts brought against States should be substantiated."[79] Still, the responsibility to release the evidence underlying attribution was styled as a voluntary, non-binding norm of responsible State behavior, not as a legal obligation.

Indeed, key States, such as the United States and the United Kingdom, have taken the position that there is no legal duty to accompany public acts

---

and attribution of malicious cyber activity that threatens United States national interests." *Id.*

78. Press Release, NATO, Brussels Summit Declaration, Issued by the Heads of State and Government Participating in the Meeting of the North Atlantic Council in Brussels 11–12 July 2018, Press Release (2018) 074 (July 11, 2018) (last updated Aug. 30, 2018), https://www.nato.int/cps/en/natohq/official_texts_156624.htm

> Cyber defense is part of NATO's core task of collective defense. We must be able to operate as effectively in cyberspace as we do in the air, on land, and at sea to strengthen and support the Alliance's overall deterrence and defense posture. We therefore continue to implement cyberspace as a domain of operations. . . . We continue to work together to develop measures which would enable us to impose costs on those who harm us. Individual Allies may consider, when appropriate, attributing malicious cyber activity and responding in a coordinated manner, recognizing attribution is a sovereign national prerogative. We are determined to deliver strong national cyber defenses through full implementation of the Cyber Defence Pledge, which is central to enhancing cyber resilience and raising the costs of a cyber-attack. We reaffirm our commitment to act in accordance with international law, including the UN Charter, international humanitarian law, and human rights law, as applicable.

79. Rep. of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (2015), transmitted by Letter Dated 22 July 2015 from the Secretary-General Comm. Established Pursuant to Resolution 68/243 (2014) Addressed to the General Assembly, ¶ 28(f), U.N. Doc. A/70/174 (July 22, 2015); G.A. Res. 70/237 (Dec. 30, 2015) (endorsing the report).

of attribution with disclosure of any underlying evidence.[80] Although there are various reasons for the reluctance to commit to releasing evidence, the most commonly cited is that such a practice can risk revealing intelligence sources and methods and cyber capabilities. Yet, absent supporting evidence, the credibility of public attribution is open to challenge. Likewise, collective attribution is less likely when intelligence on an incident is not shared.

Evidentiary issues might also hamper regional mechanisms for cyber-related sanctions. For example, the EU cyber-related sanctions regime is based on a list of natural or legal persons responsible for hostile cyber operations. Inclusion on the list by the Council requires the proposal of a member State or the EU High Representative for Foreign Affairs and Security Policy. Although those targeted by restrictive measures may submit observations leading to reevaluation, the sanctions regime neither specifies the requisite evidentiary threshold for inclusion, nor mandates the sharing of the underlying intelligence with other member States.[81]

It is worth noting that some commentators have expressed doubt as to the effectiveness of recent collective attribution statements, noting the limited number of States involved, a frequent lack of transparency surrounding the process of attribution, the failure to identify specific international law obligations that the operations breached, and the lack of political will in following up with the imposition of sanctions and other responses against the responsible State.[82] Arguably, these shortcomings hinder the development of substantive law in the field of cyber security, since they provide few indications of those cyber operations that States consider unlawful—information that is essential in both the interpretation of existing legal rules in the cyber context and the crystallization of new rules of customary international law.[83]

---

80. Brian J. Egan, *International Law and Stability in Cyberspace*, 35 BERKELEY JOURNAL OF INTERNATIONAL LAW 169, 177 (2017); Wright, *supra* note 4; *see also* TALLINN MANUAL 2.0, *supra* note 3, at 83.

81. Such evidentiary standards must meet, however, the requirements set by the Court of Justice of the European Union for application of individual sanctions. *See, e.g.*, Case C-176/13 P, Council v. Bank Mellat, ¶127, EU:C:2016:96 (Fifth Chamber) (E.C.J.).

82. *See* Anushka Kaushik, *Public Attribution and Its Scope and Efficacy as a Policy Tool in Cyberspace*, OBSERVER RESEARCH FOUNDATION (ORF) (Oct. 21, 2019), https://www.orfonline.org/expert-speak/public-attribution-and-its-scope-and-efficacy-as-a-policy-tool-in-cyberspace-56826/; *see also* Roguski, *supra* note 70.

83. Eichensehr, *supra* note 66, at 30; Finnemore & Hollis, *supra* note 5, at 12.

IV. PROPOSALS FOR INTERNATIONAL ATTRIBUTION MECHANISMS

Against this backdrop, there have been numerous calls for the establishment of an international attribution mechanism that would foster public confidence in the attribution claims of national security agencies and private cyber security companies. Although such public and private bodies have considerable professional expertise, their work tends to lack transparency,[84] and their governmental affiliation or commercial interests sometimes render their claims suspect.[85]

An independent international mechanism could lend credibility to attribution in the cyber realm, thereby limiting the ability of responsible States to deny involvement and facilitating collective attribution and response. Like the OPCW Technical Secretariat, such a mechanism could prove useful for State and non-State actors in certain situations.[86] To be sure, such a mechanism should complement, not replace, existing attribution mechanisms and practices. Optimally, it should find ways to harness the evidence gathering and analytical wherewithal of State agencies, and the technical expertise resident in the public and private cyber security sectors.

Several initiatives aimed at promoting an international attribution mechanism have been launched in recent years. In 2014, the Atlantic Council, a Washington, D.C.-based think tank, proposed the establishment of a Multilateral Cyber Adjudication and Attribution Council (MCAAC).[87] The MCAAC would serve as an inter-State body entrusted with investigating cyber incidents with a view to attributing responsibility to States when appropriate. It could then issue recommendations for de-escalation or refer the matter to other political or adjudicative bodies. The Atlantic Council's proposal envisioned the gradual development of such a mechanism as ad hoc coalitions of States undertook joint investigations of hostile cyber operations, sometimes in collaboration with private actors.

---

84. Florian J. Egloff & Andreas Wenger, *Public Attribution of Cyber Incidents*, 244 CSS ANALYSES IN SECURITY POLICY, no. 2, May 2019, at 1, 1.

85. Sasha Romanosky & Benjamin Boudreaux, *Private Sector Attribution of Cyber Incidents: Benefits and Risks to the U.S. Government* 20 (RAND Working Paper WR-1267-OSD, 2019).

86. One concern raised in this context by Eichensehr is the proliferation of attributions. Eichensehr, *supra* note 66. Still, creating a new centralized attribution institution may, over time, limit reliance on decentralized and multiple attribution bodies.

87. JASON HEALEY ET AL., ATLANTIC COUNCIL, CONFIDENCE-BUILDING MEASURES IN CYBERSPACE: A MULTISTAKEHOLDER APPROACH FOR STABILITY AND SECURITY 10–12 (2014).

In 2016, Microsoft published a proposal for an attribution premised on the investigation of incidents by international experts comprising a public-private organization.[88] The organization would conduct a technical investigation of certain operations that fell within its purview and report its findings, including evidence of attribution. Only in certain cases would the findings be published publicly. The proposal envisioned the mechanism as also offering a form of peer review of attribution claims made by other public or private entities.

The following year, the RAND Corporation proposed a Stateless attribution mechanism consisting of a consortium of private experts specializing in cyber technology and policy that would, on a discretionary basis, investigate and attribute incidents, as well as provide analysis concerning the severity of the incident and the sophistication of the operation. While the consortium would not be involved in follow-up action, other stakeholders might use its findings for that purpose.[89]

Eventually, in 2019, Microsoft joined forces with Mastercard and the Hewlett Foundation to establish the CyberPeace Institute.[90] The Institute's mandate includes promoting accountability in cyberspace by facilitating collaborative research into the behavior of those launching hostile cyber operations and ways to fend off such operations, and through publishing information on the techniques, practices, and effects of attack tools. In other words, it is less an attribution mechanism than an entity producing information that can help other entities effectively attribute.

With the exception of the CyberPeace Institute, which has been established, neither these nor other initiatives[91] have gained much momentum. Based on discussions during the Federmann Cyber Security Research Center's workshops with legal, policy, and technical experts, diplomats and other State officials, academics, and industry executives, several tentative reasons for this lack of progress can be identified.

To begin with, and perhaps most significantly, some major State actors in the field of cyber security appear uninterested in developing an international attribution mechanism, largely out of a sense that the mechanism

88. *See* CHARNEY ET AL., *supra* note 13.

89. JOHN S. DAVID II ET AL., STATELESS ATTRIBUTION: TOWARD INTERNATIONAL ACCOUNTABILITY IN CYBERSPACE (2017).

90. *See supra* note 15.

91. *See, e.g.*, SERGE DROZ & DANIEL STAUFFACHER, ICT FOR PEACE FOUNDATION, TRUST AND ATTRIBUTION IN CYBERSPACE: A PROPOSAL FOR AN INDEPENDENT NETWORK OF ORGANISATIONS ENGAGING IN ATTRIBUTION PEER-REVIEW (2018).

would be redundant. After all, powerful and technologically savvy States have developed processes for technical attribution that rely upon their own technical forensic capacity, as well as their intelligence assets, especially signals intelligence and human intelligence. When required, they collaborate with their partners, some of whom also field impressive capabilities to perform attribution; occasionally, they even turn to private cyber security companies to provide specialized expertise. Once such States can attribute, they have the offensive tools and the political and economic clout to respond meaningfully to hostile cyber operations, either alone or in collaboration with other States.

Moreover, some States appear skeptical about the very push for legal accountability. Arguably, the proposed mechanism would help to clarify the law applicable to cyber operations, thereby limiting the operational flexibility that results from legal ambiguity. In the view of skeptical States, legal clarity constrains the cyber operations of "rule of law" States, but proves ineffective in limiting operations by adversaries that do not share that commitment to the rule of law. Viewing international law as asymmetrically disadvantageous, these States would prefer to rely on self-help, like robust cyber defenses, offensive tools, and credible warnings, rather than international law, to safeguard their cyberspace.[92] They fear that referral of incidents to an international attribution mechanism might over time deprive them of the discretion ambiguity offers in terms of attribution and response options.[93]

Some workshop participants also suggested the initiatives have a number of shortcomings that have impeded acceptance. For instance, they opined that the Atlantic Council proposal is short on detail and that a few of its features—especially vis-à-vis attribution follow-up—could be regarded as overly ambitious. Its inter-State aspects also pose a risk of politicization.[94] The Microsoft proposal is likewise lacking in detail. Further, in that it proffered a hybrid institution, States may have been apprehensive about sharing the allocation of State responsibility with non-State actors. The fact that the proposal was put forward by a powerful global commercial entity also generated some skepticism, fair or not, as to Microsoft's motives. The RAND proposal is more detailed but envisages a Stateless mechanism over whose

92. *See* Michael N. Schmitt, *Grey Zones in the International Law of Cyberspace*, 42 YALE JOURNAL OF INTERNATIONAL LAW ONLINE 1 (2017).

93. *See* Herbert Lin, *Attribution of Malicious Cyber Incidents: From Soup to Nuts*, 70 JOURNAL OF INTERNATIONAL AFFAIRS 75 (2017).

94. *See, e.g.*, Mueller et al., *supra* note 15, at 117.

configuration, mandate, and *modus operandi* States would have limited influence. This arrangement is unlikely to appeal to States, especially when comparable investigative services are available from private companies.

Finally, numerous participants were of the view that the case for an international attribution mechanism would have been strengthened had the initiatives more directly identified the specific constituencies likely to be effectively served and the contexts in which the proposals would prove most useful. In other words, the proposals could be characterized as overbroad in the sense that they called for a major restructuring of cyber attribution and the underlying concept of accountability. Thus, some participants argued that a more sophisticated approach would have been to associate the proposed mechanism with specific needs—a need to increase global capacity to make credible attribution claims, a need to encourage collective attributions, and a need to support multilateral follow-up efforts.

V. THE WAY FORWARD

Effective application of international law to any domain of international relations hinges on the interaction between legal norms, fact-finding processes that identify violations and attribute responsibility to a State or non-State actor, and follow-up measures, which can include shaming, making claims in diplomatic or adjudicative fora, and imposing sanctions, including countermeasures.[95] The legitimacy of each link in the chain is premised on the legitimacy of the preceding links; thus, the legitimacy of accusations and responses depends on the legitimacy of the underlying legal rules and attribution process.[96] Specifically, the legitimacy of the attribution process undergirds the ability of an accuser to convince relevant target audiences, including third States that might join collective attribution statements or support multilateral sanctions.[97]

As explained above, fact-finding provides a basis for legal and political claims. Thus, fact-finding mechanisms are especially useful in fields where

95. On countermeasures under the law of State responsibility, see Articles on State Responsibility, *supra* note 8, art. 22, at 75–76; *see also* Michael N. Schmitt, *"Below the Threshold" Cyber Operations: The Countermeasures Response Option and International Law*, 54 VIRGINIA JOURNAL OF INTERNATIONAL LAW 697 (2014).

96. Egloff & Wenger, *supra* note 84, at 2.

97. Sven Herpig & Thomas Reinhold*, Spotting the Bear: Credible Attribution and Russian Operations in Cyberspace*, *in* HACKS, LEAKS AND DISRUPTIONS RUSSIAN CYBER STRATEGIES 33, 40 (Nicu Popescu & Stanislav Secrieru eds., 2018).

legal compliance is chronically deficient, such as human rights law, where there is a special need to further legitimize norm-implementation efforts and increase deterrence, or where unique technical challenges present themselves, as in the case of unconventional weapons. Indeed, the use of the OPCW's technical facilities in connection with the Salisbury incident illustrates the benefits of fact-finding in cases requiring complex scientific analysis, as well as the growing interest of States in internationalizing attribution processes.

Moving forward, we maintain that the goal of creating an international attribution mechanism remains viable and that such an entity would prove valuable, albeit primarily in three contexts. First, an international attribution mechanism could prove useful for States with a limited independent capacity to effectively generate accountability. This includes States that, on the one hand, are sufficiently advanced technologically so as to render them highly vulnerable to hostile cyber operations, but, on the other hand, lack the technological capability to conduct their own forensic investigations[98] and access to high-quality intelligence that can support attribution to a State actor.[99] In this context, an international attribution mechanism could help level the playing field.

Furthermore, small and mid-size States often lack the diplomatic, economic, technical, or military means to effectively respond, when appropriate, to hostile cyber operations,[100] thereby making them dependent on their ability to mobilize other States to support them. Yet, relying on other States or private cybersecurity companies for attribution assistance risks having the information provided dismissed as politically biased or profit driven. As a result, it might not generate the level of legitimacy needed to mobilize third States.

An international attribution mechanism could prove especially useful in this regard. A properly crafted mechanism would be more likely to be per-

98. Eichensehr, *supra* note 66, at 57–58.

99. KARINE BANNELIER & THÉODORE CHRISTAKIS, CYBER-ATTACKS – PREVENTION-REACTIONS: THE ROLE OF STATES AND PRIVATE ACTORS 45 (2017). *But see* Jason Rivera, *Achieving Cyberdeterrence and the Ability of Small States to Hold Large States at Risk*, *in* ARCHITECTURES IN CYBERSPACE 7, 14–15 (M. Maybaum, A. M. Osula & L. Lindström eds., 2015) (claiming that technical attribution suffices to demand some response from the territorial State).

100. Joe Burton, *Small States and Cyber Security: The Case of New Zealand*, 65 POLITICAL SCIENCE 216, 224 (2013).

ceived as independent, impartial, and professionally reliable. Its factual conclusions would presumably enjoy significant traction, which would enhance the victim State's ability to generate assistance in imposing accountability.

Second, the creation of an international attribution mechanism would signal the growing interest of States in collective attribution, as broad condemnation and multilateral responses are more likely to promote accountability than the reactions of a single State. Collective attribution is most effective when there is a high level of confidence in the initial attribution determination. Such confidence can derive from close relations among relevant States, such as ongoing cooperation between their intelligence agencies, or through legitimacy-enhancing measures, like transparency.

But in our estimation, an international attribution mechanism would be especially likely to foster collective attribution because of the legitimacy it would presumably enjoy.[101] In particular, a credible attribution process would empower States supportive of the rule of law in international relations to take a principled stand beside the victim State.[102] The same logic that led to the integration of fact-finding mechanisms in the realm of human rights law to publicly name and shame human rights violators and, in certain cases, to mobilize third States to sanction them applies *mutatis mutandis* here.

Third, an international attribution mechanism could play an important role in connection with the operation of cyber-related sanction regimes, such as that of the EU. Such regimes rely upon trust in individual attribution determinations by member States. However, since they involve sanctions on natural or legal actors that may be associated with foreign States, the smooth operation of the regime generally requires a higher level of confidence than might otherwise be the case with collective attribution.

An independent professional mechanism could offer verification of individual attribution claims, assuaging the concerns of States about the reliability of the attribution determinations upon which they are being asked to take action. Moreover, since cyber-related sanction regimes like the EU's have implications under the domestic law of the member States, such as the freezing of assets and travel restrictions, domestic courts may examine the evidence underlying the sanction decision or seek credible assurance that the fact-finding process was independent, fair, and provided those involved an

101. *Cf.* TIMOTHY O'RIORDAN, RAY KEMP & MICHAEL PURDUE, SIZEWELL B: AN ANATOMY OF INQUIRY 84 (1988).

102. *Cf.* Antonio Cassese & Andrew Clapham, *International Law*, *in* THE OXFORD COMPANION TO POLITICS OF THE WORLD 408, 409 (Joel Krieger ed., 2d ed. 2001).

opportunity to contest the findings against them.[103] An international attribution mechanism could, should it offer features along these lines, help satisfy this demand.

In sum, the configuration of the international attribution mechanism must be responsive to the goal of supporting States facing capacity issues, those interested in collective attribution, and cyber-related sanctions regimes. We believe additional research is called for in order to enumerate and ascertain specific "client" preferences and expectations. Key institutional design choices would emanate from such mapping of needs. These include (1) whether the optimal composition is public, private, or hybrid in nature; (2) the triggers to initiation of the mechanism's investigation; (3) the extent to which a mechanism should be tasked with responsibility beyond technical attribution, such as attribution under the law of State responsibility; (4) the necessary arrangements for access to forensic evidence and intelligence materials, including confidential sharing of information; (5) the opportunity of entities to whom attribution is made to contest evidence collected against them; and (6) whether, and if so when, attribution decisions and supporting evidence should be made public.

VI. CONCLUDING THOUGHTS

International fact-finding is a venerable institution. It offers a credible process for ascertaining facts underlying international incidents, and, sometimes, the attribution of legal responsibility for violation of international law norms. Fact-finding mechanisms have been utilized extensively in certain fields of international law, particularly human rights and weapons control, to generate accountability and deterrence. In such fields, it can play an integral role in implementing the rule of law and developing and interpreting relevant international law norms.

We believe there is merit in the prospect of an independent international attribution mechanism for cyber operations, one along the lines of, but not necessarily identical to, the OPCW's Technical Secretariat. Cyber operations represent a field of activity plagued by normative ambiguity and limited accountability, where reliance on the victim State's attribution capacities, or those of other States or private cybersecurity companies, may not measure up to the challenges. An independent attribution mechanism could lead to

103. *Cf.* Combined Cases C-402/05 P and C-415/05 P, Kadi v. Council and Comm'n, ¶ 352, 2008 E.C.R. I-6351.

attribution determinations enjoying a higher degree of legitimacy, thereby serving as a stabilizing force in international relations.

To date, proposals to establish an international attribution mechanism have not acquired momentum, either because they contain features that States find unattractive or because they were not developed to a degree that made States comfortable. However, we believe progress remains possible by focusing on the three logical constituencies for such a body—States with limited technological, intelligence, and diplomatic capacity; States interested in generating broad collective attribution of attacks perpetrated against them; and international and regional organizations operating a cyber-related sanctions regime. Such a focus, combined with greater granularity, would significantly improve the prospects for the establishment of an international attribution mechanism and its eventual utilization by the international community.