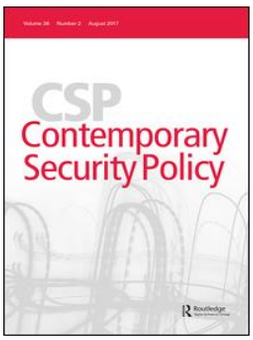

Routledge
Taylor & Francis Group

Contemporary Security Policy

ISSN: 1352-3260 (Print) 1743-8764 (Online) Journal homepage: https://www.tandfonline.com/loi/fcsp20

Contested public attributions of cyber incidents and the role of academia

Florian J. Egloff

To cite this article: Florian J. Egloff (2019): Contested public attributions of cyber incidents and the role of academia, Contemporary Security Policy, DOI: 10.1080/13523260.2019.1677324

To link to this article: https://doi.org/10.1080/13523260.2019.1677324

Published online: 12 Oct 2019.

CONTEMPORARY SECURITY POLICY
https://doi.org/10.1080/13523260.2019.1677324

Routledge
Taylor & Francis Group

Contested public attributions of cyber incidents and the role of academia

Florian J. Egloff [a,b]

[a]Center for Security Studies, ETH Zürich, Zürich, Switzerland; [b]Centre for Technology & Global Affairs, Department of Politics and International Relations, University of Oxford, Oxford, UK

ABSTRACT

Public attributions of cyber incidents by governments and private industry have become prevalent in recent years. This article argues that they display a skewed version of cyber conflict for several operational and structural reasons, including political, commercial, and legal constraints. In addition, public attribution of cyber incidents takes place in a heavily contested information environment, creating fractured narratives of a shared past. The article uses three cyber incidents (Sony Pictures, DNC, and NotPetya) to show how actors cope with this contested information environment and proposes a changed role of academia to address some of the problems that emerge. To become competent in contesting public attribution discourses, universities would have to work more across physical, disciplinary, and academic boundaries. The main implications for democracies are to be more transparent about how attribution is performed, enable other civilian actors to study cyber conflict, and thereby broaden the discourse on cybersecurity politics.

KEYWORDS Cybersecurity; attribution; threat intelligence; legitimacy; information warfare; democracy

The "attribution problem," the difficulty of finding out who did it, used to be one of the most discussed questions in the study of cyber conflict (Boebert, 2010; Clark & Landau, 2010; Dunn Cavelty, 2008, p. 20; Kello, 2013, p. 33; Lindsay, 2015; Lupovici, 2016; Morgan, 2010; Nye, 2017; Rid & Buchanan, 2015). Yet, recent empirical practice suggests that various actors not only have the capability to attribute, but have decided to share their findings in public. Public attributions claims, however, often remain contested.

So what is public attribution? Public attribution in its most elementary form is the blaming of a particular actor as responsible for a cyber incident.[1] It can be undertaken by a variety of actors, including governments, companies, NGOs, and academia. Governmental attribution is of particular importance, as government action to assign blame is an inherently political act.

CONTACT Florian J. Egloff florianegloff@ethz.ch

Governments thus have a strong incentive to use public attribution as a political tool, thereby making it interesting to the study of contemporary security politics. For research on governmental public attribution as an element of security policy, one can split the public attribution process into two phases: mechanisms that lead to public attribution and what happens after an incident is publicly attributed. Little research exists on either phase with regard to attribution of cyber incidents. This is problematic, as our understanding of contemporary security policy rests on understanding what drives threat narratives, how and why those particular ones are introduced publicly, and how contestation of threat narratives takes place in the public sphere.

My main contribution is to focus on the second phase of attribution, namely, what happens after a government goes public about a cyber incident.[2] Understanding this phase is important, as public attributions of cyber incidents are one of the main sources from which the public learns about who is attacking whom in cyberspace, thereby shaping the threat perception of the general public. Most attribution judgements are published by governments, the private sector, and a small number of civil society actors. To situate the knowledge space, in which attribution claims are introduced to, I reflect on this source of knowledge about cyber conflict by identifying how this knowledge structurally shapes our understanding of cyber conflict, in particular due to operational and (political, commercial, and legal) structural factors. In short, due to the commercial incentives on the private sector side and the political bias on the government side, the public data about cyber conflict structurally induces distrust into the representativeness of the public attribution statements.

A second contribution is to focus on the contestation of public attribution claims in democracies and the consequences such contestation brings. Contestation is fundamental to democratic politics. The open debate, the ability of everyone to freely voice opinions, and the emergence of truth trough democratic discourse is foundational to the public sphere of democratic polities. Thus, the ability to contest is a sign of healthy democratic politics. As this article will show, however, this openness to contestation, coupled with the information poor environment, creates particular problems in the area of cybersecurity. Two main questions are pursued: How do actors engaging in public attribution cope with the contestations that follow the introduction of their attribution claims? In what way could academia address those problems?

Investigating the contestation of public attribution claims and the consequences of such contestation enables us to better understand the politics of public attribution. The article argues that public attribution of cyber incidents takes place in a heavily contested information environment, creating fractured narratives of a shared past. Due to the secrecy attached surrounding the attribution processes by governments, particularly due to concerns of intelligence

agencies about sources and methods, governments are often reluctant to reveal the evidence underlying the attribution judgments. These are ideal enabling conditions for other actors to contest governmental claims. In a series of empirical examples, I reflect on the drivers of contestation after an incident is publicly attributed and show how attackers and other constituencies with various political and economic motivations drive particular narratives.

In a final part, I propose how academia could be a partial remedy to this situation. Academia, so far, has not been a strong participant in the discursive space around particular attributions. This is despite its commitment to transparency and independence theoretically making it a well-placed actor to contribute an independent interdisciplinary contribution on the state of cyber conflict. Thus, I argue for an increasing need for academic interventions in the area of attribution. This includes interdisciplinary research on all aspects of attribution (not just in cybersecurity), and conducting independent research on the state of cyber conflict historically and contemporarily. One of the main implications of this research on contestation of attribution claims for democracies are to be more transparent about how attribution is performed, to enable other civilian actors to study cyber conflict, and to thereby broaden the discourse on what is one of the main national security challenges of today.

The article is structured in three parts: First, public attribution is situated in the context of the literature on the contestation of claims in public opinion formation and the general biases underlying current public attribution claims are explained. Second, three empirical examples are used to reflect on the interaction of public attribution with a contested information environment, namely, the public attributions of the intrusions into Sony Pictures Entertainment in 2014, of the intrusions into the Democratic National Committee (DNC) in 2016, and of the NotPetya incident in 2017. In the third part, I assess the implications of contested attributions and suggest how this calls for a different role of academia, particularly with regard to research on attribution in universities. The conclusion offers a perspective on the biggest research opportunities and challenges in the study of cybersecurity and attribution in the future.

Public attribution fosters a skewed picture of cyber conflict

Contemporary security policy takes place within a heavily contested information environment. To better understand how security politics takes place with regard to cybersecurity, we first have to reflect on how our knowledge of cyber conflict is constructed. At an elementary level, we do not directly observe who is using cyber insecurity to further their interests against whom. Thus, our knowledge space of cyber conflict is constructed by

different information providers. In this first part, I first situate the contestation of public attribution claims within the literature on attribution of cyber incidents, before reflecting on how two operational and three structural factors shape the public data about cyber conflict in such a way so as to structurally induce distrust into its representativeness.

Situating public attribution in the literature on attribution of cyber incidents

The literature on cyber conflict has focused on documenting the shift from the attribution problem hindering successful policy responses, towards attribution being often possible for state agencies. For example, in 2015, Rid and Buchanan challenged three assumptions prevalent in parts of the literature: Namely, that attribution is intractable due to the nature of cyberspace, that the difficulty in attribution is finding the evidence, and that attribution is either solved or unsolved (Rid & Buchanan, 2015, p. 6).

More recently, scholars are focusing on public attribution, including whether and how to build an international organization aiding attribution processes (Eichensehr, 2019, 2020; Finnemore & Hollis, 2017, pp. 475–476, 2019; Grindal, Kuerbis, Badiei, & Mueller, 2018; Schulzke, 2018; Solomon, 2018). Reflections on the legal, and naming and shaming aspects of public attribution help to clarify the international normative function of public attribution (Eichensehr, 2020; Finnemore & Hollis, 2019), whilst analysis of institutional policy proposals shows potential ways towards improving transparency and credibility problems (Egloff & Wenger, 2019; Eichensehr, 2019, 2020; Grindal et al., 2018; Solomon, 2018). Empirically, in the last five years, public attribution of cyber incidents has moved from being incredibly rare, to becoming a more routine national security policy option in international politics. Despite the increasing prevalence of public attribution claims, little research—with the exception of Schulzke (2018, discussed below)—has zoomed in on the two phases of public attribution, namely, first, the mechanisms that lead to public attribution and second, what happens after an incident is publicly attributed. Splitting public attribution processes into these two phases is a useful conceptualization, as there are a whole series of political processes giving rise to a government considering public attribution, whilst a different set of political challenges inform the handling of the situation after having introduced the public attribution claim.

The lack of research is unfortunate, as both parts are relevant to understanding where attribution sits within the larger domain of the politics of cyber insecurity. First, better understanding the mechanisms leading to public attribution is important to situate the actor's own understandings of the domain and their activity in a larger context. As an example, for the international relations scholarship it matters whether public attributions are aimed

at domestic or international audiences. Furthermore, the incentives to go public will shape the type of attributions representing the public record of cyber conflict, whether they be court cases, public ministerial statements, technical reports, or joint diplomatic statements delivered with allies. Second, better understanding what happens after an incident is publicly attributed is important to assess the utility of public attribution as a policy tool, but also, to better understand the wider politics of cyber insecurity. For example, if public attributions are used by elites to frame our understanding of cyber conflict, the type of framing, and the effect thereof are worthy of further study. This article focuses mainly on this second part of the process, namely, what happens after an incident is publicly attributed.

Schulzke (2018) has suggested some theoretical reasons why we should pay careful attention to the effects of public attribution of cyber incidents. Drawing on psychological research, he argued that people do not like ambiguity when searching for explanations of unexpected threatening events (p. 957). In the attribution of cyber incidents, the relatively long time-frame between the event and a confident attribution statement results in people already having drawn their own conclusions, before "better" information may be available. Schulzke draws out four theoretical mechanisms particularly poignant for public opinion formation of attribution of cyber incidents. First, due to the information poor environment, causal narratives offered by elites gain large weight in media coverage (Baum & Groeling, 2010; Bennett, Lawrence, & Livingston, 2007; see also Stone, 1989). This effect is stronger in the cyber environment compared to kinetic incidents, as in many cases the victim has the ability to keep the occurrence of a cyber incident a secret. Second, existing hostilities influence subsequent threat perception, with attributional uncertainty having the potential to reinforce current strategic narratives. Schulzke argues that this, again, is particularly exacerbated compared to kinetic attacks, as evidence is harder to grasp and attribution, in general, is more ambiguous resulting in even stronger power of initial framings of blame (p. 959).

Third, information about cyber incidents is regularly not forthcoming, reinforcing the perception of hidden processes and making cyber attack causal narratives very similar to conspiracy theories (i.e., "a proposed explanation of some historical event (or events) in terms of the significant causal agency of a relatively small group of persons-the conspirators-acting in secret." Definition from Keeley, 1999; see also Schulzke, 2018, p. 962). Fourth, because of time delays in attribution processes it is hard to hold policymakers accountable. In addition, because cybersecurity as policy issue is distributed amongst different stakeholders, responsibility is diffused and susceptibility to partisanship increases. From these four mechanisms, Schulzke (2018) concludes,

> the public should be sensitized to the problem and to a security context in which blame takes time to establish. It would be prudent for policymakers and journalists to avoid a rush to judge who is responsible for an attack, regardless of how obvious the answer may initially seem. This would promote greater openness to information that is uncovered by investigators. Political scientists can play a role in this by exploring how these novel threats fit into what previous research has uncovered about opinion formation and conflict processes. Future research should continue to investigate the political challenges associated with attribution and to consider what additional steps could be taken to manage this problem. (p. 964)

Agreeing with the desirability of Schulzke's prescriptions, and following his call for more research in this area, this article introduces a set of arguments beyond the four mechanisms raised. Schulzke mostly draws on psychological, opinion formation, and media research, arguing about the attribution environment in abstract. This article contributes another element: contestation of public attribution claims and the consequences such contestation brings. To do so, the article first introduces factors skewing the public discourse on cyber conflict. This is an important reflection in order to contextualize our knowledge of cyber conflict. Second, the drivers of contestation after an incident is publicly attributed are discussed, with particular focus on how contestation is partially driven by the attacker and partially by other constituencies with various political and economic motivations. This has consequences for the opinion formation processes. Third, the article concludes drawing out the likely consequences for the overall discursive environment in public attribution, and recommends a changed role for academic interventions in this space. Only few places in the world have specialized programs to perform interdisciplinary research on cybersecurity and international relations. While outliers exist (see research of the Citizenlab; for a different example see Demchak & Shavitt, 2018), a sustained academic engagement with the process that produces the realities of cyber conflict and an active positioning within that is sorely needed.

A skewed (public) picture of cyber conflict

To better understand the consequences of governments and private actors publicly revealing the perpetrators behind cyber campaigns, we need to understand the biases underlying the data and judgements made. Those biases in the data and judgements will shape our knowledge space on cyber conflict. There are at least two operational, and three structural factors that jointly skew our picture of cyber conflict.

The first two factors are operational. First, offensive actors hide their tracks. The actors engaging in offensive behavior often have incentives to hide their tracks to achieve their goals. Thus, the attackers may sometimes be the only ones knowing they are engaging in offensive behavior. Of course, some

actors want to advertise the origin of cyber activity broadly, but this remains the minority of publicly known incidents (Poznansky & Perkoski, 2018). Second, the victims often have incentives to keep their victimhood secret, limiting the amount of information we have about the impact of cyber conflict. This is both due to operational and reputational reasons: Operationally, it can be advantageous to not disclose your knowledge to the attacker; reputationally, organisations may come to the conclusions that it is better not to disclose having been breached.

There are three structural factors that further skew our public knowledge of cyber conflict. First, the security companies, which are one of the main sources of how we learn about cyber conflict, have limited visibility, mostly defined by the technologies employed and the markets they serve. Furthermore, they have a specific prism of what kinds of threats they investigate further (on the political choices made by security companies, see Stevens, 2019). That prism results in highly detailed knowledge about a subset of actors engaging in offensive measures. Only a subset of the research on these investigated actors is then published and picked up in the public's awareness of cyber conflict. In addition, most of the companies investigating threats (threat intelligence companies) are based in Western states—though this is changing, for example, with China-based threat intelligence teams increasingly reporting on threat activity within China (one example is Qihoo360). Partially due to their client base, partially due to political sensibilities, and partially due to different political priorities resulting in a different target set, threat intelligence companies rarely publicly reveal Western operations (for an analysis of threat intelligence reporting focusing on civil society, see Maschmeyer, 2019).

Second, there exists an attribution asymmetry. The rights and responsibilities of governments and private actors are still in political dispute. This means, there still is a lacking baseline on who has to provide security for whom, and for what price, in what circumstance (see Dunn Cavelty & Egloff, 2019). The result is highly unequal investments with regard to whom attribution capabilities are used for: The financially potent have the means to buy attribution, whilst some of the political targets have the public visibility for security companies to show off their skills. Outside of those two categories, most organizations and citizens never get in contact with an attribution investigation. Third, recently, some governments, most prominently the members of the Five-Eyes signals intelligence alliance (consisting of the United States of America, United Kingdom, Canada, Australia, and New Zealand), have started to publicly attribute some cyber intrusions. Governments, however, have their own, politically derived incentives to publish some results of investigations and not others.

The political motivations behind public attributions, coupled with the scarce evidence offered for the attribution judgements, skews our knowledge

of cyber conflict. Because of the operational factors leading to underreporting, the commercial incentives on the private sector side, and the political bias on the government side, the public data about cyber conflict structurally induces distrust into the representativeness of the attribution statements made by these actors. However, due to the secrecy of the cyber incidents, the knowledge introduced in to the public domain by these actors is key in providing the discursive baseline for contestations about cyber conflict.

The remainder of this article focuses on this understudied element: The contestation phase that follows a public attribution. This phase is important, as by introducing attribution claims into the public discourse, actors are laying political blame for specific actions onto other actors. The next part shows how different actors are challenging these political acts in a contested information environment.

Contested public attributions in democracies

Contestation of the attribution offered by a government regularly follows attribution statements. This part focuses on this contestation phase, and the implications contestation has on the perception of cyber conflict, and the trustworthiness of attribution claims by society.

Contestation is fundamental to democratic politics (see, for example, Dryzek, 2002). The open debate, the ability of everyone to freely voice opinions, and the emergence of truth trough democratic discourse is foundational to the public sphere of democratic polities. Thus, the ability to contest is a sign of healthy democratic politics, or, in Dryzek's terms: "Contestation is democratic to the extent that it is engaged by a broad variety of competent actors under unconstrained conditions" (p. 77). Thus, ideally, there are a broad set of competent actors participating the contestation in the public sphere. As the series of empirical examples in this section will show, however, whilst the discourse on public attribution is generally open to contestation, the problems raised with the actors (see above), coupled with the information poor environment about cyber incidents to the broader public, creates particular opportunities for motivated, sometimes adversarial, intervention in the area of cybersecurity.

In abstract, there are different phases that contestation occurs within. At some point in time, an intrusion is discovered by the victim. The fact of the existence of the intrusion may, but does not have to, leak into the public domain. If it does, public contestation around who might be behind the intrusion starts at that point. If it does not, contestation starts the latest with the first public statement attributing the intrusion.

Contestation is undertaken by several different communities and is case-specific. As a generalized matter, contestation will depend on the type of evidence that the attributor offered. The government has some choice over

whether, and how much to reveal. Thereby, trade-offs have to be made on how much detail is revealed: revealing more detail is giving the attacker more insight into the governmental processes, particularly with regard to sources and methods. Concealing the incident entirely runs into the risk of not being in control of the narrative, should the fact of the existence of the incident leak. Revealing part of the incident, but not others, opens up contestation as to the credibility of the evidence. Overall, the less clear the matching between the evidence disclosed and the conclusions proffered, the more opportunity there is for antagonistic communities to introduce doubt unto the claims made.

This does not mean that all claims always have to be supported by public evidence: Indeed, we trust other people's claims based on trust in their person, their procedures, or their institutions all the time, often based on a common sense, that is, "our acceptance of the intractable facts about the world and our already existing shared experience and understanding about the social world" (Muirhead & Rosenblum, 2019, p. 127). However, with regard to the attribution of cyber incidents, there is not yet a wealth of shared experience and understanding to draw upon, let alone intractable facts, opening up the space to fundamental contestation.

In this contestation process, the attacker is able to influence the contestation. Hence, the attacker can offer alternative interpretations vying to persuade the same audiences the victim wants to convince. As such, Guitton (2014) was right when he cast public attribution as a "game to convince an audience." We now have evidence that different actors actively engage in this type of contestation, and that some of the actors are becoming more skilled at it. This can be best observed using empirical examples. For this reason, to illuminate the contestation phase of public attribution, the article observes some short empirical examples of contested attributions.

The universe of cases of public attributions in democracies is growing, with my current research project on the politics of public attribution already having identified around 50 cases across Western Europe and North America (Center for Security Studies, n.d.). The empirical examples included in this article are selected, because they are particularly instructive to observe the contestation phase, and to see the development of government policy to address some of the challenges of engaging in public attribution.[3] The Sony attack in 2014 is one of the earliest examples, where a government publicly blamed another government for a cyber incident, and tried to convince its audience by offering up evidence. It is particularly instructive, as it shows the contestation that can occur domestically, when a government claims attribution. The DNC case of 2016 is one of the (rare) examples, where we have corroborated public documentation of the adversary directly engaging in countermessaging, trying to muddy the attribution claims offered by the private sector and government. It adds empirical illustration to the

phenomenon of contestation by demonstrating the active nature of some adversaries in interfering in (domestic) political discourses. Finally, the NotPetya case of 2017/8 is one of the newer cases where a new policy of diplomatic collaboration to address the problem of not being able to share all the evidence is observable. Governments, recognizing their trust deficit in claiming attribution in public, have teamed up with other governments to lend more political weight and credibility to their claims: The NotPetya case illustrates this well.

Cyber intrusion at Sony Pictures Entertainment 2014

On November 24 2014, a wiper malware was activated on a large part of SPE's infrastructure, crippling the company's ability to continue their work. The malware issued a warning that company documents would be released if demands were not met by 11:00 at night. Having let that deadline pass, the group Guardians of Peace (GOP) published several movies, SPE internal documents, and e-mail archives of SPE executives over the next month, thereby alerting the public of the intrusion at Sony Pictures Entertainment.

Given the upcoming release date of the movie *The Interview*, speculations about possible North Korean connections were raised. North Korea reacted by issuing a press statement denying responsibility, but praising the attacking group for their actions and condemning SPE for producing a film "abetting a terrorist act" (KCNAWatch, 2014). This was followed by a statement by GOP on December 8 2014, which directly connected the showing of a movie to their actions. It demanded that SPE should "stop immediately showing the movie of terrorism which can break the regional peace and cause the War!" (Gallagher, 2014).

Lacking any indications of canceling the movie, it was a threat of terrorist attacks against moviegoers issued on the Pastebin platform on December 16 2014 that changed the dynamic. Despite the U.S. Department of Homeland Security's claim of having no intelligence about a plot against movie theaters, many movie theaters opted-out of showing the movie (Perera, 2014; Seal, 2015; United States Department of Homeland Security & Federal Bureau of Investigations, 2014). The next day SPE issued a press statement canceling the release, which had been scheduled for Christmas Day. By that time, the public discourse in the United States was still focused on whether it could really be North Korea behind the intrusion. For example, on the December 17, Kim Zetter, a renowned journalist covering information security, in an article weighing up different theories concluded: "Regardless of whether the Sony, Saudi Aramco and South Korea attacks are related, the evidence indicating they're nation-state attacks is circumstantial. And all of the same evidence could easily point to hacktivists. Our money is on the latter" (Zetter, 2014).

On December 18 2014, the White House announced it was considering a proportional response (White House, 2014). On December 19 2014, the FBI officially attributed the GOP's actions to the North Korean government (FBI, 2014). The same day, U.S. President Barack Obama confirmed the attribution to North Korea (Obama, 2014). The FBI did not offer specific evidence, but explained that their judgement was based in part in overlap of previously used malware, infrastructure, and tools used in previous attacks carried out by North Korea. Some experts arrived at the same conclusion. For example, Brian Krebs (2014) laid out in detail the evidence why it is plausible for North Korea to be behind the intrusion. Similarly, Nicholas Weaver (2014) explained why the U.S. government is credible in claiming North Korea provenance. Bruce Bennet (2014) of RAND explained how this fits within the larger North Korean political and security context. And finally, Dmitry Alperovitch (2014) of Crowdstrike asserted that his company's own assessment supports the attribution to North Korea.

However, this did not assuage the skeptics. The generic language used by the FBI led to skeptical statements from some information security specialists. For example, Robert Graham (2014) entitled his blog post "The FBI's North Korea evidence is nonsense," whilst Marc Rogers (2014a, 2014b), DEFCONs head of security, published two contributions explaining his doubt of the FBI's attribution claims.

Bruce Schneier (2014) entitled his December 24 blog post "Did North Korea really hack Sony?" and voiced his deep skepticism of the North Korea attribution. By the December 26 2014, NPR featured a segment on "Doubts Persist On U.S. Claims Of North Korean Role In Sony Hack" (Shahani, 2014). Some people also used the skepticism of the Sony attribution to feature their own companies. For example, Kurt Stammberger of Norse Inc. and Jeffrey Carr of Taia Global both claimed that their own company's analyses led them to different judgements (Biddle, 2014; Taia Global, 2014). Thus, the original FBI/Obama attribution was not convincing to parts of the security community, which was used to see arguments to be supported with data, and which was inherently skeptical of the use of government authority to give weight to a truth claim. At the same time, the Russian government expressed solidarity with North Korea on the film being "aggressive and scandalous," and charged the United States with escalating tensions without presenting direct evidence linking the intrusions to North Korea (Lukashevich, 2014).

There was still much speculation about the provenance of the Sony attack. So much so, that the then director of the FBI, James Comey, released more details on the Sony attribution on January 7 2015 (Comey, 2015). In a speech at Fordham University, he informed the public that one element they based their attribution on was an operational security mistake, where the North Koreans forgot to mask their true IP addresses. Immediately, this was

challenged by skeptics. Kim Zetter (2015) at Wired again covered the critics, this time leaving the last word to Richard Bejtlich, who offered some context:

> I don't expect anything the FBI says will persuade Sony truthers. The issue has more to do with truthers' lack of trust in government, law enforcement, and the intelligence community. Whatever the FBI says, the truthers will create alternative hypotheses that try to challenge the "official story." Resistance to authority is embedded in the culture of much of the "hacker community," and reaction to the government's stance on Sony attribution is just the latest example.

Bejtlich was right. Some parts of the security community continued to voice their skepticism. Particularly, Jeffrey Carr's Taia Global followed up with a report alleging Russian hackers still being on Sony's networks, and that this claim sheds doubt on the veracity of the U.S. government's attribution (Taia Global, 2015; see also Pearson, 2015). Carr's source offered as bona fides some documents that the hacker claims to originate from the Sony network, which Carr tried to verify. Problematically, Carr's source for this claim was a well-known Russian hacker with self-claimed previous contracting relationships to the Russian state (Best, 2019; Pearson, 2015). Of note: both Jeffrey Carr and Kurt Stammberger have since left the cybersecurity industry (Collier, 2018).

This empirical example shows some of the dynamics that occur in the contestation of a governmental attribution claim. While the government insists on its legitimacy to claim attribution without revealing sources and methods, other audiences use this predicament for their own argumentative strategies. Firstly, the perpetrator can contest the attribution claim by denying sponsorship and demand proof for the attribution claim. Secondly, skepticism of government authority will fuel distrust in a governmental claim without the evidence being offered.[4] Thirdly, interested third-parties can use this situation to influence the contestation, as had happened in the Sony case with the Russian government aligning itself with the North Korean narrative. In such a contested information environment, multiple "truths" will continue to co-exist. For example, Seth Rogen, the director of the film *The Interview*, claimed in 2018 that he still does not believe North Korea to be behind the hack (Marchese, 2018). The U.S. government, meanwhile, reinforced its attribution claim multiple times to gain legitimacy, eventually resulting in a detail-rich criminal complaint of a North Korean citizen (*United States of America v. Park Jin Hyok*, 2018).

Cyber intrusions into the Democratic National Committee 2016

The intrusions into the Democratic National Committee (DNC) before the U.S. presidential election in 2016 are a second example for how contestation of the attribution takes place in a politicized information environment.

In April 2016, Crowdstrike was called to help with incident response for the DNC. They discovered at least two threat actors were active on the DNC networks, one had been there since at least summer 2015, the other since April 2016. On June 14 2016, the DNC, together with the cybersecurity firm Crowdstrike, publicly attributed the intrusions into its networks to two separate Russian espionage groups, APT28 and APT 29 (or FancyBear and CozyBear in Crowdstrike's terminology). Immediately, the media enquired whether Crowdstrike was right. Various other security companies confirmed their findings, including Fidelis, ThreatConnect (2016), SecureWorks (2016), and Mandiant (Kopan, 2016a; Nakashima, 2016).

Just as immediate was the attacker's reaction, which was later confirmed to be the Russian Main Intelligence Directorate (also known under its former name GRU). On June 15 2016, they created a fake hacker persona named "Guccifer 2.0" and attempted sowing confusion and doubt over the attribution claim (*United States of America v. Netyksho et al.*, 2018). They claimed to have hacked the DNC and as bona fides offered documents, which the hacker claimed to originate from the DNC network. Over the next few months, the attackers used the Guccifer 2.0 persona to distribute documents on the Democratic Party's campaign and to shed doubt about the attribution to Russia. RT (2016), a Russian government funded media network, immediately picked up on the content and the contradictory claim of hacking provenance, and further distributed this interpretation.

However, by late July 2016, the cybersecurity community had identified parts of the broader Russian subversion campaign, including some of the influence elements (Gioe, 2018; Rid, 2016a, 2016b, 2017). Information security specialists and some journalists were quick to point out that Guccifer 2.0 was likely a front of the Russian intelligence services. Meanwhile, the U.S. government officially did not attribute cyber intrusion at the DNC.[5] Despite the broad sourcing and unusual clarity of evidence, it took another month for members of the legislative to go public (September 22 2016), and two months for the U.S. government to release a meagre public statement by the executive (October 7 2016) (Feinstein & Schiff, 2016; United States Department of Homeland Security & Director of National Intelligence, 2016).

Despite this, Republican presidential nominee Donald J. Trump further fueled the uncertainty about the origin of the leaked material by during the first presidential debate on September 26 2016, implicating other theories of provenance such as Russia, China, or a 400 lbs hacker (Kopan, 2016b). This is important, as the presidential debates are moments of highest political media exposure, and do lead to follow-up media stories, giving the multiple theories of provenance angle a bigger discursive platform. He continued to reiterate this position even after the election. On the December 11 2016 he implied on FoxNews Sunday interview with Chris Wallace that the

intelligence community has "no idea if it's Russia, or China, or somebody. It could be somebody sitting in a bed someplace" (FoxNews, 2016).

Only by the beginning of 2017, the government released an intelligence community assessment attributing various interference activities, including Guccifer 2.0, to the Russian government (National Intelligence Council (Office of the Director of National Intelligence), 2017). On January 12 2017, the GRU again used the Guccifer 2.0 persona to dispute the relationship to Russia. By alleging the falsification of evidence, it further fueled speculation about other theories of provenance, which continued to be covered by publications such as *The Nation* (Lawrence, 2017). In the end, the United States released a detailed indictment assigning blame to the Russian military intelligence service (*United States of America v. Netyksho et al.*, 2018).

Much of the delay in the U.S. public response is explained by domestic political concerns of the executive being seen to unduly intervene in the election process. At the domestic level, the procedures were underdeveloped and unprepared for a coordinated response between the federal government and the state-based election officials. Inter-agency discussions and indecisiveness about what to do next significantly slowed down this process. For example, in August 2016, then FBI Director James Comey drafted an op-ed publicly attributing the Russian activities that was debated in the inter-agency process but was never published (U.S. Office of the Inspector General, 2018).

The contestation of the provenance of the intrusions was enabled by this lack of a public response. As a consequence, the use of the materials, also by the traditional media, was less problematized, as there was no governmental voice claiming illegitimate interference into the election. This is consistent with research on the influence of elites on initial framings, in its absence opening up space in the initial coverage of the incident to the attacker's narrative (see for example Baum & Groeling, 2010). This effect will diminish over time, as the official narrative settles (Bennett, Lawrence, & Livingston, 2006). However, because of the uncertainties in this initial contestation, doubts on are going to persist, and significant work on behalf of the government has to be undertaken to gain credibility for its version of events. In the case of the election interference in the United States in 2016, this included the appointment of a special counsel, who brought detailed indictments against Russian activity as well as published two reports on the outcome of the investigation (Mueller, 2019).

NotPetya incident in 2017

Finally, NotPetya, a third empirical example, can demonstrate how the meaning-making activities of governments have changed, partially, in order to address the problem of not being willing to release clear and convincing evidence in a timely manner in the public domain.

NotPetya was a severe destructive cyber campaign aimed primarily at Ukraine on the June 27 2019. It employed a wiper worm, which caused destruction worldwide. As with any large cyber incident, the question quickly arose, who is behind this campaign. Pretending to be ransomware, the worm itself displayed none of the elements a cyber criminal campaign would entail. The Ukrainian security service soon claimed for the Russian special services to be behind the attack (SBU, 2017). This claim was reinforced by cybersecurity companies, who independently found technical links between the NotPetya attack and the BlackEnergy group, a group who was also associated to the previous attacks on the Ukrainian energy grid (Cherepanov, 2017; Kaspersky Blog, 2017). Russia immediately denied these allegations (Brewster, 2017). Thus, the matter rested in a stand-off. This is until the Five-Eyes, the signals intelligence cooperation between the United States, United Kingdom, Canada, Australia, and New Zealand, decided to publicly attribute the campaign to the Russian military.[6] Again, Russia swiftly denied the allegations (Russian Embassy in the USA, 2018). However, the political momentum behind this attribution differed from the Ukrainian response. The British attorney general, Jeremy Wright (2018), outlined the reasoning:

> If more states become involved in the work of attribution then we can be more certain of the assessment. We will continue to work closely with allies to deter, mitigate and attribute malicious cyber activity. It is important that our adversaries know their actions will be held up for scrutiny as an additional incentive to become more responsible members of the international community.

This joint approach of the Five-Eyes was corroborated by the U.S. State Department, which highlighted the importance of partners in buttressing each other's attribution claims and responses (Office of the Coordinator for Cyber Issues, 2018).[7] Thus, the Five-Eyes governments had decided that they use public attribution as a way of shaping the environment. By building an international coalition, attributing blame publicly to an actor for specific actions that are deemed undesirable to the international community, the coalition of states is changing the operational environment.

The intent, as outlined by the British government official, is to establish boundaries for responsible behavior, respectively clearly labeling the behavior deemed "irresponsible." Indeed, the United Kingdom's hopes to attain benefits with pursuing public attribution more broadly, including making cyberspace "more transparent as counter-normative and destructive behaviour (i.e. Wannacry and NotPetya) are attributed" leading to "greater stability in cyberspace as clear lines of unacceptable behaviour are drawn," increasing the legitimacy of attribution by undertaking attribution with allies, and using attribution as a first step "to enable wider response options to impose costs on the responsible actors" (UK government, n.d.). Thus, the British government

identifies its activities in public attribution as boundary drawing behavior that gains more legitimacy as it is undertaken in coalitions of states. Though an in-depth international legal analysis is out of scope, one can also note, that by acting in coalition states prepare the ground for establishing state practice, one of the sources for international customary law (see further Finnemore & Hollis, 2019).[8] This is in congruence with the literature on international norms: In order for norms to be effective, norm violations have to be acknowledged (Finnemore & Hollis, 2017; Miller, Nakashima, & Entous, 2017). The norm shaping activity seems to be targeted at encouraging state actors to not indiscriminately deliver effects.

What the years of experience with public attribution since Rid and Buchanan's (2015) seminal article have shown that, contrary to their suggestions, specificity in the data presented to support a public attribution claim is not a necessity to advance such a claim successfully. In the examples raised in this article (Sony, DNC hack, NotPetya), the government did not readily offer up specific data to back up their attribution claims. The Sony hack thereby is particularly instructive: The pointing to declassified evidence actually weakened the government's truth claim, as it enabled other commentators to find alternative explanations for the particular data. In the DNC and NotPetya case, the governments initially were able to build on the private sector claims, and used the government's reputation of possessing capable (signals) intelligence capabilities to convince some audiences of the trustworthiness of their claims.

Implications of contested public attributions for the role of academia

Whilst the first section of this article reflected on the operational and structural factors leading to a skewed picture of cyber conflict, the previous section identified in three empirical examples the contestation around the provenance of cyber intrusions. This last section reflects on where this leaves the knowledge space on public attribution, and proposes how academia, if it became a competent stakeholder in the contestation of the public attribution discourse, could partially remedy some of the problems identified.

The three empirical examples showed that public attribution takes place in a contested information environment, with the attacker also having a voice. Attribution claims are introduced mainly by two sets of stakeholders: governments and cybersecurity companies. Whilst cybersecurity companies provide technical data, forensic links, and estimates about who might be behind a cyber intrusion, they are often not the political entity laying blame on a specific attacker. In the Sony case, Mandiant provided the incident response services to Sony, but refrained from publicly attributing the incident. The DNC intrusions are the exception, where Crowdstrike was asked by the DNC to go public with specific attribution claims. In the NotPetya case,

private industry provided much of the remedy to the infection, and offered analysis connecting the operation to previous campaigns, but refrained from making political attributions. Rather, it was governments that, in a coordinated manner, laid blame onto the Russian military.

There are long-term ramifications stemming from attributions in such contested information environments. First, the uncertainty about the truth claims persist. Thus, even several years after a specific cyber incident, there are people asking about whether we can really know who is behind the incident. This is due to the original event being clouded in this contested information environment, where the multiplicity of narratives offered to the public create an impression that one really cannot know who is behind a cyber incident. The secrecy of the attribution processes thereby lays fertile ground for conspiratorial thinking (Schulzke, 2018). This is beneficial to perpetrators of cyber intrusions, who, despite implausible deniability, can continue operating with relative impunity (Cormac & Aldrich, 2018).

Indeed, whilst not unique to the area of cyber incidents, the destabilization of the perception of the achievability of attribution is particularly impactful in an information poor environment, where outside verification is difficult (For an example of the destabilizing an official narrative laying blame onto a specific actor in a physical incident, see Ramsay and Robertshaw's (2018) analysis of the narratives proffered in the wake of the poisoning of Sergei and Yulia Skripal). For democracies, thereby, secrecy surrounding the attribution processes poses a particular challenge in this contested information environment: Legitimacy of state action is dependent on it being explainable, accountable, and ultimately transparent (even if in hindsight). For this democratic accountability, the public should be able to get an accurate understanding of the underlying conflict dynamics that the use of public attribution is part of (see also Nincic, 2003; Nothaft, Pamment, Agardh-Twetman, & Fjällhed, 2018).

Second, the actors providing the attribution judgements are motivated by political and financial incentives. The strongest cybersecurity policy actors have, however, been (signals) intelligence agencies, who by design are not very public entities. So far, only the United States government has released how they approach attribution processes of cyber incidents (Office of the Director of National Intelligence, 2018). Traditionally, intelligence agencies have been the locus with the most resources and abilities to pursue attribution processes at the state level. The recent political use of their analytic work in public is democratically problematic, as the underlying processes for reaching their conclusions are kept secret, usually for at least thirty years. The governments choose, based on political considerations, which incidents to attribute publicly. Oversight bodies remedy some of the trust deficit, though their remit is much wider than overseeing attribution processes. Yet, state intelligence agencies remain one of the main sources of information about cyber conflict.

Thus, the electorate is dependent on other sources of information that can triangulate governmental information. Whilst governments may have good reason for not disclosing their sources and methods, private sector structurally does not provide a remedy, as it has its own structurally induced incentives to be only partially transparent. Particularly, claims made based on data gathered in customer engagements often legally have to be cloaked in secrecy. As previously noted, structurally, we end up with a lack of transparency both about the baseline of the actual events as well as about how actors have reached their judgements about the events. Similarly problematic are the trust issues associated with cybersecurity companies. The politicization of the companies themselves, as well as the selection of knowledge that they publicly produce, make the establishment of trust in a country as well as across countries challenging. However, shared discursive spaces for trust in knowledge about cyber conflicts across societies are much needed, particularly as future conflicts are likely to rely on digital interaction to an even higher degree than today.

Finally, academia, so far, has not been a remedy, as researchers use the public record as the baseline for studying cyber conflict (see Valeriano & Maness, 2014). In order for academia to be a check on the biases introduced, it would have to establish itself as a trustworthy source of data and interpretation. For that to happen, interdisciplinary knowledge on attribution processes is required. For example, the ability to judge a company report's attribution claims depends as much on the ability to understand the data, the judgements made, and the hypotheses analyzed, as it does to understand the specific company's access to the data it may not make public, but rely upon, to make its analytic judgements. Thus, there is a need for additional trustworthy sources of data and interpretations of data. In addition, there exists a call for a less elitist, more democratic access to attribution knowledge (Schulzke, 2018). One answer could be a network of distributed and regionally trusted public knowledge-creators (as, for example, proposed by Deibert in Solomon, 2018). Universities may provide one such possible locus to remedy the trust problem. Their strong rootedness in academic, transparent, and peer-reviewed research affords them societal trust, which other actors, due to their structurally induced incentives do not, and possibly cannot, have. In addition, universities are well positioned to be independent stewards of data and analytic methods. They can transparently analyze, provide context, and integrate new phenomena of digital conflicts into shared knowledge structures.

However, to be such sources of trustworthy information on attribution, universities have to engage, even more than today, in work across both physical, disciplinary, and academic boundaries. They have to transcend physical boundaries, as part of the trust problem is associated with national political interests. A cross regional collaboration therefore could ensure that, should

national political priorities taint a specific attribution process, other universities are able to balance that process. They have to transcend disciplinary boundaries, as attribution processes inherently draw on multi-disciplinary knowledge. Finally, they have to transcend academic boundaries, as much of the data resides with governments, companies, and civil society. Thus, academics have to be willing to collaborate across multiple entities to for their engagement in attribution research. The outcome is in line with what Ronald J. Deibert, the head of the CitizenLab at University of Toronto called for, namely, "to empower civilian institutions in multiple countries with resources and capabilities to do independent research on threats to cyberspace in the public interest regardless of boundaries, and regardless of whose national or commercial interests are concerned" (Deibert, 2017). Should such a network of independent institutions with capabilities in the attribution space exist, they could become a stabilzing element in the contested information environment (see also Eichensehr, 2019, 2020). Whilst universities are often accused to be risk-averse institutions, in the context of research on attribution, this would be an asset.

Conclusion

Attribution research has come a long way: From accepting the "technical" impossibility toward arguing about the constructed nature of the attribution problem, we have seen the full spectrum of positions. Meanwhile, both the practice of the actors studied, as well as their capabilities and attitudes towards attribution, have changed. This has left research transformed. We are currently witnessing an international political contestation of what attribution is, can be, and should be. Despite this, little research examined the two phases of public attribution, namely, the mechanisms that lead to public attribution and what happens after an incident is publicly attributed. This article made a contribution to address this research gap.

At a mundane, everyday, level of analysis, the actors shining the light, the one's possessing the capability to perform (or to order on their behalf) attribution investigations, are fundamentally shaping our view of cyber conflict. This has far-reaching implications on our research practices, as our scholarly analyses are, through the (re-)use of this data, reconstituting a reality reflected through political and economic prisms of these actors.

Attribution claims are introduced, contested, and even the possibility to do attribution is put into question. Disinformation tactics are used to muddy specific attribution claims, leaving an electorate exposed to the coexistence of "multiple truths" and a fractured narrative of the past. This article has detailed some of the implications introducing attribution claims into contested information environments have. Through a study of the contestation phase of public attribution, it showed a multitude of voices contributing to

the adoption of a particular truth claim and, in particular, the influence attackers and motivated third parties can have in the discourse. The Sony incident highlighted the difficulty a government can have in convincing an electorate of its claims, when there is no record of accomplishment in making attribution claims in public. The DNC intrusion showed how the attacker can take part in the meaning-making activities, actively trying to dispel the notion that the government knows who is behind a cyber incident. Finally, the NotPetya incident showed how actors seemed to have learned from the contested cases. In particular, the coordination of attribution claims across different countries and entities was specifically designed to bolster the legitimacy and credibility of the attribution claims at the international level.

Further research should investigate the parameters shaping the different choices of the data presented and credibility garnered. For example, two rationales could explain the initial lack of detail offered. First, from an intelligence perspective, the sources and methods informing the judgment may be particularly time sensitive. Second, the lack of detail may be motivated by the fact that technical evidence never "proves" responsibility and is usually open to interpretation. By not offering any data, but using strong estimative language such as "almost certain" or "highly likely," attributing entities wager their reputations as competent entities making attribution judgements. Other governments, for example the Swiss, chose yet an opposite route, publicizing detailed technical reports, but not publicly blaming a particular actor for an intrusion (GovCERT.ch, 2016).

This leads to a particular role for academia and a need for more research on attribution. First, academia can make independent assessments of attribution claims. An example: In July of 2016, enough information was available in the public domain to document and preliminarily assess the Russian influence campaign into the American elections. Few voices from U.S. academia were audible at the time, with Thomas Rid (at the time still in the United Kingdom), being one of the only academics willing to speak out (Rid, 2016a). One would hope for more academics to engage in the public discourse, as they have the potential to represent a more trustworthy source of information. This requires a deeper interdisciplinary engagement, trusted contacts in the information security community, and an access to data repositories to pursue such an investigation. Engaging in such public contestation comes with personal risks to individuals that could be mitigated partially, if more institutions were capable and willing to contribute to the public discussion of attribution claims (on personal risks, see for example Satter, 2019).

Second, academia can explain the conditions under which a truth claim is believed. Particularly, the "technical" core of the problem is still considered settled knowledge—rather than being contested. Research in this area includes an explanation of how evidentiary standards emerge, stabilize, and destabilize

over time, and how trust in the speaker, the process, or the institution can be a precondition for introducing successful attribution claims.

Finally, the politics of attribution need to be further unraveled. For example, attribution asymmetries and their impact need to be better explored. Research into whom attribution capabilities are used for, and whom not, is crucial to understand the power dimension of attribution. Attribution plays a key role, as it can rebalance asymmetric power relationships transforming "the attacks from nowhere" into "attacks by a specific entity." Identifying the perpetrator is empowering, opening-up specific strategies of resistance for victims. The study of the politics of publicly identifying a perpetrator, a power move, can tell us more about the functioning of cyber and security politics at the national and international level.

There are promising prospects for this. More data is available from underreported parts of the world than ever. Threat intelligence companies based in non-Western countries are increasingly contributing to a richer, more diverse, analytic picture of the threat landscape. Furthermore, universities have a unique opportunity to contribute to the broadening of the discourse, for example, by choosing independent focuses of analysis, combining the research tradecraft of international relations with those of applied computer science. The biggest challenge in this will be the universities themselves: It is overcoming the disciplinary confines in order to shine a light onto the emerging digital futures that for some, despite all the conveniences they bring, also bring the horrors of more domination and control.

Notes

1. A more precise definition would split the attribution process into sense-making and meaning-making processes (Egloff, 2018, p.148, 165). Shortly defined, the sense-making process in attribution refers to the ongoing knowledge-generation process that establishes what happened, whereas the meaning-making process refers to deliberate actions that influence how others interpret a particular cyber intrusion. Public attribution is a specific kind of meaning-making process, which can then be split analytically into the two phases introduced in this article.
2. For research on the former phase, see Center for Security Studies (n.d.).
3. Part of these empirical examples draw on material first introduced in Egloff, 2018.
4. A claim that is also supported by research on value similarity and trust, see Visschers and Siegrist (2008).
5. At working level, the FBI had the DNC intrusions on the radar since 2015, but there is no public evidence of it briefing the White House before 2016.
6. New Zealand did not independently assess it, but joined the Five Eyes in the condemnation, whilst Canada attributed NotPetya to actors in Russia.
7. For the U.S. legal view of attribution, see Egan (2017).
8. Thanks to Dr. Matteo Bonfanti for pointing this out.

Acknowledgments

The article builds on a conference paper presented at the CSS Cyber Conference in September 2018, as well as on my previous work and ongoing research project on the Politics of Public Attribution. I thank all the conference participants, and in particular Myriam Dunn Cavelty, Jasper Frei, and Miguel Alberto Gomez, for their valuable feedback. Thanks also to Jasper Frei for his help with referencing and formatting of the paper.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Florian J. Egloff is a Senior Researcher in Cybersecurity with the Center for Security Studies at the ETH Zürich, Switzerland. His research focuses on the politics of cybersecurity, particularly with regard to intelligence policy, and the role of non- and semi-state actors in cybersecurity. Florian's current projects focus on the politics of public attribution and the use of cyber intrusions for political purposes. Prior to working at ETH Zürich, Florian wrote his DPhil (Ph.D.) in Cyber Security at the University of Oxford on *Cybersecurity and Non-State Actors: a Historical Analogy to Mercantile Companies, Privateers, and Pirates*. Florian is also a Research Associate at the Centre for Technology and Global Affairs at the Department of Politics and International Relations and teaches at the Centre for Doctoral Training in Cyber Security (both at the University of Oxford).

ORCID

Florian J. Egloff https://orcid.org/0000-0002-0290-667X

References list

Alperovitch, D. (2014, December 19). How should the U.S. government respond to North Korea's attack on Sony? *PBS News Hour Web page*. Retrieved from https://perma.cc/VD29-FA4K

Baum, M. A., & Groeling, T. (2010). Reality asserts itself: Public opinion on Iraq and the elasticity of reality. *International Organization*, *64*, 443–479. doi:10.1017/S0020818310000172

Bennet, B. W. (2014, December 11). Did North Korea Hack Sony? *The RAND Blog*. Retrieved from https://perma.cc/8GAW-7246

Bennett, W. L., Lawrence, R. G., & Livingston, S. (2006). None dare call it torture: Indexing and the limits of press independence in the Abu Ghraib Scandal. *Journal of Communication*, *56*, 467–485. doi:10.1111/j.1460-2466.2006.00296.x

Bennett, W. L., Lawrence, R. G., & Livingston, S. (2007). *When the press fails: Political power and the news media from Iraq to Katrina*. Chicago, IL: University of Chicago Press.

Best, E. (2019, March 20). Leaker, Liar, Hacker, Hoaxer: The Russian contractor who infiltrated anonymous. *Emma Best Blog*. Retrieved from https://perma.cc/NM3P-NQNM

Biddle, S. (2014, December 31). Researcher: Sony hack was likely an inside job by a woman named "Lena". *Gawker Web page*. Retrieved from https://perma.cc/FME5-QN2X

Boebert, W. E. (2010). A survey of challenges in attribution. In N. R. Council (Ed.), *Proceedings of a workshop on deterring cyberattacks: Informing strategies and developing options for U.S. Policy* (pp. 41–52). Washington, DC: The National Academies Press.

Brewster, T. (2017, July 3). NotPetya Ransomware Hackers 'Took Down Ukraine Power Grid'. *Forbes Web page*. Retrieved from https://perma.cc/T9ED-4MRF

Center for Security Studies. (n.d.). The Politics of Public Attribution. *Center for Security Studies Webpage*. Retrieved from https://perma.cc/2CU3-T3YH

Cherepanov, A. (2017, June 30). TeleBots are back: Supply-chain attacks against Ukraine. *Welivesecurity Web page*. Retrieved from https://perma.cc/97MG-N9PR

Clark, D. D., & Landau, S. (2010). Untangling attribution. In N. R. Council (Ed.), *Proceedings of a workshop on deterring cyberattacks: Informing strategies and developing options for U.S. Policy* (pp. 25–40). Washington, DC: The National Academies Press.

Collier, K. (2018, September 7). The indictment of North Korea for the Sony hack shows how cybersecurity has evolved. *BuzzFeed News Web page*. Retrieved from https://perma.cc/64PB-46R4

Comey, J. B. (2015, January). *Addressing the Cyber Security Threat*. Remarks delivered at the International Conference on Cyber Security at Fordham University, New York City, NY.

Cormac, R., & Aldrich, R. J. (2018). Grey is the new black: Covert action and implausible deniability. *International Affairs*, *94*, 477–494. doi:10.1093/ia/iiy067

Deibert, R. J. (2017, January 4). The DHS/FBI Report on Russian hacking was a predictable failure. *JustSecurity*. Retrieved from https://perma.cc/XSL2-4FFZ

Demchak, C. C., & Shavitt, Y. (2018). China's maxim – leave no access point unexploited: The hidden story of China Telecom's BGP Hijacking. *Military Cyber Affairs*, *3*(1), 1–9. doi:10.5038/2378-0789.3.1.1050

Dryzek, J. (2002). *Deliberative democracy and beyond: Liberals, critics, contestations*. Oxford: Oxford University Press.

Dunn Cavelty, M. (2008). *Cyber-security and threat politics: U.S. Efforts to secure the information Age*. London: Routledge.

Dunn Cavelty, M., & Egloff, F. J. (2019). The politics of cybersecurity: Balancing different roles of the state. *St Antony's International Review*, *15*(1), 37–57.

Egan, B. J. (2017). International Law and stability in cyberspace. *Berkeley Journal of International Law*, *35*, 169–180. doi:10.15779/Z38CC0TT2C

Egloff, F. J. (2017). Cybersecurity and the age of privateering. In G. Perkovich, & A. Levite (Eds.), *Understanding cyberconflict: Fourteen analogies* (pp. 231–247). Washington, DC: Georgetown University Press.

Egloff, F. J. (2018). Cybersecurity and non-state actors: A historical analogy with mercantile companies, privateers, and pirates (Doctoral dissertation). University of Oxford, Oxford.

Egloff, F. J., & Wenger, A. (2019). Public attribution of cyber incidents. *CSS Analyses in Security Policy*, *244*, 1–4. doi:10.3929/ethz-b-000340841

Eichensehr, K. E. (2019). Decentralized Cyberattack attribution. *AJIL Unbound*, *113*, 213–217. doi:10.1017/aju.2019.33
Eichensehr, K. E. (forthcoming 2020). The Law & politics of Cyberattack attribution. *UCLA Law Review*, *67*.
FBI. (2014, December 19). Update on Sony investigation. *Press Release Web page FBI*. Retrieved from https://perma.cc/Z745-YR3S
Feinstein, D., & Schiff, A. (2016, September 22). Feinstein, Schiff statement on Russian hacking. *Senator Dianne Feinstein's Web page*. Retrieved from https://perma.cc/DF4G-44EN
Finnemore, M., & Hollis, D. B. (2017). Constructing norms for global cybersecurity. *American Journal of International Law*, *110*, 425–479. doi:10.1017/S0002930000016894
Finnemore, M., & Hollis, D. B. (2019). Beyond naming and shaming: Accusations and international Law in cybersecurity. *Temple University Legal Studies Research Paper*, *14*, 1–31.
FoxNews Sunday. (2016, December 11). Chris Wallace Hosts FoxNews Sunday. Interview with President-Elect Donald Trump. *YouTube Web page*. Retrieved from https://www.youtube.com/watch?v=MKWt6D0V6yk
Gallagher, S. (2014, December 8). Sony pictures attackers demand: 'Stop the Terrorist Film!'. *ars Technica Web page*. Retrieved from https://perma.cc/T2BL-QVRD
Gioe, D. V. (2018). Cyber operations and useful fools: The approach of Russian hybrid intelligence. *Intelligence and National Security*, *33*(7), 1–20. doi:10.1080/02684527.2018.1479345
GovCERT.ch. (2016). APT Case Ruag – Technical Report. *MELANI: GovCERT Publication Online*. Retrieved from https://perma.cc/2XKP-4FAX
Graham, R. (2014, December 19). The FBI's North Korea evidence is nonsense. *Errata Security Blog*. Retrieved from https://perma.cc/F2ZV-JZXP
Grindal, K., Kuerbis, B., Badiei, F., & Mueller, M. (2018). Is it time to institutionalize cyber-attribution? *Georgia Tech Online Publication*. Retrieved from https://perma.cc/6AZ8-LLVB
Guitton, C. (2014). *Achieving attribution* (Doctoral dissertation). King's College London, London.
Kaspersky Blog. (2017, June 30). From BlackEnergy to ExPetr. *Kaspersky Blog*. Retrieved from https://perma.cc/W7L5-8K22
KCNAWatch. (2014, December 7). Spokesman of policy Department of NDC Blasts S. Korean authorities' false rumor about DPRK. *KCNAWatch*. Retrieved from https://perma.cc/52LQ-A9NQ
Keeley, B. L. (1999). Of conspiracy theories. *Journal of Philosophy*, *96*(3), 109–126. doi:10.2307/2564659
Kello, L. (2013). The meaning of the cyber revolution: Perils to theory and statecraft. *International Security*, *38*(2), 7–40. doi:10.1162/ISEC_a_00138
Kopan, T. (2016a, June 21). DNC hack: What you need to know. *CNN Web page*. Retrieved from https://perma.cc/WR76-QQQF
Kopan, T. (2016b, September 28). Is Trump right? Could a 400-pound couch potato have hacked the DNC? *CNN Web page*. Retrieved from https://perma.cc/ZWV3-LTRN
Krebs, B. (2014, December 23). The case for N. Korea's role in Sony hack. *Krebs on Security Blog*. Retrieved from https://perma.cc/Z96T-ZVAP
Lawrence, P. (2017, August 9). A new report raises big questions about last year's DNC hack. *The Nation Web page*. Retrieved from https://perma.cc/53HA-H4LE

Lindsay, J. R. (2015). Tipping the scales: The attribution problem and the feasibility of deterrence against cyberattack. *Journal of Cybersecurity*, *1*, 53–67. doi:10.1093/cybsec/tyv003

Lukashevich, A. (2014, December 25). Briefing by Foreign Ministry Spokesman Alexander Lukashevich, 25 December 2014. *The Ministry of Foreign Affairs of the Russian Federation Web page*. Retrieved from https://perma.cc/7DBQ-4XDS

Lupovici, A. (2016). The "attribution problem" and the social construction of "violence": Taking cyber deterrence literature a step forward. *International Studies Perspectives*, *17*, 322–342. doi:10.1111/insp.12082

Marchese, D. (2018). In conversation: Seth Rogen. New York (Vulture). Retrieved from https://perma.cc/ZG9W-K6BY

Maschmeyer, L. (2019). Blind spots – Tracking targeted threats to civil society in reporting by the Infosec industry. *Tech & Policy Initiative, Working Paper Series*, *1*(2), 58–90.

Miller, G., Nakashima, E., & Entous, A. (2017, June 23). Obama's secret struggle to punish Russia for Putin's Election Assault. *The Washington Post Web page*. Retrieved from https://perma.cc/PAS9-RBS8

Morgan, P. M. (2010). Applicability of traditional deterrence concepts and theory to the cyber realm. In N. R. Council (Ed.), *Proceedings of a Workshop on deterring cyberattacks: Informing strategies and developing options for U.S. Policy* (pp. 55–76). Washington, DC: The National Academies Press.

Mueller, RS,III. (2019). *Report on the investigation into Russian interference in the 2016 presidential election* (2 vols). Washington, DC: U.S. Department of Justice.

Muirhead, R., & Rosenblum, N. L. (2019). *A lot of people Are saying: the new conspiracism and the assault on democracy*. Princeton, NJ: Princeton University Press.

Nakashima, E. (2016, June 20). Cyber researchers confirm Russian government hack of Democratic National Committee. *The Washington Post Web page*. Retrieved from https://perma.cc/5D48-N2UM

Nincic, M. (2003). Information warfare and democratic accountability. *Contemporary Security Policy*, *24*, 140–160. doi:10.1080/13523260312331271849

Nothaft, H., Pamment, J., Agardh-Twetman, H., & Fjällhed, A. (2018). Information influence in Western democracies: A model of systemic vulnerabilities. In C. Bjola, & J. Pamment (Eds.), *Countering online propaganda and extremism: The dark side of digital diplomacy* (pp. 352–366). London: Routledge.

Nye, J. S. (2017). Deterrence and dissuasion in cyberspace. *International Security*, *41*(3), 44–71. doi:10.1162/ISEC_a_00266

Obama, B. H. (2014, December 19). Remarks by the President in year-end press conference [Transcript]. *Obama White House Web page*. Retrieved from https://perma.cc/AQ6V-MEXN

Office of the Coordinator for Cyber Issues. (2018, May 31). Recommendations to the President on deterring adversaries and better protecting the American people from cyber threats. *Publication U.S. Department of State Web page*. Retrieved from https://perma.cc/A8UG-QEXN

Office of the Director of National Intelligence (ODNI). (2017). Background to "Assessing Russian activities and intentions in recent US elections": The analytic process and cyber incident attribution. *U.S. Office of the Director of National Intelligence Publication*. Retrieved from https://perma.cc/29XQ-UYQM

Office of the Director of National Intelligence (ODNI). (2018). "A guide to cyber attribution". *U.S. Office of the Director of National Intelligence Publication*. Retrieved from https://perma.cc/9UJL-72BG

Pearson, J. (2015, February 4). Infamous hacker 'Yama Tough' says Russians hacked Sony. *Motherboard Web page*. Retrieved from https://perma.cc/F4HW-XH79
Perera, D. (2014, December 16). DHS: No credible threat to Sony movie launch. *Politico Web page*. Retrieved from https://perma.cc/RTU6-JYZZ
Poznansky, M., & Perkoski, E. (2018). Rethinking secrecy in cyberspace: The politics of voluntary attribution. *Journal of Global Security Studies*, *3*, 402–416. doi:10.1093/jogss/ogy022
Ramsay, G., & Robertshaw, S. (2018). *Weaponising news: RT, sputnik and targeted disinformation*. London: King's College London.
Rid, T. (2016a, July 25). All signs point to Russia being behind the DNC hack. *Motherboard Web page*. Retrieved from https://perma.cc/L322-ZWWK
Rid, T. (2016b, October 20). How Russia pulled off the biggest election hack in U.S. History, *Esquire Web page*. Retrieved from https://perma.cc/SCQ4-PL67
Rid, T. (2017). Disinformation: A primer in Russian active measures and influence campaigns. *U.S. Senate Select Committee on Intelligence Web page*. Retrieved from https://perma.cc/3FA6-22XWPL67
Rid, T., & Buchanan, B. (2015). Attributing cyber attacks. *Journal of Strategic Studies*, *38*, 4–37. doi:10.1080/01402390.2014.977382
Rogers, M. (2014a, December 21). Why I *still* don't think it's likely that North Korea hacked Sony. *Marc's Security Ramblings Blog*. Retrieved from https://perma.cc/DG6Z-2ZE9
Rogers, M. (2014b, December 24). Fooled again. No, North Korea didn't hack Sony. *Daily Beast Web page*. Retrieved from https://perma.cc/SH3W-K5MB
RT. (2016, June 16). 'Guccifer 2.0' releases hacked DNC docs revealing mega donors, Clinton collusion. *RT Web page*. Retrieved from https://perma.cc/ADD5-E77W
Russian Embassy in the USA [RusEmbUSA]. (2018, February 15). #Peskov on allegations that pin NotPetya cyber attack on @Russia: We categorically reject such accusations, we consider them unsubstantiated, groundless. This is nothing more than the continuation of the Russophobic campaign lacking any evidence 7/2017 → http://tass.com/politics/955000. *Twitter Post*. Retrieved from https://perma.cc/MNK5-JTZP
Satter, R. (2019, January 26). Undercover agents target cybersecurity watchdog. *Associated Press Web page*. Retrieved from https://perma.cc/L3SL-4M2V
SBU. (2017, July 1). SBU establishes involvement of the RF special services into Petya. A virus-extorter attack. *SBU Web page*. Retrieved from https://perma.cc/HE27-EQBD
Schneier, B. (2014, December 24). Did North Korea Really Attack Sony?. *Schneier on Security Blog*. Retrieved from https://perma.cc/Y43F-B92L
Schulzke, M. (2018). The politics of attributing blame for cyberattacks and the costs of uncertainty. *Perspectives on Politics*, *16*, 954–968. doi:10.1017/S153759271800110X
Seal, M. (2015, February 4). An exclusive look at Sony's hacking saga. *vanity fair web page*. Retrieved from https://perma.cc/9ZKF-BJDF
SecureWorks. (2016, June 16). Russian threat group targets Clinton campaign. *SecureWorks Blog*. Retrieved from https://perma.cc/HWQ7-MPDF
Shahani, A. (2014, December 26). Doubts persist on U.S. claim of North Korean role in Sony hack. *NPR Web page*. Retrieved from https://perma.cc/FN56-7EK2
Solomon, H. (2018, 18 May). RightsCon report: Universities should form cyber attribution network. *IT World Canada Web page*. Retrieved from https://perma.cc/226K-LL4E

Stevens, C. (2019). Assembling cybersecurity: The politics and materiality of technical malware reports and the case of Stuxnet. *Contemporary Security Policy*. Advance online publication. doi:10.1080/13523260.2019.1675258

Stone, D. A. (1989). Causal stories and the formation of policy agendas. *Political Science Quarterly*, *104*, 281–300. doi:10.2307/2151585

Taia Global. (2014, December 24). Taia global linguists establish nationality of Sony hackers as Russian, not Korean. *Webarchive Web page*. Retrieved from https://perma.cc/YT3M-RPER

Taia Global. (2015). The Sony Breach: From Russia, No Love. *Webarchive Web page*. Retrieved from https://perma.cc/WG8R-ZCJM

ThreatConnect. (2016, June 17). Rebooting Watergate: Tapping into the Democratic National Committee. *ThreatConnect Web page*. Retrieved from https://perma.cc/7WLC-8ZN3

United Kingdom of Great Britain and Northern Ireland, Foreign and Commonwealth Office. (n.d.). *UK's approach to the attribution of cyber incidents*. Personal copy.

United States Department of Homeland Security & Director of National Intelligence. (2016, October 7). Joint Statement from the Department of Homeland Security and Office of the Director of National Intelligence on Election Security. *Homeland Security Web page*. Retrieved from https://perma.cc/L5UT-QX6X

United States Department of Homeland Security & Federal Bureau of Investigations. (2014, December 24). November 2014 Cyber Intrusion on USPER I and Related Threats. *First Look Media Web page*. Retrieved from https://perma.cc/X6YN-XY26

United States of America v. Netyksho et al., No. 1:18-cr-215. (2018). *United States District Court, Western District of Washington*. Retrieved from https://perma.cc/V6CP-LJMB

United States of America v. Park Jin Hyok, No. MJ18-1479. (2018). *United States District Court, Central District of California*. Retrieved from https://perma.cc/SQA4-DUXR

United States Office of the Inspector General. (2018, June). A review of various actions by the Federal Bureau of Investigation and Department of Justice in Advance of the 2016 Election. *U.S. Department of Justice Publication*. Retrieved from https://perma.cc/8EVS-P9BF

Valeriano, B., & Maness, R. C. (2014). The dynamics of cyber conflict between rival antagonists, 2001–11. *Journal of Peace Research*, *51*, 347–360. doi:10.1177/0022343313518940

Visschers, V. H., & Siegrist, M. (2008). Exploring the triangular relationship between trust, affect, and risk perception: A review of the literature. *Risk Management*, *10* (3), 156–167. doi:10.1057/rm.2008.1

Weaver, N. (2014, December 18). Why it's possible North Korea was behind the Sony hack. *Mashable Web page*. Retrieved from https://perma.cc/F3BD-TSZE

White House. (2014, December 18). Press briefing by the Press Secretary Josh Earnest. *Obama White House Web page*. Retrieved from https://perma.cc/JS4N-JTYL

Wright, J. (2018, May 23). Cyber and International Law in the 21st century. *Gov.UK Web page*. Retrieved from https://perma.cc/UTT9-NKPB

Zetter, K. (2014, December 17). The evidence that North Korea hacked Sony is flimsy. *Wired Web page*. Retrieved from https://perma.cc/JMG6-FQA4

Zetter, K. (2015, January 8). Critics say new evidence linking North Korea to the Sony hack is still flimsy. *Wired Web page*. Retrieved from https://perma.cc/F8W2-ZLA