JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 2021, 1–12
doi: 10.1093/cybsec/tyab002
Research Paper

Research Paper

Attribution and Knowledge Creation Assemblages in Cybersecurity Politics

Florian J. Egloff [1]* and Myriam Dunn Cavelty

[1]Senior Researcher in Cybersecurity, Center for Security Studies (CSS), ETH Zürich, Haldeneggsteig 4, IFW, 8092 Zürich, Switzerland and Research Associate, Centre for Technology and Global Affairs, Department of Politics and International Relations, University of Oxford, Oxford, United Kingdom. E-mail: florianegloff@ethz.ch. Twitter: https://twitter.com/egflo (@egflo); [2]Senior Lecturer for Security Studies and Deputy for Research and Teaching, Center for Security Studies (CSS), ETH Zürich, Haldeneggsteig 4, IFW, 8092 Zürich, Switzerland. E-mail: dunn@sipo.gess.ethz.ch. Twitter: https://twitter.com/CyberMyri (@CyberMyri).

*Correspondence address. Center for Security Studies (CSS), ETH Zürich, Haldeneggsteig 4, IFW, 8092 Zürich, Switzerland. Email: florianegloff@ethz.ch

Received 2 July 2020; revised 13 January 2021; accepted 21 January 2021

Abstract

Attribution is central to cybersecurity politics. It establishes a link between technical occurrences and political consequences by reducing the uncertainty about who is behind an intrusion and what the likely intent was, ultimately creating cybersecurity "truths" with political consequences. In a critical security studies' spirit, we purport that the "truth" about cyber-incidents that is established through attribution is constructed through a knowledge creation process that is neither value-free nor purely objective but built on assumptions and choices that make certain outcomes more or less likely. We conceptualize attribution as a knowledge creation process in three phases – incident creation, incident response, and public attribution – and embark on identifying who creates what kind of knowledge in this process, when they do it, and on what kind of assumptions and previous knowledge this is based on. Using assemblage theory as a backdrop, we highlight attribution as happening in complex networks that are never stable but always shifting, assembled, disassembled and reassembled in different contexts, with multiple functionalities. To illustrate, we use the intrusions at the US Office of Personnel Management (OPM) discovered in 2014 and 2015 with a focus on three factors: assumptions about threat actors, entanglement of public and private knowledge creation, and self-reflection about uncertainties. When it comes to attribution as knowledge creation processes, we critique the strong focus on existing enemy images as potentially crowding out knowledge on other threat actors, which in turn shapes the knowledge structure about security in cyberspace. One remedy, so we argue, is to bring in additional data collectors from the academic sector who can provide alternative interpretations based on independent knowledge creation processes.

Key words: Attribution, assemblage, cybersecurity politics, knowledge creation process, threat intelligence

1 Introduction

Attribution is central to the politics of cybersecurity. The process of attribution, which involves several phases and spans different communities, establishes a link between technical occurrences and their political implications. An attribution judgement can be viewed as a result of a knowledge creation process through which

uncertainties about who is behind an intrusion and what the likely intent of it was are reduced, thereby giving some cyber-incidents political significance.[1]

Shared and accepted knowledge about the crucial parameters of a cyber-incident is necessary if technical, legal or political action is to be taken in retaliation. It is thus little surprising that, as part of the increasing budgets for cybersecurity, the investment in attribution capabilities – manifested in skilled people with the right knowledge, organisations and processes to enable such people, and technology and data to support both – has increased in the private and the public sector (for an example, see [1]). Such investments are in part underlying the increasing number of attribution judgements offered in the public domain.

Given such developments, attribution practices and processes are not only considered a practical-operational issue but have also become the focus of scholarly publications (exemplary, but not exhaustively, see [2–18]). The literature has covered the scope and limits of attribution processes, as well as the uncertainty existing at an international level regarding the judgements reached by individual stakeholders from the public and the private sector. The majority of existing publications on attribution assumes that there are "facts" about cyber-incidents that improved attribution capabilities can uncover. As a notable exception, Lupovici describes the "attribution problem", i.e. the difficulty or even impossibility of establishing who is behind a cyber-intrusion, as socially constructed and highlights both the human agency in the construction of the underlying technology and the practices that stabilize the interpretation of the "attribution problem" [19].

Conceptually, attribution processes can be split into first, mechanisms that lead to a public attribution [20] and second, what happens after an incident is publicly attributed [21].[2] Both parts open up relevant questions of how attributive knowledge is created, established, and disseminated. In the former, one can distinguish the sense-making process of attribution, which refers to the knowledge creation process that leads to an attribution judgement. Subsequently, there is a meaning-making process in which an attribution judgement is communicated to others in order to change the uncertainty structures associated with the particular intrusion and to exert political effects [16, 22]. We are deliberately not focusing extensively on the latter here. However, it is noteworthy for future research that when attribution judgements have been introduced publicly, they enter a contested information environment, where attackers and other motivated parties contribute to destabilizing attribution claims leading to fractured narratives around the responsibility for a specific intrusion [17].

In this article, we aim to look at how attributive knowledge is created, established, and disseminated in the sense-making phase. We argue that attribution processes have to be scrutinized more carefully in order to understand how public knowledge about cyber-incidents that includes certainty about the perpetrators stabilizes and affects cybersecurity politics. We ask what "truths" – understood here as temporarily stable and accepted knowledge about the parameters of cyber-incidents – are created by whom, when, and based on what kind of assumptions and previous knowledge in attribution processes?

The writing of this article was triggered in part by the observation that there is an increasing "normalization" of enemy images in attribution judgments, evident in technical reports and governmental statements. They follow very familiar and long-standing patterns of enmity that overlap with existing enemy images, particularly US strategic rivals (see e.g. [23]). This has a direct impact on scholars in the field of cybersecurity politics. Attribution judgments from the private and public sector provide the largest empirical basis for scholars studying cyber-conflict. For example, based on such attributions, the scholarly community has concluded that many of the more spectacular cyber-incidents that we are aware of are not just isolated occurrences but should rather be understood in the context of great power rivalry [24, 25]. That made us wonder, are these really the only actors that are active on our networks or might we have become blind towards other developments in the process of honing in on these strategic rivals? Academically speaking, are our inferences about the strategic use of cyberspace and the effects on international politics correct or is it possible that we are missing important dynamics because we depend on data that looks to be highly selective?

In a critical security studies' spirit, we purport that the "truth" about cyber-incidents that is established through attribution is constructed through a knowledge creation process that is neither value-free nor purely objective. This process is based on a series of assumptions and choices made by different groups of people along the way. It is our goal to identify these assumptions and choices in each step of the attribution process to understand how they focus the attention of analysts on specific aspects while eclipsing others. This is in line with work from (critical) intelligence studies that holds against the assumption that "raw data" or neutral facts can ever be collected in the intelligence cycle, given how previous decisions on what to collect and the established practices and methods that are used to target data, create very specific cognitive and heuristic dependencies [26]. However, we do not aim to judge these processes as suboptimal or skewed and will refrain from using the word "bias". Rather, by highlighting assumptions along the way, we want to open up the discussion about the possibility of alternatives. Our contribution is thus to be understood as an intervention to inject alternative views and possibilities into the discourse, not to replace one way of doing things with another.

Analytically, we focus on knowledge creation assemblages. The concept of "the assemblage" highlights complex networks that are never stable but always shifting, assembled, disassembled and reassembled in different contexts, with multiple functionalities [27]. This suits attribution processes particularly well, since they always consist of several phases (we look at three, as outlined in the next section), involving a different set of actors and knowledge steps. To illustrate an attribution process with its assumptions and choices, we use the intrusions at the US Office of Personnel Management (OPM) discovered in 2014 and 2015. The case consists of two known intrusions, allowing us to trace the assumptions driving

1 We deliberately do not want to define the concept of "attribution" further than this nor are we interested in excluding particular forms of attribution. Important for this article and our argument are the knowledge creation processes that are geared towards the reduction of uncertainty. What form this judgment takes – for example, whether it attributes to a particular network or geographic location without identifying an actor group or identifies threat groups or individuals – plays no substantial role here. Furthermore, we do not pay attention to what political actions follow an attribution judgment; i.e. whether states decide to indict someone, impose sanctions, or other retaliatory measures is secondary for our argument.

2 A note about the use of the word "public" here. Since we are primarily interested in the creation of public knowledge – knowledge that is accessible and shared in the public space – we do not look at attributions that are non-public. However, the steps in the knowledge creation process that we are describing in this article are the same for all attributions, whether they are publicly communicated or not.

knowledge creation more than once in a single case. The case is very well documented, allowing us to delve into technical and organizational details. We use a qualitative document analysis, guided by several theoretical concepts we outline below.

The article has three parts. In the first, we situate our undertaking in existing critical cybersecurity studies research, with specific attention on the theoretical concept of "assemblage". The attribution assemblage and its knowledge creation process can best be dissected by paying attention to three lenses: assumptions about threat actors, entanglement of public and private knowledge creation, and self-reflection about uncertainties in the process. We then introduce three temporal phases of the attribution knowledge creation process: incident creation, incident response, and public attribution. In the second part, we use each of the three lenses to analyse an empirical case across the three phases of the attribution knowledge creation process. In the conclusion, we summarize our findings and discuss their implications.

2 Cybersecurity and Knowledge Creation Assemblages

Two fundamentally different meta-theoretical views shape the way we go about scholarly projects: The positivist worldview stands for the belief that it is possible to represent the objective truth about a study object if adequate methods are used. In contrast, the post-positivist worldview stands for the belief that there is no truth outside of our representation of it. Our ways of pursuing knowledge are never neutral but subjective and embedded in a historically grown system of practices that tell us "how to do things the right way". The first view dominates research on cybersecurity politics.

Intellectual disagreement on how to study issues of politics are part and parcel of academia – debates about ontology, epistemology, and methodology are at the heart of some of the most fruitful key debates in International Relations (IR) and security studies [28]. At the same time, however, they tend to sharply divide the discipline. We are not interested in adding to this division – nor do we believe that it is fruitful to fight old, entrenched battles over different conceptions of science and their respective value. Rather, we intend this article to be an invitation for cybersecurity scholars and practitioners to reflect upon normalized practices, without claiming superior intellectual ground.

In this article, we purport that cybersecurity understood from a post-positivist vantage point takes its known shape through a series of knowledge creation processes and that it is those we need to study in order to understand various political forms and implications. In order to systematize the empirical study of attribution that follows, this section does two things. First, it introduces relevant post-positivist literature in order to identify key assumptions that guide this article, particularly highlighting the concept of the assemblage as analytically fruitful. Second, it will present a generic, three-phase attribution model as a knowledge creation process and will briefly discuss methodological issues that arise when studying processes that are partially hidden from the public eye.

2.1 Cybersecurity as Assemblage

Post-positivist cybersecurity studies are carving out a niche for themselves in a variety of (mainly European-based) journals [29–34]. In contrast to the relatively narrow set of questions traditional international relations scholars focus on by using strategic studies' concepts and theories with roots in the Cold War [35–37] there is no single topical focus in the post-positivist literature. Nonetheless, it is united by the assumption that cybersecurity comes into being through an interactive, non-hierarchical, multi-layered assemblage of people, objects, technologies and ideas, is thus co-produced between a wide range of users, institutions, laws, materials, protocols, etc. [22, 31].

In the words of one of the originators of the concept, an assemblage is "a multiplicity which is made up of many heterogeneous terms and which establishes liaisons, relations between them (...) Thus, the assemblage's only unity is that of co-functioning" [38]. The most radical philosophical consequence of the theory of assemblages is that it does not assume we already know the finished shape or product of what we analyse. An assemblage has no essence and no fixed defining features but is contingent on "social and historical processes to which it is connected" [39]. There is no finality, but a continuous, observable effort in the form of multifaceted practices to produce and stabilize cybersecurity, including the creation of technical and political facts in shifting networks, an idea influenced by science and technology studies (STS) [40].

There are three interrelated lenses into the "knowledge creation assemblages" that we will focus on. Below, we explain their importance.

2.1.1. Public-private production of attributive knowledge

First, it is noteworthy how the knowledge creation efforts of public and private actors are intermeshed in intriguing ways in the threat intelligence space. Cybersecurity companies[3] have played a role from very beginning of the cybersecurity story. Not only were technical experts often called upon for testimonies in parliamentary settings, they were also paramount in forging common images of "good" and "bad" hackers, whereas the "good" hacker is usually employed by a company, follows the law and is considered "a professional". This way, hacking was gradually turned into "a service rather than a risk, and hackers become a valuable resource rather than a threat" (p.114 [34]), creating the foundation for a striving segment of the IT security market, which in turn effects the attribution space. Furthermore, private entities have come to the fore as "norm entrepreneurs" in emerging technology governance arrangements, giving them a much more active role in the shaping of political matters than previously acknowledged [41, 42]. Calls for international attribution standards or an international attribution organization is one of the demands made in this context [7, 43–45].

More recently, there is a growing interest in how certain cybersecurity companies – especially threat intelligence companies – are connected to state policy and practice at the national and international level. In her recent publication, Stevens calls Symantec's reports on Stuxnet a "landmark occurrence in the emergence of commercial cybersecurity expertise in the context of strategic state cyber-operations", which makes it "an important constitutive element in wider practices of hardening facts about threats" (pp.130-131, [33]). The point that intrigues us about this is how the traditional division between public and private, between national security interests and commercial interests are blurred in the attribution process. Technical reports get entangled with politics, while at the same time, political attributions are entangled with economic and

3 If we say "cybersecurity companies" we mean companies that "supply" cybersecurity, i.e. whose services one buys to increase the security of one's digital networks.

commercial incentives. Knowledge created by the private sector feeds into larger processes of political attribution, thus shaping which incidents become visible, are communicated and acted upon (similar points have been raised by [31, 46]).

2.1.2. Uncertainty in attributive knowledge creation

The second lens we want to focus on is how these knowledge creation assemblages deal with uncertainties – how are they perceived, communicated and managed? (see also [47, 48]). We observe that most of the ideas behind basic collection practices in private threat intelligence companies and in the more traditional intelligence agencies do not differ in fundamentals though they may differ in some of the details. Therefore, we turned to intelligence studies to get further insights. In general, intelligence studies have expended significant effort examining reasons for and remedies to "intelligence failures". An "intelligence failure" is present if the intelligence community fails to provide the right type of knowledge to policy makers in a timely manner even though they should have been able to, which says as much about expectations as it does about actual performance [49]. The assumption behind the classification of "failures" as such is that if the system of knowledge production were optimized, the probability of failures would be reduced.

Uncertainty is a key aspect of such "failures". The fundamental question in intelligence circles has often been how uncertainties can be communicated or better managed in general [58]. The father of modern intelligence analysis Sherman Kent proposed a standard set of verbal expressions still in use today (words of estimative probability, WEPs) that an analyst can use to express uncertainty in what they think is a uniform and unambiguous way – a practice we also see in attribution processes [50]. The "truth" about cyber-incidents can only stabilize when uncertainties are removed from the narratives – a practice we will focus on in our analysis of the knowledge creation assemblage.

2.1.3. Actor-centric and actor-agnostic knowledge creation processes

The third lens is related to assumptions that guide attribution methods, which are directly related to the active reduction of uncertainties from the very beginning. One of the key reasons why attribution is considered achievable today in contrast to initial debates that centred on the anonymity of action in cyberspace, is a shift away from mainly technical considerations towards the premise of human habits. Habits develop in any persistently conducted human activity, which means that we can track and trace the known habits of threat actors across time and space. To further explain this, we distinguish between actor-centric and actor-agnostic knowledge creation processes. Actor-centric knowledge creation processes focus on creating knowledge about specific threat actors and derive security controls from that knowledge. The paradigm they adopt is: the more knowledge you have about threat actors, the better you can defend yourself against them. This stands in contrast to actor-agnostic knowledge creation processes, which focus on other forms of protection. The paradigm they adopt is: the more you know about one's own network, technologies, data, and practices, the superior your judgements about what is considered 'malicious' behaviour will be (exemplary for this view, see U.S. NSA official Rob Joyce's presentation [51]). We expect the interplay between the two paradigms and particularly the choice of one paradigm over another to be highly relevant for the knowledge creation assemblage.

In the next subsection, we want to describe the details of this knowledge creation process step by step in order to then identify the assumptions and choices behind it as well as the omissions that are taken into account, deliberately or not.

2.2 Sense-Making in Three Phases

We classify the attribution knowledge creation processes (sense-making) into three temporal phases. In the first, a security event is noticed and, upon initial assessments, established as an incident; the second is about incident response, and the third looks at the dissemination of attributing information. Whilst public attribution is part of a meaning-making process (communicating an attribution judgement), in contrast to previous work by one of the authors [17] here we do not look at the further effects of public, official attributions and what kind of further political processes they set in motion (the meaning-making process part of attribution). Rather, we are particularly interested in public attribution's function as an updating of the knowledge assemblage, where knowledge from public attributions is used again as an input for defensive practices that lead to incident creations.

For the purpose of our conceptualization, it is a public attribution if any actor that is part of the knowledge creation process disseminates knowledge about a cyber-incident publicly. This judgment can be, and often is, contested, but since the attribution process always aims to reduce uncertainties and attribution judgments contain authoritative statements about the "truth", an attribution judgment narrows interpretative possibilities and leads to a sedimentation of knowledge. Throughout this process, the entanglement of public and private actors and processes, assumptions and previous knowledge about threat actors, as well as practices to reduce, manage, silence or foreground uncertainties are paramount.

2.2.1 Phase 1: How Security Events Become Incidents

Social order in cyberspace is produced by the heterogeneous relations within and through relevant assemblages. The ultimate aim of cybersecurity as a practice is to stabilize these assemblages, which are there to execute a specific performance, namely the uninterrupted provision of specific data flows for the efficient functioning of the economy, society, and the state. The success of such a stabilization is the degree to which it does *not* appear to be a network that demands effort for keeping it together, but rather a coherent, independent entity that "just works" [52]. This desired state, however, is repeatedly challenged by security events, some of which are elevated to what we call "cyber-incidents". Latour calls these moments of disruptions are called *depunctualization* [53] because they make network performances "break down". Luckily for researchers, in such moments, parts of the assemblage become visible to the observer, allowing us to study previous hidden aspects of the knowledge creation process [54].

Generally, in larger organisations, security relevant events are monitored in security operation centers (SOCs), which have in recent years increasingly been tasked with integrating incident response, threat intelligence, and threat-hunting capabilities.[4] SOCs are tasked with sorting through security relevant events (often called 'alerts') and do initial assessments of whether an "incident" should be established. Thereby, an incident requires incident response and thus leads us to the next phase in the knowledge generation process.

4 Gartner Identifies the Top Seven Security and Risk Management Trends for 2019. https://www.gartner.com/en/newsroom/press-releases/2019-03-05-gartner-identifies-the-top-seven-security-and-risk-ma (12 August 2019, last accessed).

As the discovery of a security related event is a "disruption" and offers the potential for radically new knowledge to emerge, it is useful to reflect on where security related events come from and on what type of knowledge they are based on. We distinguish security related events that are generated as a result of actor-centric or actor-agnostic security knowledge creation processes. Actor-centric approaches have gained prominence in recent years, with the threat intelligence market advertising itself as an enabler to "spot" known actors on one's networks. To further understand actor-centric security controls, and how they interact with other actor-centric knowledge creation processes (specifically threat intelligence) it is insightful to look at an example. We use threat hunting as such an example (Note: there are other techniques, such as using technical threat intelligence feeds directly to "spot" malicious traffic, or, in its most elementary form, the part of an anti-virus product that relies on previously seen malware (i.e. signatures)).

Threat hunting has been proposed as one measure to reduce the uncertainty about the presence of threat actors in an organisation's network, particularly those threat actors that are expected to be able to circumvent the baseline security controls [55–57]. Threat hunting relies on the stance of assuming that your network has been breached and looking for second and third order consequences of threat activity (beyond "just" looking for security relevant events). Threat hunting, as the name implies, relies on previous hypotheses on what types of threat actors might be active on your network and, if they were, where evidence of their activity might be found. Thus, the hunting aspect of the practice refers to the active search for activity previously hypothesized about.

In order to judge the impact an actor-centric practice such as threat hunting might have, it is pertinent to observe where those hypotheses come from, i.e. what knowledge creation assemblages they are embedded within. As said, the practice aims to reduce the uncertainty around the presence of threat actors on an organisation's network. The information about those threat actors can be gathered from public and private sources: other security professionals, past breaches, public reporting, or, in recent years an industry tailored to this specific information need has emerged: threat intelligence [58].

Threat intelligence aims to provide actionable information that enables an organisation to better defend itself against the threats it faces. Threat intelligence as marketed includes many services, from atomic indicators of compromise, to tactics, techniques, and procedures of specific threat actors, to contextual analysis of the goals, strategies, and long-term motivations of threat actors. Threat intelligence is most adept at tracking persistently operating actors, whilst having more difficulty at identifying novel, one-off operations. Threat actor-centric security controls are natural follow-ons, operationalizing threat intelligence into data collection and monitoring practices and hypotheses tailored to one's own networks. If the threat intelligence is useful, it follows that organisations – ceteris paribus – that adopt threat actor-centric practices are more likely to find specific threat actors that they were informed about, than those that they were not informed about.

Consequently, through the practice of threat intelligence, organisations that adopt actor-centric practices, reinforce the knowledge creation assemblage that threat intelligence shapes. Threat intelligence serves as a spotlight that illuminates a very specific part of the threat space but leaves other parts in the dark. A recent study of publicly available threat intelligence reports found public reporting with regard to civil society organisations to be skewed towards traditional adversaries of the West (China, Russia, Iran, North Korea) [42]. Furthermore, a study of two commercial threat intelligence providers showed that the two providers cover very different indicators, even as they pertain to the same threat actors, suggesting a low-level of coverage of their overall operations, even for the most tracked threat actors [59]. The input of threat intelligence into security controls contributes to a self-reinforcing cycle, as threat intelligence is partially based on the output of incident response and attribution processes. We will return to this aspect in the analysis below. Here it suffices to point out that actor-centric defensive practices are more likely to reinforce pre-existing knowledge assemblages than generic defence strategies that are not tailored to a particular actor.[5]

To conclude, the discovery of incidents itself can skew the knowledge creation process in a certain direction. Timely indicators of compromise are often provided by cybersecurity vendors, who themselves are not actor-agnostic. Thus, via threat intelligence, the indicators nudge the "discovery" towards very particular types of intrusions (and away from others).

2.2.2 Phase 2: Incident Response

This section covers the second generic phase of our overall analysis of knowledge creation practices, namely, the incident response process. Incident response starts when an incident is established (output of Phase 1) (For an incident definition, see [60]). Most incidents are minor and can be dealt with swiftly. Some are major and need in-depth incident response, often involving cybersecurity companies offering specialized incident response capabilities and sometimes involving government resources (law enforcement, intelligence, and/or technical help). Here, we are interested in these major incidents, particularly, incidents that are deemed to be caused by threat actors (i.e. incidents assessed to be based on intentional human activity, not based on technical sources or inadvertent human mistakes).

When evidence of activity by a threat actor on a network is found, large uncertainties open up: questions such as what happened, and, more specifically, how did the threat actor pursue their goals on your networks, what were those goals, and were they achieved, loom large. Incident response processes are thus specific kind of knowledge creation processes where we can see the knowledge creation assemblage at work. Those incident response processes geared towards finding the original breach(es) (as opposed to those that focus on recovery at the detriment of evidence collection) try to establish the answers to the question of what happened. Rid and Buchanan refer to this as the "tactical" (what?) and "operational" (how?) parts of the attribution process [1].

In incidents with significant consequences, a relevant audience will want to know who, and why, they engaged in this malicious activity against the organisation (referred to by Rid and Buchanan as the "strategic" part of the attribution process). We refer here to the relevant audience, as this audience may be incident/organisationally specific, and dependent on who else may know of the incident. The audience often includes senior decision-makers that have the responsibility to decide what to do in response to an incident. Particularly, decision-makers may care a lot about the intent of the intruder, an element of attribution that can often be hard to substantiate with robust data.

5 It is important to highlight once more that there are many defensive information controls that lead to indicators of compromise, which are not tied to threat intelligence. We just note that the trend to using threat intelligence as a 'lead' to find actors on the network directs energy and attention towards a specific type of knowledge creation process.

The whole process of incident response and attribution may happen away from public scrutiny. Most jurisdictions do not mandate public breach disclosure, unless a special legally protected type of data was touched (examples are personally identifiable data or health data), or the breach is expected to have a significant business impact (i.e. material event), in which case, publicly listed companies may carry a financial obligation to disclose [61]. Only few incidents are constructed under the scrutiny of the public eye. Thereby, the publicity may shape the knowledge creation and knowledge dissemination practices. This public element is the focus of the next section.

2.2.3 Phase 3: Public Attribution Practices

Under public attribution practices, we understand any public knowledge dissemination by any actor involved in the knowledge creation process about a cyber-incident. To relate it to our conceptualization of incident response as a knowledge creation practice above, public attribution is a knowledge dissemination practice that establishes public knowledge. In recent years, public attributions of cyber-incidents have been increasing, including attributions to attain political effects [54]. A public statement assigning blame to a specific party can thereby address multiple audiences.

For example, public attribution can be used as a tool for uncertainty reduction in a particular audience. For incidents that are publicly known (e.g. data breaches), there exists uncertainty about the context and purpose of the intruder. This context and purpose are important particularly to third parties that are affected by the incident (e.g. customers, suppliers, investors, citizens). Public knowledge of an incident without any context around the possible identity of the intruder and its purpose fuels the uncertainty about the response that other parties would demand of an organisation.

Any data released publicly by the diverse public and private actors involved in the knowledge creation processes will be linked to and embedded in existing knowledge. The newness of the data thereby has the potential to change and update existing knowledge assemblages. Particularly the current public representation of a threat actor that the data released is associated with, feeds back into the threat intelligence process. Thereby, this association of one's public data dissemination into current knowledge is not a process that the breached organisation can fully control. As soon as information about the intrusion becomes public, other actors may start to observe the public elements of the intrusion and publish their interpretations of it (e.g. on a threat intelligence company's blog post). Thus, publicity starts these processes of (re-)assembly of the knowledge creation assemblage, with other actors acting upon that new knowledge and reconfiguring themselves to the changed circumstances.

An additional complication in the case of cyber-incidents is that the full data, on which one bases one's attribution claims, are rarely disclosed, due to sensitivities around proprietary data, as well as sources and methods of the intelligence provider. This brings the discloser in a credibility dilemma: the less data one discloses, the more one has to rely on a general sense of trust in one's own institution by the audiences addressed, opening up possibilities for adversarial strategies to discredit one's attribution claims (see [17]).

Publicly disclosed information about intruders is picked up by the specialist press and sometimes by the more generalist media. Truth claims about particular incidents have the potential to reinforce or challenge pre-existing strategic narratives. To best illustrate this, in our next section, we will look at a specific case and trace the knowledge creation processes throughout those three phases using the three lenses introduced above.

3 OPM: The Attribution Knowledge Creation Assemblage in Action

This section applies the concepts we introduced above to a particularly impactful and well-documented set of intrusions at the United States Office of Personnel Management (OPM). Why did we choose this case? First, the case consists of *two* sets of intrusions that took place between ca. 2012 and 2015. The first intrusion (here intrusion A) was discovered in 2014, the second (intrusion B) in 2015. In intrusion A the intruders familiarized themselves with the network layout at OPM. In intrusion B (presumably using information derived from intrusion A), the intruders exfiltrated personnel records pertaining to people applying for a US government security clearance. This creates an interesting dynamic: we are able to witness the limits of the knowledge creation process, as the organisation gets hacked a second time during the remediation of the first. It provides us with a great illustration of the power of actor-agnostic security controls and how they interact with actor-centric incident response practices.

Second, we deliberately chose a very well documented case in order to be able to highlight the details of the knowledge creation processes. Stuxnet, another case we could have used to show the working of the assemblage, has already received a lot of attention [41, 62–64]. In other cases, such as the Sony Pictures Hack, it would be more interesting to study the contestation process (and knowledge re-assemblage) that happens in the meaning-making phase (as was done in [22]). It should be noted here anew that we do not claim that the attribution in the OPM case could have been done better or that the attribution to China is wrong. Our aim is to "simply" demonstrate how the knowledge creation assemblage works.

Thirdly, the OPM case is also an important case that scholarship should treat in-depth. It is embedded into international political processes, as it represents not economic espionage (something the U.S. at the time tried to get China to stop), but rather classical espionage at scale. Further, it represents somewhat of a turning point, as classical espionage at scale is something that the U.S. would later come to recognize as strategically significant and reorient its policy of restraint to persistent engagement, which seeks to engage adversaries continuously, i.e. before becoming a victim as was the case in OPM [65]. Finally, scholarship would also come to debate whether and how espionage at scale ought to be regulated more [22, 25, 66–71]. All three reasons make OPM worthy of in-depth study.

Building on our conceptual toolset developed above, here we will analytically focus on three particular lenses onto the knowledge creation assemblages in the following order: (1) actor-centric and actor agnostic knowledge, (2) public and private actors, and (3) whether and how uncertainty is represented. These three are present across the different temporal phases of the knowledge creation process (incident creation, incident response, and public attribution). We start with the actor-centric and actor-agnostic lens, as it is particularly relevant for the first two temporal phases (in the third phase, public attribution, one necessarily has an actor-centric perspective). Through each lens, we will examine different ways the practices interact with the larger knowledge creation assemblage. Thereby, we deliberately stay relatively technical in focus, as our

aim is to follow the knowledge creation process as closely as possible.

To do so, we will draw on the majority staff report of the Committee on Oversight and Reform by the U.S. House of Representatives dated 7. September 2016 [72]. We acknowledge that this is an imperfect and politically one-sided, i.e. adversarial to the government, source, though, we also note that the part of the document that we are interested in, i.e. the timelines of the intrusions rather than the political appreciation thereof (especially pp. 51-172), is often sourced in documentary evidence from the time of the intrusions by the actors involved or in congressional testimony by actors involved. We thus consider the document reliable to retrace the knowledge creation processes at the time. Where appropriate, we also draw connections to the broader knowledge creation assemblages and how they existed at the time.

3.1 Actor-centric and actor-agnostic knowledge creation processes in action

Both actor-centric and actor-agnostic knowledge creation processes are present in the OPM intrusions. Intrusion A was originally discovered by a perimeter protection appliance named Einstein, possibly operated by an Internet Service Provider, who flagged it to US-CERT, which then reported it to OPM in March 2014.[6] The Einstein appliance (version 1 or 2) at the time used unclassified threat information to detect and block known threats on classified [73]. We can thus classify it as an actor-centric perimeter defence, using known threat information. Verifying this initial warning from US-CERT by analysing their record, OPM had enough forensic evidence of adversary activity within their networks to elevate the security related event into an incident (p. 53, [54]). Interestingly, the incident response knowledge creation process that it triggered would find that the threat actor was active on OPM's network since at least July 2012 (p. 64, [54]).

After five days, OPM had established the "who and what" and "what [the hackers] are interested in (p. 54 [54])." From then on, they worked on the how the intruders were able to get into and operate within the network. For the incident response OPM used actor-agnostic technologies, such as full-packet capture of any traffic going to the command-and-control server and traffic traversing to/from the most sensitive/high-value part of their system, as well as forensic imaging software. In conjunction with US-CERT, OPM used that to monitor the intruders until 27. May 2014, when they removed all compromised systems, exchanged all potentially compromised account credentials, and forced all Windows administrators to use hardware based personal identity verification cards (p. 60, [54]).

Intrusion B was originally discovered with an actor-agnostic technology (Websense). A contractor had been tasked to assist the adoption of a new Websense functionality and noticed the a "certificate error for the domain called opmsecurity.org" on 14 April 2015 (p. 84, [54] quoting [74]). OPM discovered that an "alert" to this unknown SSL certificate was discovered on 24. February 2015 and that traffic was leaving the OPM network since December 2014 (p. 85, [54] quoting [75]). An initial investigation showed four malicious binaries, three suspicious IP addresses, and a number of indicators in the domain and certificate registration that produced red flags for the OPM security team: thus, an incident response process was initiated.

During the incident response, both actor-agnostic and actor-centric knowledge creation processes were leveraged. Particularly, actor-agnostic technologies were used to extract those four malicious binaries and further in reverse engineering the malware. The result was then compared to previously known malware (in this case PlugX variants) (p. 99, [54]). The results of the initial investigation could also be corroborated with previously existing reporting on the same campaign (e.g. against an insurance company, Anthem) (p. 87, [54]). Furthermore, OPM would testify that they were "uncomfortable with trusting that we knew all the indicators of compromise. And so we obtained the Cylance endpoint client and deployed it [...]. Cylance was able to find things other tools could not 'because of the unique way that Cylance operates. It doesn't utilize a standard signature of heuristics or indicators, like normal signatures in the past have done, it utilizes a unique proprietary method" (pp. 101-102, [54]).[7] Thus, they were specifically looking for an actor-agnostic technology. This new technology provided them with more visibility and identified 41 pieces of malware on different parts of OPM's network (p. 102 & p. 108, [54]). It also meant that OPM, with the help of Cylance engineers, had to sort through lots of false-positives (alerts that were not security relevant). The incident responders traced the actor's activity through logs and forensic images back to the first appearance on 7 May 2014, 20 days before the remediation of Intrusion A went into effect. Had it not been for the use of actor-agnostic knowledge creation processes, Intrusion B may have been found much later, or worse, not at all.

3.2 Public and private actors are producing attributive knowledge

Public and private actors both were fundamentally intertwined in a knowledge creation assemblage during the two OPM intrusions in all three stages: incident creation, response, and attribution. In incident creation, both in Intrusion A and B, public and private actors and technologies played a role in discovery. In Intrusion A, it was likely a private actor using a public technology (Einstein) that notified US-CERT of OPM's network beaconing out to a known command-and-control server. In Intrusion B, a contractor installing a new version of a private technology noticed the intrusion and flagged it to OPM.

In the incident response processes, private and public actors and technologies worked alongside to create knowledge. This included private contractors managing other private incident response and protection technology providers (Mandiant, Cylance, Encase, CyTech), but also teams (and likely technologies) from across government (FBI, NSA, DHS). Interestingly, both CyTech and Cylance provided services in advance, without having a contractual security of getting paid. In CyTech's case, they provided services "to OPM out of a sense of duty and with the expectation that there would be a contractual arrangement put in place" (p. 133, [54]) but ended up not getting paid. Together, the public and private actors produced knowledge on what happened in each of the intrusions.

The attribution processes are less well documented. The oversight report quotes a Washington Post article which posits that the government has "chosen not to make any official assertions about attribution," but also mentions that officials have hinted at China being the leading suspect (p. 157, [54]). Drawing on both the details of the investigation at OPM (including testimony), and integrating this with private sector threat research, the oversight report makes

6 P.52 FN206 notes that it was first detected via an Einstein device and notes that it was "possible" that this device was operated by an Internet Service Provider. We follow that interpretation here.

7 The proprietary method is explained in the document on pp. 93-94 as being a classification method of every action happening on an endpoint.

concrete claims about what intruders were behind the intrusions (namely Axiom behind Intrusion A and DeepPanda behind Intrusion B). It makes these claims drawing on knowledge that was available in the public domain, and creating new knowledge by showing how the intrusions fit into existing knowledge assemblages. Note: by pointing out that much of the attributing knowledge to a particular actor is still classified, the report adds weight to the public claims it is making, as it implies that the authors have awareness of the classified assessments and, one assumes, they do not want to deliberately mislead the public (p. 167, [54]).

The report particularly lays emphasis on a malware found in Intrusion A, aliased by industry with the name Hitkit, to be strongly associated with Axiom. It quotes various industry reports to support that judgement. Note that the knowledge creation processes that lead to the public "truth" do not just rely on "official" statements made by the victim. Various indirect ways of assembling knowledge in the public domain, including through public and private actors, were important for the public attribution processes.[8] For example, on 10 July 2014, the Washington Post reported authorities having traced the Intrusion A to China, but not having identified yet whether the intruders work for the government, based on an anonymous U.S. official [76].

For Intrusion B, the report documents not only the overlapping infrastructures used against various other targets with US government personnel, but also quotes the knowledge making industry engaged in in linking these intrusions. For example, on 27 February 2015, two weeks before Intrusion B was discovered by OPM, ThreatConnect published a blog post outlining part of the attack infrastructure used in Intrusion B and attributing it to Deep Panda [77]. Furthermore, the oversight report cites from testimony documenting OPM personnel connecting the intrusion to DeepPanda, worth quoting in full:

> "So I'll use the word 'actor', the ones that were identified in prior exhibits. You had Shell Crew, or sometimes known as Deep Panda, as well as Deputy Dog, and it has many, many other names. So those were the two that, at least as it relates to industry research being done, that the malware that we found was closest related to it. By no means are we saying it was them; it's just it was a relationship or similarity" (p. 167, [54]).

Similarities and differences are what clusters these knowledge creation processes. We can also note that there are significant difficulties in naming actor groups, not only between different industry players, but also for people in government, who would be using a different set of clustering to generate their own actors/activity sets. As with Intrusion A, after Intrusion B became public, the Washington Post reported that the intrusion was conducted by Chinese state-sponsored actors based on anonymous officials and private sector reporting [78].

3.3 Uncertainty differs across the three stages of attributive knowledge creation

Uncertainty was represented differently across the three stages of interest. First, in the incident creation phase, certainty and radical uncertainty interplay with one another. On the one hand, OPM makes clear that about Intrusion A, there are a number of factors one cannot know, as OPM did not have logging in place for certain actions. Thus, the report concludes that the US "will never know with complete certainty the universe of documents the attacker exfiltrated" (p. 51 [54]), thereby recognizing fundamental constraints on the "knowable". On the other hand, it is forensic evidence that establishes certainty of "actual adversary activity" that necessitates the opening of incident response (as opposed to "just" having the presence of malware). For example, in Intrusion B, the discovery of a Windows Credentials Editor was found to be confirming the presence of an adversary with ill intent (p. 97, [54]). Thus, forensic evidence is used to close-off uncertainties.

Second, in the incident response phase, the defenders are dealing with the uncertainties of how deep the adversaries are buried in the network. If possible, the defenders need to establish the breadth and depth of compromise, both for effective remediation and for damage assessment. In Intrusion A, this phase lasted from March-May 2014, in which the defenders watched the intruders move laterally on the network and prepared the infrastructure to remediate unexpectedly (what OPM called the "big bang").

Third and finally, in the attribution phase, judgements represent uncertainties by using estimative language standardized in the intelligence community. Thus, the oversight report uses "likely", both for Intrusion A and Intrusion B, to indicate that uncertainty around the attribution to a specific threat actor (p. 17, [54]). Nevertheless, it quotes Director of National Intelligence James Clapper that referred to China s "the leading suspect" (p. 157, [54]). Thus, by representing the findings of the investigation about Intrusion A and Intrusion B as likely, and putting them into the context of the statement by the DNI, whilst not presenting any alternative explanations, the oversight report narrows the interpretative possibilities and encourages the sedimentation of knowledge in the form of "truth".

The oversight report omits the part of DNI Clapper's quote that many people, including two out of the three anonymous peer-reviewers, remember and hence were crucial for the knowledge sedimentation. It was not the uncertain, nuanced, and potentially open-ended "leading suspect" part of the quote that was spread widely in the media and scholarship, but rather this more definitive statement: "you have to, kind of, salute the Chinese for what they did. If we had the opportunity to do that, I don't think we'd hesitate for a minute' [79]. Thus, by attaching the operation to the Chinese (government) and integrating and legitimising it using the U.S. operational framework as a basis, the DNI left no doubt about the provenance of the intrusions.

4 Conclusions

Cybersecurity knowledge creation processes in the form of public attribution judgments fundamentally shape what we, the general public, know about cyber-threats and their political implications. This article offered a science-technology studies inspired lens towards understanding such knowledge creation processes. We argued that through using the concept of the knowledge creation assemblage, we can shed light onto important dynamical aspects that explain how knowledge about cyber-incidents is created. To do so, we conceptually split up the knowledge creation process into three temporal phases (incident creation, incident response, and public attribution). This conceptual slicing was used to describe which sets of knowledges, actors, and technologies are informing each phase, and, through so doing, to analytically highlight how each phase interacts with the larger knowledge creation assemblage.

8 In principal-agent theory these processes would be interpreted as forms of «proxies». In assemblage theory, we note that the knowledge creation assemblage is fully and continuously intertwined between public and private actors.

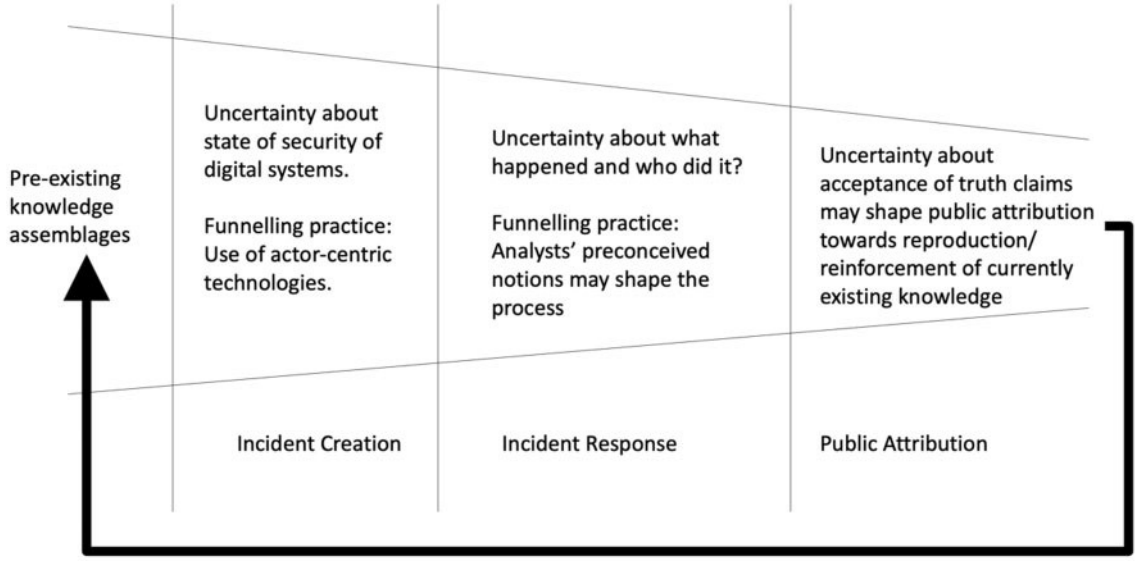

Figure 1: Self-reinforcing tendencies in cybersecurity assemblages

We used three analytic lenses (actor-centric and actor-agnostic, public and private, and how is uncertainty represented) to unravel the knowledge creation processes in these three phases in an empirical case with two intrusions. In the first analytic lens, we highlighted how actor-centric technologies and processes lead to a reinforcement and differentiation of current knowledge. Contrary to that, actor-agnostic technologies and processes have the possibility to broaden and disrupt it.

The second analytic lens showed how public and private actors seamlessly work together in the incident creation and response processes observed. With some private actors operating out of sense of "duty", it also reconfirmed how blurry the distinction public/private gets in the field of cybersecurity [23, 26, 80]. In the attribution knowledge creation processes, the private sector reports were particularly relevant, building up an overall knowledge assemblage into which the new knowledge gained as a result of the incident response knowledge creation processes are integrated. Importantly, the oversight report that we drew on used private sector actor grouping aliases (Axiom and Deep Panda) to assign responsibility.

The third analytic lens identified how uncertainty is present/absent in the different stages. It is both a driver and a constraint for the knowledge creation processes observed. Thus, whilst uncertainty about the status of one's network can be a driver for practices trying to discover security relevant events, the creation of an incident is a moment of radical uncertainty. It is at that point that an organisation uses the incident response knowledge creation processes to displace that uncertainty. Attribution, in that sense, is an afterthought from an organisational perspective. The resulting attribution judgements and the way that they are represented were found to make use of language standardized in the intelligence community. Thereby uncertainty is represented with words of estimative probability.

What can we conclude? Knowledge creation processes are messy and contingent and without further in-depth case studies, generalizations beyond the case we studied in these pages will be impossible. However, it does seem apparent that assumptions and choices in these processes produce certain outcomes – or rather, make certain conclusions more likely than others. One of the consequences such practices inevitably have is that they feed into additional knowledge creation processes, for example, into academia. Following our three analytical lenses, we draw the following conclusions:

First, the public-private nexus demands more scrutiny. Foremost, there is a need for systematic study of private threat intelligence reports. We know that market logics are not usually based on fairness but favour the financially potent in the public and the private sector who can pay for attribution services (p. 61, [17]). A recent empirical analysis of commercial threat reporting shows convincingly that "high end threats to high-profile victims are prioritized in commercial reporting while threats to civil society organizations, which lack the resources to pay for high-end cyberdefense, tend to be neglected or entirely bracketed" [81]. If the goal is more security for everyone, this is a problem. Maschmeyer, Deibert, and Lindsay offer a promising route by studying the publicly available threat reporting and comparing it with targeting of civil society organisations, thereby demonstrating a systematic skewing of the public threat reporting due to the commercial incentives to do so. This work can be extended by studying the private threat reporting and assessing to what degree this public skewing is also present in the private intelligence products (for a promising start, see [58]). However, we could expect that, as the threat intelligence market matures and digitalization deepens across the globe, threat intelligence companies from different parts of the world may broaden the insight into the diverse actors engaged in offensive cyber behaviour, leading to a more global insight into cyber conflict overall.

Second, it is necessary to focus even more closely on the representations or non-representation of uncertainties in attribution processes. From a critical scholar's perspective, the current practices mask the socially constructed nature of intelligence and by extension the practical handling of uncertainties [82]. Uncertainties have a tendency to disappear from the discourse and from view; they get masked by the practices themselves. This highlights the need to "pay attention to how information is defined, created, managed, and used in particular contexts." (p. 659, [26]). In fact, a focus on the knowledge creation assemblage reveals that "uncertainty" manifests differently in different phases of the process and that the practices of reducing those uncertainties are manifold. Those particular variations in uncertainties could also be linked to particular "funnels" in the attribution process (see Figure 1), where particular types of practices that react to different types of uncertainties narrow (i.e. funnel) the knowledge construction process.

Each phase is associated with different uncertainties, which, coupled with specific ways of acting, could further exacerbate the self-reinforcing tendencies of the intelligence cycle, here denoted with a "funnel". Actor-centric technologies are an example of a funneling practice reacting to the uncertainty about the state of security of digital systems in the incident creation process. During the incident response phase, analysts' preconceived notions (not discussed in this paper) may become relevant in shaping the attribution process. Finally, during the public attribution phase, uncertainty about the acceptance of truth claims may shape, what knowledge seeps out in the public domain, and leads to concerns about the reproduction and reinforcement of currently existing threat narratives. We suggest that further researching the nuances between the interactions of different uncertainties and their practices with funneling tendencies could lead to a more fine-grained analysis in the future.

Third and finally, the actor-centric and actor-agnostic lens showed the impacts current trends in knowledge creation practices can have on the overall knowledge structure. Crucially, we are not suggesting abandoning actor-centric practices. Rather, we highlight the importance of reflecting on the trade-offs between practices that lead to the discovery of known actors vs. practices that have the potential to lead to the discovery of unknown actors. We expect, in practice, a combination of both will be required.

Closely related to this is the question of who could provide different data to challenge dominant threat images. We note academia has mostly focused on being interpreters of the analysis produced by public attribution statements and has used it to study cyber-conflict, thereby partially recreating the knowledge assemblages shaped by the actor-centric and actor-agnostic trade-offs of other actors. Only few academic institutions (notably the CitizenLab, CrySySLab, and Civilsphere) have independently collected forensic artefacts, upon which incident response knowledge creation processes were used, leading to an independent assessment, one that differed from the one represented by the other actors described.

But universities themselves are connected to targeted intrusions, be it as victims, operational infrastructure, or as host of victims (see e.g. in times of SARS-CoV-2 vaccine research the fears of cyber espionage affecting the integrity of clinical trials [83]). In that sense, universities could contribute more to an independent systematic data collection on cyber-incidents, allowing for comparisons between their datasets and private actors' telemetry, thereby contributing an independent viewpoint into cyber-conflict. Universities should invest in more interdisciplinary knowledge on attribution practices, thereby empowering themselves to be in a better position to assess other attribution outcomes of industry and government reports.

The measures addressed above ameliorate the transparency in knowledge creation assemblages. Knowing the type of knowledge creation processes that make up the knowledge assemblage that is "the cyber-threat" enables a more transparent discussion of the shaping of realities of cyber-conflict faced by different actors internationally. Thereby, we can work towards challenging and reformulating the current narrative of cyber-conflict into a more open conversation about how different communities worldwide experience cyber-conflict.

Acknowledgments

A previous version of this article has been presented at the 2019 Conference on Cyber Norms at The Hague, at a workshop on "Exploring the socio-cultural fabric of digital (in)security" in August 2020, and at a research colloquium at the Center for Security Studies at ETH Zürich in October 2020. We received further comments from Timo Steffens and Lilly Pijnenburg Muller. On all occasions, we received valuable feedback. We gratefully acknowledge Jasper Frei's feedback and assistance with the referencing. Special thanks go to three anonymous reviewers whose comments helped us sharpen our argument further.

Reference

1. Raiu C. Attribution 2.0. In: *Area41 Conference, Zurich*, 2018. https://perma.cc/5FZ4-6SJR.
2. Rid T, Buchanan B. Attributing cyber attacks. *J Strateg Stud* 2015;**38**:4–37.
3. Lin, H. Attribution of malicious cyber incidents: from soup to nuts. *J Int Aff* 2016;**70**:75–137.
4. Deibert, RJ. Toward a human-centric approach to cybersecurity. *Ethics Int Aff* 2018;**32**:441–24.
5. Eichensehr, KE. Decentralized cyberattack attribution. *AJIL Unbound* 2019;**113**:213–17.
6. Eichensehr, KE. The law & politics of cyberattack attribution. *UCLA Law Rev* 2020;**67**:520–98.
7. Finnemore M, Hollis DB. Beyond naming and shaming: accusations and international law in cybersecurity. *Eur J Int Law* 2020;**31**:969–1003.
8. Delerue, F. *Cyber Operations and International Law. Cambridge Studies in International and Comparative Law*. Cambridge: Cambridge University Press, 2020.
9. Mikanagi, T, Mačák, K. Attribution of cyber operations: an international law perspective on the Park Jin Hyok Case. *Camb Int Law J* 2020;**9**:51–75.
10. Steffens, T. *Attribution of Advanced Persistent Threats: How to Identify the Actors Behind Cyber-Espionage*. Berlin: Springer, 2020.
11. Grindal, K, Kuerbis, B, Badii, F *et al. Is It Time to Institutionalize Cyber-Attribution?, Internet Governance Project White Paper.* Atlanta: Georgia Tech, 2018.
12. Grotto, A. Deconstructing cyber attribution: a proposed framework and lexicon. *IEEE Secur Priv* 2020; **18**:12–20.
13. Guerrero-Saade, JA. Draw me like one of your french apts—expanding our descriptive palette for cyber threat actors. In: *Virus Bulletin Conference, Montreal*, 2018. 1–20.
14. Guerrero-Saade, JA, Raiu, C. Walking in your enemy's shadow: when fourth-party collection becomes attribution hell. In: *Virus Bulletin Conference, Madrid*, 2017. 1–15.
15. Bartholomew, B, Guerrero-Saade, JA. Wave your false flags! deception tactics muddying attribution in targeted attacks. In: *Virus Bulletin Conference*, Denver, CO, 2016. 1–9.
16. Guerrero-Saade, JA. The ethics and perils of apt research: an unexpected transition into intelligence brokerage. In: *Virus Bulletin Conference, Prague*, 2015. 1–9.
17. Lindsay, JR. Tipping the scales: the attribution problem and the feasibility of deterrence against cyberattack. *J Cybersecur* 2015;**1**:53–67.
18. Romanosky, S, Boudreaux, B. Private-sector attribution of cyber incidents: benefits and risks to the U.S. Government. *Int J Intell CounterIntelligence* 2020:1–31.
19. Lupovici, A. The "Attribution Problem" and the social construction of "Violence": taking cyber deterrence literature a step forward. *Int Stud Perspect* 2016;**17**:322–42.
20. Egloff, FJ. Public attribution of cyber intrusions. *J Cybersecur* 2020;**6**:1–12.
21. Egloff, FJ. Contested public attributions of cyber incidents and the role of academia. *Contemp Secur Policy* 2020;**41**:55–81.
22. Egloff FJ. Cybersecurity and non-state actors: a historical analogy with mercantile companies, privateers, and pirates. DPhil Thesis. University of Oxford, 2018.
23. Roth, F, Stirparo, P, Bizeul, D, *et al.* APT Groups and Operations. https://web.archive.org/web/20201217131838/ https://docs.google.com/spreadsheets/u/1/d/1H9_xaxQHpWaa4O_Son4Gx0YOIzlcBWMsdvePFX68EKU/pubhtml
24. Harknett, RJ, Smeets, M. Cyber campaigns and strategic outcomes. *J Strateg Stud* 2020:1–34.

25. Kostyuk, N, Zhukov, YM. Invisible digital front: can cyber attacks shape battlefield events? *J Conflict Resolut* 2019;**63**:317–47.
26. Räsänen, M, Nyce, JM. The raw is cooked: data in intelligence practice. *Sci Technol Human Values* 2013;**38**:655–77.
27. DeLanda, M. *Assemblage Theory*. Edinburgh: Edinburgh University Press, 2016.
28. Dunn Cavelty, M, Wenger, A. Cybersecurity meets security politics: complex technology, fragmented politics, and networked science. *Contemp Secur Policy* 2019;**41**:5–32.
29. Stevens, T. *Cybersecurity and the Politics of Time*. Cambridge: Cambridge University Press, 2016.
30. Balzacq, T, Dunn Cavelty, M. A theory of actor-network for cyber-security. *Eur J Int Secur* 2016;**1**:176–98.
31. Collier, J. Cybersecurity assemblages: a framework for understanding the dynamic and contested nature of security provision. *Politics Gov* 2018;**6**: 13–21.
32. Shires, J. Enacting expertise: ritual and risk in cybersecurity. *Politics Gov* 2018; **6**:31–40.
33. Stevens, C. Assembling cybersecurity: the politics and materiality of technical malware reports and the case of Stuxnet. *Contemp Secur Policy* 2019;**41**:129–52.
34. Tanczer, LM. 50 shades of hacking: how IT and cybersecurity industry actors perceive good, bad, and former hackers. *Contemp Secur Policy* 2020;**41**:108–28.
35. Maness, RC, Valeriano, B. The impact of cyber conflict on international interactions. *Armed Forces Soc* 2016;**42**:301–23.
36. Borghard, ED, Lonergan, SW. The logic of coercion in cyberspace. *Secur Stud* 2017;**26**:452–81.
37. Kello, L. *The Virtual Weapon and International Order*. London: Yale University Press, 2017.
38. Deleuze, G, Parnet, C. *Dialogues*. New York: Columbia University Press, 2002, 69.
39. Nail, T. What is an assemblage? *SubStance* 2017;**46**:24.
40. Latour, B, Woolgar, S. *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press, 2013.
41. Hurel, LM, Lobato, LC. Unpacking cyber norms: private companies as norm entrepreneurs. *J Cyber Policy* 2018;**3**:61–76.
42. Gorwa, R, Peez, A. *Big Tech Hits the Diplomatic Circuit: Norm Entrepreneurship, Policy Advocacy, and Microsoft's Cybersecurity Tech Accord*. https://berlinpolicyjournal.com/big-tech- hits-the-diplomatic-circuit/ (30 June 2020, last accessed).
43. Egloff, FJ, Wenger, A. Public attribution of cyber incidents. In: Merz, F (ed.). *CSS Analyses in Security Policy No. 244*. Center for Security Studies: Zurich, 2019, 1–4.
44. Solomon, H. RightsCon report: universities should form cyber attribution network. *IT World Canada Web Page*, 2018. https://perma.cc/226K-LL4E
45. Mueller, M, Grindal, K, Kuerbis, B, and Badiei, F. Cyber attribution: can a new institution achieve transnational credibility?. *Cyber Defense Rev* 2019;**4**:107–22.
46. Leander, A. Understanding US national intelligence: analyzing practices to capture the chimera. In: Best, J, Gheciu, A (eds.), *The Return of the Public in Global Governance*. New York: Cambridge University Press, 2014, 197–221.
47. Canton, B. The active management of uncertainty. *Int J Intell CounterIntelligence* 2008;**21**:487–518.
48. Slayton, R. Governing uncertainty or uncertain governance? information security and the challenge of cutting ties. *Sci Technol Human Values* 2021;**46**:81–111.
49. Jensen, MA. Intelligence failures: what are they really and what do we do about them? *Intell Natl Secur* 2017;**27**:261–82.
50. Kent, S. Words of estimative probability. *Studies in Intelligence*, 1964.
51. Joyce, R. *USENIX Enigma 2016—NSA TAO Chief on Disrupting Nation State Hackers, YouTube*, 2016. https://www.youtube.com/watch?v=bDJb8WOJYdA.
52. Callon, M. Techno-economic networks and irreversibility. In: Law, J (ed.), *A Sociology of Monsters: Essays on Power, Technology and Domination, Sociological Review Monograph*, Vol. **38**. New York: Routledge, 1991, 153.
53. Latour, B. *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge: Harvard University Press, 1999.
54. Best, J, Walters, W. "Actor-Network Theory" and international relationality: lost (and found) in translation: introduction. *Int Politic Sociol* 2013; **7**:346.
55. FOR508. Advanced Incident Response, Threat Hunting, and Digital Forensics. https://www.sans.org/event/threat-hunting-and-incident-response-summit-2019/ course/advanced-incident-response-threat-hunting-training (12 August 2019, last accessed).
56. FOR572. Advanced Network Forensics: Threat Hunting, Analysis, and Incident Response. https://www.sans.org/event/threat-hunting-and-incident-response-summit-2019/course/ advanced-network-forensics-threat-hunting-incident-response (12 August 2019, last accessed).
57. FOR578. Cyber Threat Intelligence. https://www.sans.org/event/threat-hunting-and- incident-response-summit-2019/ course/cyber-threat-intelligence (12 August 2019, last accessed).
58. Work, JD. Evaluating commercial cyber intelligence activity. *Int J Intell CounterIntelligence* 2020;**33**:278–308.
59. Bouwman, X, Griffioen, H, Egbers, J, Doerr, C, Klievink, B, van Eeten, M. A different cup of TI? The added value of commercial threat intelligence. In: *Proceedings of the 29th USENIX Security Symposium*, 433–50. San Diego, CA: USENIX Association, 2020.
60. ISO. *ISO/IEC 27000:2018*. Geneva, 2018.
61. Exemplary, see Security and Exchange Commission, 17 CFR Parts 229 and 249 [Release Nos. 33-10459; 34-82746] Commission Statement and Guidance on Public Company Cybersecurity Disclosures. https://www.sec.gov/rules/ interp/2018/33-10459.pdf (12 August 2019, last accessed).
62. Langner, R. Stuxnet: dissecting a cyberwarfare weapon. *IEEE Secur Privacy* 2011;**9**:49–51.
63. Lindsay, JR. Stuxnet and the limits of cyber warfare. *Secur Stud* 2013;**22**: 365–404.
64. Slayton, R. What is the cyber offense-defense balance? Conceptions, causes, and assessment. *Int Secur* 2017;**41**:72–109.
65. U.S. Cyber Command. *Achieve and Maintain Cyberspace Superiority: Command Vision for US Cyber Command*, 2018. https://web.archive.org/web/20210108165528/https://www.cybercom.mil/Portals/56/ Documents/USCYBERCOM%20Vision%20April%202018.pdf?ver=2018-06-14-152556-010
66. Libicki M. The coming of cyber espionage norms. *Paper Presented at the 9th International Conference on Cyber Conflict (CyCon)*, 30 May–2 June 2017.
67. Boeke, S, Broeders, D. The demilitarisation of cyber conflict. *Survival* 2018;**60**:73–90.
68. Georgieva I. The unexpected norm-setters: intelligence agencies in cyberspace. *Contemp Secur Policy* 2020;**41**:33–54.
69. Egloff, FJ, Maschmeyer, L. Shaping not signaling: understanding cyber operations as a means of espionage, attack, and destabilization. *International Studies Review*, 2020.
70. Chesney, R, Smeets, M, Rovner, J, Warner M, Lindsay, JR, Fischerkeller, MP, Harknett, RJ, Kollars, N. Policy roundtable: cyber conflict as an intelligence contest. *Texas National Security Review*, 2020.
71. Lindsay, JR. Cyber conflict vs. cyber command: hidden dangers in the american military solution to a large-scale intelligence problem. *Intell Natl Secur* 2020;**36**:1–19.
72. U.S. Congress, House of Representatives, Committee on Oversight and Government Reform. *The OPM Data Breach: How the Government Jeopardized Our National Security for More Than a Generation*. Washington, DC: Government Printing Office, 2016.
73. Written testimony of Dr. Andy Ozment, Assistant Secretary for Cybersecurity and Communications, U.S. Department of Homeland Security, Before the U.S. House of Representatives Committee on Oversight and Government Reform, regarding the DHS Role in Federal Cybersecurity and the Recent Compromise at the Office of Personnel Management, pp. 2–5. https://docs.house.gov/meetings/GO/GO00/20150616/103617/HHRG-114-GO00-Bio-OzmentA-20150616.pdf (30 June 2020, last accessed).

74. H. Comm. On Oversight and Gov't Reform. *Interview of Brendan Saulsbury, Senior Cybersecurity Engineer, SRA, Ex. 4* (17 February 2016). https://archive.org/stream/ReportFromThe CommitteeOnOversightAndGovernmentReformOnTheOPMBreach/Report%20from % 20the%20Committee%20on%20Oversight%20and%20Government% 20Reform%20on%20the%20OPM% 20Breach_djvu.txt (30 June 2020, last accessed).
75. AAR Timeline—Unknown SSL Certificate (15 April 2015) at HOGR020316-1922 (OPM Production: 29 April 2016). https://archive.org/stream/ReportFromTheCommitteeOnOversightAndGovernmentReformOnTheOPMBreach/Report%20from%20the%20Committee% 20on%20Oversight%20 and%20Government%20Reform%20on% 20the% 20OPM%20Breach_djvu.txt (30 June 2020, last accessed).
76. Chinese hackers go after U.S. workers' personal data. *Washington Post*, 10 July 2014. http://www.washingtonpost.com/world/ national-security/chinese-hackers-go-after-us- workers-personal-data/2014/07/10/92db92e8-0846-11e4-8a6a-19355c7e870a_story.html (30 June 2020, last accessed).
77. The Anthem Hack: All Roads Lead to China. https://web.archive.org/web/20200520133650/https://threatconnect.com/blog/the-anthem-hack-all-roads-lead-to-china/(30 June 2020, last accessed).
78. Chinese breach data of 4 million federal workers. *Washington Post*, 4 June 2015. https://www.washingtonpost.com/ world/national-security/chinese-hackers-breach-federal -governments-personnel-office/2015/ 06/04/889c0e52-0af7-11e5-95fd- d580f1c5d44e_story.html (30 June 2020, last accessed).
79. Clapper JR, Brown T. *Facts and Fears: Hard Truths from a Life in Intelligence*. New York: Viking, 2018.
80. Egloff, FJ. Cybersecurity and the age of privateering. In: Perkovich George and Levite Ariel (eds.), *Understanding Cyberconflict: Fourteen Analogies*. Washington DC: Georgetown University Press, 2017, 231–47.
81. Maschmeyer, L, Deibert, RJ, Lindsay JR. A tale of two cybers—how threat reporting by cybersecurity firms systematically underrepresents threats to civil society. *J Inf Technol Polit* 2020;**18**:1–20.
82. Kreuter, N. The US intelligence community's mathematical ideology of technical communication. *Tech Commun Q* 2015;**24**:217–34.
83. Grierson, J, Devlin, H. Hostile States Trying to Steal Coronavirus Research, Says UK Agency. *TheGuardian*, 3 May 2020. https://www.theguardian.com/ world/2020/may/03/hostile- states-trying-to-steal- coronavirus-research-says-uk-agency