COLUMBIA | SIPA
School of International and Public Affairs

Tech & Policy Initiative, Working Paper Series 2

COLUMBIA | SIPA
School of International and Public Affairs

It is with great pleasure that I invite you to read this second volume of Columbia SIPA's Tech & Policy Initiative's Working Paper Series. Building on the insights of the first volume, the second volume features working papers produced by SIPA-supported expert and next generation scholars who are engaging critical areas related to the impact of digital technology on society and institutions. The papers are multi-disciplinary and forward-looking, engaging complex subjects including the critical areas of Internet and data governance, the dynamics of cyber conflict and cyber sovereignty, how digital technology has impacted traditional economic sectors and business models, or other areas.

This academic work was undertaken with vital support of the Carnegie Corporation of New York as part of SIPA's Tech & Policy Initiative, an ambitious effort to explore the digital world and SIPA's core fields of study. Since its inception, the Tech and Policy Initiative has sought to bridge the gap between policymakers, academics and practitioners in cybersecurity, internet governance and the digital economy through convening, research, training and other activities. The Tech & Policy Initiative draws on many disciplines and talented researchers within SIPA, in other parts of Columbia University, and outside entities to develop insights that will translate into better and more effective policies, and to inform government policies and private sector actions.

This volume also includes papers prepared for SIPA's 2017 Global Digital Futures Policy Forum, an annual conference that brought together more than 100 scholars, private sector leaders, legal experts, entrepreneurs, technologists, and others to discuss the challenges of internet fragmentation. It also features papers prepared for the 2017 State of the Field of Cyber Conflict, an annual conference convening academics and practitioners to advance discussion of new developments in cybersecurity, evolving theory alongside operational and practical concerns.

We hope you will find these papers insightful and an illustration of the many opportunities and challenges associated with tackling this complex and ever-evolving field.

Merit E. Janow

Merit E. Janow

Dean, School of International and Public Affairs

Professor of Practice, International Economic Law and International Affairs

TABLE OF CONTENTS

Section 3: Digital Economy 293

Section 4: Internet Governance 348

Section 1:

Next Generation Cyber Fellows Sponsored, by Carnegie

Knowing Thyself, Not Just Thine Enemy: Explaining Delays in the Development of International Rules for Cyber Conflict

Justin Key Canfil[1]

Introduction

The word “cyberattack” has become a household word in recent years. The problems associated with cyber insecurity are pervasive and widespread; by some estimates, the global annual cost of cyberattacks will come to $6 trillion by 2021 (Morgan, 2016). Cyber is considered by the overwhelming majority of scholars to be what Jervis (1978) calls an “offense-dominant” domain, its primary characteristic being that $1 spent on offense costs more than $1 to offset with defense. In Jervis’ (1978) framework, the cyber world is also “doubly dangerous” because offensive tools are often indistinguishable from those used for active defense or espionage.

The traditionally preferred tool for managing offense-dominance is deterrence, but deterrence extremely difficult in cyberspace (Libicki, 2009; Kramer, Starr, and Wentz, 2009). Over the objections of some scholars (Valeriano and Maness, 2015; Rid, 2013), worries about a “cyber Pearl Harbor” abound in the media and defense circles (Stavridis 2019; Clarke and Knake, 2011). Recent news that hackers allegedly connected with the Russian government possessed the ability to hold US critical infrastructure at risk has only exacerbated these concerns (Perlroth and Sanger, 2018). Nor is the US the only vulnerable party. Capability also correlates with vulnerability, since developed countries tend to be the most powerful operators but are also the most-networked (Significant Cyber Incidents, 2018), and US adversaries have been targets, as well (Healey, 2019). The stakes, it would seem, are shared by all.

Despite universal interest in mitigating the risks of cyber conflict, however, powerful states cannot seem to agree on how international law applies. Some Western experts were optimistic after a number of significant breakthroughs in 2015 (Healey and Maurer, 2017), but subsequent

[1] PhD Candidate, Columbia University Department of Political Science; Affiliated Scholar, Department of Law, London School of Economics. j.canfil@columbia.edu - The author would like to thank Columbia University SIPA and the Carnegie Corporation for their generous research support.

negotiations at the United Nations Group of Global Experts (UNGGE) broke down in 2017 when some participants backtracked (Korzak, 2018).

These problems were not a surprise. The international community has been aware of potential problems associated with an ungoverned cyberspace for many decades. Healey and Grindal (2012) mark the earliest inception of cyber conflict as having occurred in 1986. Thirty-plus years later, and despite considerable efforts, no well-articulated model of binding international law applicable to cyberspace has gained universal acceptance.

Despite cyber's novelty, it is not the first time the international community has debated how to regulate a new technology. History has much to teach. This paper argues that, much like in previous cases, uncertainty over the risks and benefits posed by horizon innovations in cyberspace may play an important role in why efforts to solve more immediate problems have been frustrated.

The Problem

The militarization of cyberspace is widely regarded by scholars as posing significant challenges for international politics. Broadly speaking, emergent military technologies can be systematically costly regardless of whether or not they are individually beneficial to possess. New military technologies can destabilize the security environment in several key ways. First, they may shift the balance of power, breaking power-parity equilibria by undoing what states have learned about their relative strength vis-a-vis adversaries. Second, technological breakthroughs can induce arms races or create first-strike incentives. Third, they can bypass or exploit loopholes in existing arms control agreements, rendering cooperative frameworks ineffective.

In cyberspace more specifically, low cost and easy accessibility is thought to even the playing field for less powerful actors (Nye, 2010). Meanwhile, the difficulty of attribution in cyberspace hamstrings traditional deterrence models, upon which stability in other domains has historically relied (Libicki, 2009; Long, 2016). Resulting mutual vulnerability raises the risk of inadvertent or uncontrolled escalation (Kramer et al, 2009). A rich literature on cyber deterrence has emerged, with no consensus yet among scholars about how best to make the model work.

Where deterrence fails, there is arms control. The factors outlined above imply that states should have a discernible collective interest in regulating activity in this space through the development of coherent and binding international law. At the most basic level, law provides stability by instituting "rules of the road" (North, 1990) – benchmarks for actors to evaluate political behavior (Hadfield and Weingast, 2012; Milgrom et al, 1990). This in turn allows members of its community to more easily monitor others' activity and label it as appropriate or inappropriate. Theory implies that, armed with this knowledge, would-be defectors are less likely to take actions that might be deemed inappropriate for fear of being sanctioned by others in the system, materially or otherwise (Guzman, 2010; Keohane, 1999). Even when only marginally enforceable, regulation can reduce uncertainty and dampen the security dilemma.

The usefulness of laws regulating technological *means* is evidenced by the multitude of arms control treaties in existence, as well as in the *lex specialis* body of Hague Law, which regulates certain military technologies in accordance with basic international humanitarian law principles. In cyberspace, the international community has formally debated such regulation at least 1998, when the Russian Federation first broached the topic at the United Nations General Assembly (UN General Assembly A/RES/53/70). The US and other countries have also emphasized the importance of developing some type of cooperative arrangement (Maurer, 2011), although each differs on its ideal formulation. Given their collective interest in mitigating the security dilemma, it is puzzling that states have thus far proven unable to agree on anything but the most basic legal topography for cyber conflict.

Cyber law negotiations are just one case where states have deliberated over the utility of regulation for a new technology. In practice, we see a great degree of variation in the time it takes to arrive at legal consensus. Some weapons are regulated very early on in anticipation of a particular technology. For example, the Strasbourg Agreement of 1675 (signed between the Holy Roman Empire and France) famously banned the battlefield use of primitive chemical weapons -- poisoned bullets -- a technology that did not become widely feasible until the mid- to late-19th century with the advent of the industrial economy (Zanders, 2003). Similarly, the 1967 Outer Space Treaty prohibited the stationing of military assets on celestial bodies a full two years before the first manned lunar landing took place (UN General Assembly Resolution 2222, XXI; Lanius, 2017). Other technologies are regulated at a slower pace, often after wartime experience, while still others are never regulated at all.

Strategic risk aversion is the most commonly cited mechanism for failure in arms control negotiations. Formal and informal accounts have classically been modeled as a prisoner's dilemma between two political players who would like to regulate a technology but cannot trust each other to honor the agreement. These types of games hinge on each player's beliefs about her opponent's intentions, and thus the credibility of her promises (Montgomery, 2006), the essence of the security dilemma. We know, for example, the security dilemma can be escaped by extending the shadow of the future. Axelrod (1986) famously showed that repeated interactions between "forgiving" players can lead to cooperation, and Majeski (2004) extends this finding to argue that cooperation is even possible when capabilities are asymmetric. Institutions such as formalized agreements and binding international law can enhance the visibility of the interactions, thereby locking in present gains (Ikenberry, 2000; Koremenos, 2005; Krisch, 2005).

Despite these predictions, bargaining failures continue to occur. Scholars have cited the presence of malicious "types" in the system who disrupt good-faith interactions (Kydd, 2000), obstacles to monitoring and verification (Fairbanks and Shulsky, 1987; Abbott, 1993; Bendor, 1993), the size of the bargaining coalition (Keohane, 1999; Olson, 1965), and expected enforcement problems as leading mechanisms for the collapse of talks. Another research program points to human fallibility. Jervis (1976) argues compellingly that decisionmakers do not always behave rationally (though they very often expect others to), while Keohane (1984) adds that the rationality of negotiators is bounded (although his argument is actually the opposite -- that institutions are attractive stopgaps for decisionmakers with a finite capacity for rational calculation). As a result, mistakes and misperceptions can unravel mutually-beneficial agreements. While undoubtedly true, this mechanism can also be bidirectional; in some cases, misperceptions have helped solidify agreements (Grynaviski, 2014).

Another school of thought argues that emotion during times of uncertainty may complicate cooperative processes (Mercer, 2010), yet this is more appropriate for crisis bargaining than arms control bargaining, since the former entails shorter time horizons and higher stress levels. Others cite the role of domestic political institutions (Miller, 1984; Putnam, 1988; Morrow, 1991; see also Moravcsik, 1997) and civil society (Bunn, 1999; Price, 1998; Rutherford, 2000; Horowitz, 2016; Carpenter, 2016), although this would imply obstacles are idiosyncratic, yielding few systematic predictions of use to use in understanding the trajectory of the cyber case. Indeed, because the public is uniquely vulnerable to cyberattacks, we might expect public opinion to exert especially high pressure on the US government to negotiate a deal to mitigate risks in the immediate term.

Although few scholars have addressed the question of bargaining over technologies, per se, a dominant literature on strategic restraint contends that technologies are susceptible to bans after formative, usually traumatic, experiences. Chemical weapons during the First World War and nuclear weapons at the end of the Second World War are commonly thought to have spurred a taboo against the use of these technologies, subsequently cemented in comprehensive international treaty regimes (e.g. Price, 1995; Tannenwald, 1999). While certainly true in some cases, the "experience" criterion cannot explain technologies that are regulated preemptively (or never at all). Nor does the modern chemical weapons taboo make an ideal example: the Hague Conventions of 1899 and 1907, ratified by 31 and 35 states, respectively, explicitly banned the use of asphyxiating shells and poisonous weapons prior to the experience of the First World War (see also Brown, 2005).

Existing theory has outlined the myriad mechanisms by which arms control efforts can fail, but it provides little traction on the question of why some technologies are regulated more rapidly than others -- in some cases preceding the emergence of a technology; in others, only after a technology is learned and understood; and in others still, not at all. Even theories that model repeated interactions provide little insight on empirical variation on negotiation duration.

For instance, in a seminal paper, Fearon (1998) argues that states signal their relative strength by displaying a willingness to endure the costs of noncooperation, thus securing better terms. In Fearon's model, states anticipate a stream of future costs and benefits from a proposed arrangement and then make a determination about whether it will pay to sign on. This model naturally assumes that players have information about their own incentives but not whether an adversary can be trusted. But the empirical record reveals that this is often not the case. For example, as early as the Johnson administration, Washington was quite concerned about whether new technologies allowing access to the deep seabed would ultimately be in the US' national interest. It therefore sought to delay any multilateral agreement from taking shape until more could be known.[2]

In fact, this paper argues that outcomes of varied duration may be explained by a different kind of risk-aversion. Regardless of whether there exists any uncertainty about an adversary's intentions, contracting parties may have uncertainty over their *own* best interests. This is corroborated by the fact that technology necessarily entails a learning process. Acquisition alone does not convey an immediate advantage: militaries must optimize doctrine and assess diffusion through a process of theorizing, demonstration, testing, or experience (Biddle, 2006; Posen, 1986 and Horowitz, 2010).

[2] Personal papers of Bromley Smith, LBJ Presidential Library. Austin TX. Accessed in December 2018.

Not until a technology's potential is learned can political actors know their own preferences over how to regulate the environment.

However, when this happens, actors might not like what they find: a technology may end up benefiting one's adversary, leading to regrets about not curbing it sooner. Bargaining parties thus face a tradeoff between (a) closing off potential technological avenues *a priori* or (b) traveling down these avenues at their peril.[3] When the risk of pursuing the latter strategy is considered too high, and when adversaries are cooperative, anticipatory or immediate regulation is possible.

For in informal illustration, imagine two players engaged in a bargaining situation.[4] In a world with complete information, each player considers how much to invest in an arms control package and how to design it -- which weapons systems should receive coverage, how strict to make the obligations, whether to include invasive monitoring provisions, and so on -- and players then consider one another's proposals. Negotiations continue until players locate a mutually satisfactory constellation of terms, in which case a regime emerges, or quit, in which case the space remains ungoverned. When considering proposals, each player looks "down the game tree," calculating whether she could do better over the long-run if terms were adjusted, and whether her opponent would accept such an adjustment. For example, suppose Player A expects Player B's industrial capacity to increase over the next 10 years, making it easier for B to acquire more of weapon system x relative to A's arsenal. All else equal, A might thus reason that a rigid arms limitation treaty would be in her favor and push for such terms.

However, note that it takes two to solidify a bargain. As Schelling (1961) and Jervis (1993) detail, arms control is only possible where states share mutual interest. This implies that, among rational self-interested actors, bargains are only reached in situations of mutual optimism or out of mutual necessity, and only when solving today's problems is prioritized over potential problems over the horizon.[5] This is doubly true in security matters, where the salience of strategic interests is maximized, since unregulated spaces are associated with instability and arms races (see Downs et al, 1985; Downs et al, 1990; Glaser, 2000; Fearon, 2011).

This problem is seriously compounded when Player A has uncertainty not only of B's future potential, but her own as well. Player A might worry about an economic recession, an unanticipated, incremental technological breakthrough by B, a leadership change in B that undoes previous cooperation, or a doctrinal innovation that gives B a comparative advantage in producing

[3] "Closing off potential technological avenues *a priori*"' can be taken to mean either that (1) states abandon the technology and never learn its potential, or (2) learn its potential but are restricted from using it due to reputational (or other) sanctions imposed by the regime upon defectors. I model the process as the first possibility, but it would be simple to model it according to the second by adding in a parameter representing the cost of defection. I decide not to because a large body of research already addresses the question of compliance, which is outside the scope of this paper. If the former approach (no learning) is accurate, states never learn that defection would pay and thus no incentive for defection is created. If wrong, the results are considered to hold for sufficiently high defection costs.

[4] Formal versions of this model have previously been presented at the Institute of World Economy and International Relations (IMEMO) 2018; the International Studies Association 2018 and American Political Science 2018 conferences, the University of Pennsylvania (Nov. 2018), and Leiden University (Nov. 2018).

[5] Mutual necessity is a tricky subject. Jervis (1993) argues that states may not always prefer a vulnerable adversary since this can heighten the security dilemma. I merely take the position that adversarial states prefer stable asymmetric advantage over stable parity.

or fielding the weapon in question. Moreover, Player A may worry that her counterpart B has a relatively more accurate picture of the future -- heightening suspicion over B's proposals.

As the illustration shows, State *A*'s (and reciprocally State *B*'s) willingness to cooperate, as well as the acceptable range of terms, depends crucially on how much each player values freedom of action today. Should a player expect future circumstances to change such that arms buildup is preferable to arms control, the harder it will be to draw that player to the negotiating table now. Similarly, should either state perceive that its opponent believes the same, the former may doubt the ability of the arms control regime to bind the latter in the long-run. Either scenario can cause the collapse of negotiations that at present offer mutual beneficial.

This exercise generates several takeaways. When at least one side is optimistic in the nascent stages of a breakthrough (for simplicity, collapsed into time period *t_1)*, anticipatory regulation is least likely. Only when the technology is later revealed t_2 to be mutually disadvantageous does convergence become likely. Conversely, when both sides are pessimistic in *t_1*, anticipatory regulation is most likely. If the technology is later revealed to present asymmetric advantages, incentive shifts can create compliance pressures for the regime. States may also initially be pessimistic and be proven right. Of course, depending on the strictness of regime provisions, states may never have occasion to learn the true value of the technology (for instance, if the prohibition includes a ban on testing), reducing *ex post* compliance pressures but raising *ex ante* selection pressures (Fearon, 1998).

Empirical Observations

To test this mechanism, I employ a “large-*n* qualitative” study (Fortna, 2004) on 15 technological cases. In the interest of space limitations, I include a discussion of three example cases in detail, below. These technologies are selected because, as extreme cases, they illustrate the mechanism more vividly. While this exercise does not yet purport to offer a truly a systematic investigation, it provides useful information through a comparison of cases with maximum variation on the dependent variable (speed and strictness of regulatory convergence). Further study, in the form archival research and computerized text analysis, will investigate the role of elite beliefs about the future of the security environment for each of these cases.

Example Case: Anticipatory Regulation

In many cases, regulatory consensus on new technologies came swiftly. Consider the 1899 Hague Convention, which in part banned the use of early modes of chemical and aerial warfare. States have long realized the destabilizing effects of battlefield chemical weapons technology. As already mentioned, the first treaty to ban the use of chemical weapons was signed between two countries in 1675, long before the technology was truly feasible. States confronted the issue anew in the 19th century, when rudimentary chemical warfare became truly feasible. Ideas for poisonous weapons were proposed but rejected on moral grounds by the British during the Crimean War (1854) and both the Confederate and Union Armies during the American Civil War (1861 and 1862/1864, respectively). Then, during the Boer War in the last years of the 19th century, the British reversed their earlier position by filling shells with picric acid in at least one instance -- one of the first cases of chemical weapons being systematically deployed on the battlefield. The deployment was an utter failure, with gaseous clouds reportedly wafting back to hit Britain's own troops (Romano et al, 2007). In light of the increasing viciousness of warfare during that time, states convened in The Hague in 1899 to consider formalized humanitarian restrictions on certain weapons technology.

The historical record shows that Tsarist Russia pushed hard for the inclusion of a rule prohibiting the use of chemical weapons (Mazanec, 2015). As the weakest and least industrialized of the negotiating parties, Russia's military leadership must have worried about the country's ability to support advanced research and development into chemical weapons, whereas other countries had robust private chemical industries. Indeed, Russia was decidedly unable to reciprocate against Japanese chemical attacks during the Russo-Japanese War only a few years later [Romano et al. 2007]. Britain's disastrous experience, which was certainly not unknown to British leadership, may have tipped others into the Russian camp. By the end of the conference, and at Russia's urging, parties arrived at mutually acceptable terms for a chemical weapons ban in Art. 23(a). These terms were ratified by 51 states (Convention II, 1899).

As time passed, noncompliance became more widespread. By the end of the First World War, it was clear that use of such weapons in war was disastrous for all parties and not particularly useful in a tactical sense (Price, 1998; Brown, 2005; Main, 2015), leading to a significant decline in use following the signing of the 1925 Geneva Protocol and later its stronger successor, the Chemical Weapons Convention (CWC). Chemical weapons represent an interesting case wherein states exhibited anticipatory pessimism rooted in moral opprobrium while the technology was still over the horizon, then early asymmetric optimism after the emergence of the technology, and finally symmetric pessimism after full information was revealed through battlefield experiences. The superpowers continued to develop and stockpile chemical weapons during the Cold War period in compliance with the 1925 agreement, which banned their use only in wartime, until the CWC was cemented to outlaw this as well. Since then, instances of chemical weapons use have been rare and isolated.

Example Case: Downstream Reversals

It is also known that Russia was a leading advocate for the regulation of air-to-ground warfare in the Hague Convention. The resulting 1899 ban was introduced four years before the first airplane and ten years before the first military aircraft, the 1909 Wright Military Flyer. The terms specifically prohibited "the discharge of any kind of projectile or explosive from *balloons or by similar means*." Aerial vehicles at the time were hardly suited for overhead bombing. Principally, these consisted of large balloons, slow and vulnerable, usually tethered to the ground at a fixed position and used for reconnaissance. The notion of aircraft, however, had captured the imagination of the public since the mid-19th century; by the 1880s gliders had shown that there was possibility in winged flight. Powered aircraft, potentially capable of carrying explosive munitions were now an invention just over the horizon -- one that negotiators clearly anticipated in some form.

Again, the lagging state of Russian industry meant that Moscow probably worried about a technological disadvantage should aircraft become weaponized. Other countries seemed not to have recognized the real potential of aerial warfare, treating the platform in its earliest days as more of a curiosity (c.f. Mitchell, 1925). Skeptical parties were convinced to sign on to a ban with a five-year duration provision, which allowed the issue to be revisited again after more was known about the technology.

In the meantime, aircraft technology was finally developed and allowed to proliferate. Militaries continued to learn about the technology's potential; nothing in Hague 1899 proscribed military experimentation, only use against other parties during wartime. Later attempts to renew the ban in the 1907 Hague Convention led to the adoption of a watered-down compromise ratified by only a handful of states: a prohibition on "the attack or bombardment of undefended towns, villages, etc., of the words 'by whatever means'" (Schindler and Toman, 1988). Ultimately, despite early concerns that aerial ordnance should be regulated, limitations in the 1899 regulatory apparatus allowed states to learn about the military utility of the technology and reverse their earlier positions.

Example Case: Impeded Consensus

Conversely, there are examples where a breakthrough technology was not immediately regulated. Consider the case of anti-satellite weapons (ASATs). These presently come in several forms: co-orbital satellites, which would have the ability to change course and target other objects in space; direct-ascent weapons, which are launched directly at their targets from lower in the atmosphere; and pulses, which can disable electronics in space by bombarding them with electrons or optic-blinding laser beams. Despite some major drawbacks -- namely their propensity to cause permanent damage to the outer space environment and spur arms races in space, ASATs are extremely useful in modern warfighting, which relies to an unprecedented degree on networked communications (Cohen, 1996; Biddle, 2006). The ability to neutralize an adversary's satellites is akin to blinding, gagging, and binding your foe.

ASATs were first envisioned in the late 1950s, when Soviet planners decided to invest in a killer satellite that could attack other satellites. Other states began developing their own methods: the United States noticed electromagnetic emissions after an upper atmosphere nuclear test in 1958, leading military planners to wonder about the potential uses for disabling outer space assets (Operation Hardtrack I Report, 1958). In 1962, another test inadvertently disabled a British satellite, proving the technology's usefulness but also its unpredictability (Plait, 2012; Hollingham, 2018). Contrary to physicists' predictions, the effects of the Starfish Prime test were much broader than anticipated. Furthermore, residual radioactive particles failed to disperse, creating an artificial radiation belt in space that lingered for months (DTIC, 1962). Meanwhile, Soviet plans for a co-orbital satellite finally became operational until 1978 (Peebles, 1983; Hostbeck, 2015). After experimentation, however, the Soviets realized that these vehicles have critical disadvantages: they are slow, complicated, and expensive to operate: due to the difficulty of maneuvering in space, engagement windows are small and time-to-engagement is much larger.

Consequently, direct-ascent weapons became the preferred route of development. Two classes of direct-ascent weapons exist: exo-atmospheric kill vehicles, such as the American ground-based interceptor (GBI), which destroy targets with kinetic force, and high-altitude missiles with a conventional kill mechanism (i.e. an explosive charge) (Hostbeck, 2015). The United States pioneered direct ascent technology beginning in the late 1950s. The first ground-based interception was accomplished in 1963 (Stares, 1985), and in 1982 the US successfully launched an air-to-space missile, the ASM-135 (Puffer, 1985). However, these too were soon recognized to have disastrous side effects: the National Aeronautics and Space Administration (NASA) determined that the debris from ASM-135 had caused substantial and long-lasting environmental damage.

Congress subsequently intervened by imposing restrictions on ASAT testing in 1985. In a rush to learn more about the technology before restrictions were levied, the Department of Defense conducted a final test shortly before the relevant statute went into effect that October (Portree, 1999), but subsequent US administrations have pressured other countries not to field ASATs and the Soviet Union unilaterally offered to ban ASAT testing in 1982 (Hostbeck, 2015). Despite these measures, no formal agreement was signed, and the technology has since proliferated among various actors in the international system. China has notably tested several such weapons in recent years, to which the US has responded in-kind (Space.com, 2008). Reports from the US Office of the Director of National Intelligence warns that China and Russia will both be able to deploy battle-ready ASAT weapon systems within a few years (Coats, 2018).

Given the utility of ASATs for modern warfare, it is unsurprising that no ban has emerged. This is true despite states' collective interest in regulating their use to safeguard the space environment, protect civilian space assets, and mitigate costly arms races that could soon be underway (Covault, 2007). ASATs simply offer too many battlefield advantages, especially to weaker states in a dyad with an interest in hamstringing their opponent's superior conventional capabilities -- for example, compare Chinese and American Pacific naval posture (see Horowitz, 2010). Over the now-50 year history of ASAT research and development, the international community has gradually learned that the costs of ASAT technology is high, but not high enough, seemingly, to offset the expected advantages of retaining freedom of action.

Discussion

As the example cases illustrate, beliefs about the future security environment can motivate or impede international accord. Comparatively weaker states in a dyad have a strong incentive to encourage the proliferation of technologies that they expect to diffuse power in the system, since such technologies offer such states a boost in operational capabilities. Similarly, weaker states may oppose the unrestricted use of technologies that concentrate power, including technologies that are too costly or complex for weaker states to adopt. Stronger states have the opposite incentives. Because international law is forged by consensus, the side favoring restrictions may have to compromise on its ideal terms or offer side payments in order to elicit the support of parties that value freedom of action, since the latter essentially have a veto on the crystallization of any universally binding legal arrangement.

What lessons does history have for ongoing cyber law negotiations? At first glance, predictions might not seem to bear out in the cyber case. As discussed, cyber technology is thought to diffuse power. Consequently, we should expect to see the most powerful states take a more critical stance while comparatively weaker states should embrace it, all else equal. Although data on cyber power is difficult to measure or collect, the US is thought to be among the strongest actors currently operating in cyberspace. However, Washington has been obstinate in its refusal to endorse a cyber treaty, insisting on the one hand that existing law should be sufficient, and on the other calling for nonbinding normative arrangements where gaps persist.

What might explain this pattern? Because cyber is so often treated as a *sui generis* technology, one might expect ambiguity over definitions and concepts to impede negotiations. Yet my research suggests that negotiators – despite, for the most part, lacking technical expertise – feel relatively confident about the definitions and concepts at stake.[6] Instead, US policy toward an international cyber treaty may be driven by divergent beliefs about the future balance of power. The US' Joint Doctrine for Cyberspace acknowledges that "permanent global cyberspace superiority is not possible due to the complexity of cyberspace" (JP 3-12 2008, p. I-2), but the US continues to enjoy an advantage and is host to a large fraction of the internet's infrastructure. Beliefs are pivotal in uncertainty, and beliefs about superiority may cloud assessments about downstream national interest (Snyder, 1989). If military operators believe they are capable of outgunning adversaries – especially if the "gloves are off" – delaying may be seen as worth the risk. [7] Similarly, for actors with long time horizons and plans to grow in strength, patience may pay off even if present circumstances place them at a disadvantage.

Crucially, however, mistakes are also precisely what make agreements *possible* in worlds where *any* party has an incentive to defect. This follows from the theoretical predictions derived from the model, wherein agreements are driven by mutual pessimism about future outlook. Consensus on cyber law has not yet been attained in part, perhaps, because the international community – or at least a fraction of it -- is simply not pessimistic enough about what the world will look like if no universally accepted, binding provisions exist to limit unrestricted cyber warfare. Expectations about how horizon innovations such as artificial intelligence, quantum computing, or a more integrated "internet of things" may alter the risk environment. The advent of such technologies may draw reluctant parties back to the bargaining table.

Conclusion

While the present costs associated with cyber conflict are widely recognized, few states can agree on what an optimal regulatory solution would look like. Beyond uncertainty about bargaining partners' true intentions, expectations about the future of the security environment may play a pivotal role in the success or failure of international regulatory proposals, and these beliefs can fluctuate over time. The difficulty of attaining consensus, short of immediate and overriding necessity, stems from the fact that states are reluctant to bind themselves to terms that may later

[6] Interviews conducted with US and European ministry-level government officials, November 2018. Names omitted to protect anonymity.

[7] From conversations at several academic workshops, including the 2016 and 2017 State of the Field Conferences at Columbia University, New York. Personal attribution prohibited due to Chatham House rules.

prove disadvantageous. The alternative is to refrain from binding oneself at all. Both entail costs, both to the individual and international society. This may explain the slow progress seen in multilateral cyber law fora. As Maurer (2011, p. 5) writes, the "first [UNGGE] group established in 2004 failed to even find the smallest common denominator which forced the Secretary-General to conclude in 2005 that, 'given the complexity of the issues involved, no consensus was reached on the preparation of a final report'" (quoting text from UN General Assembly A/60/202:2).

Frustrated with this process, one way states have sought to circumvent the need for coherent laws on the use of force in cyberspace has been to instead focus on norm-building. While some states, notably Russia, have long advocated a comprehensive cyberspace treaty, other states, like the US, dispute the feasibility of a treaty in such a rapidly changing domain and instead support the development of system of non-binding, normative guidelines. The disadvantage of norms is that they have a softer touch than do binding, formal agreements. Without the reporting requirements, withdrawal provisions, and other measures built into "hard" law, transgressions are harder to monitor the sanctions associated with violating them are often weaker and less regular. They therefore do little to dampen mistrust. Unlike formal multilateral agreements, the most effective norms arise organically, from the bottom-up.

The tentative findings of this paper buttresses existing research that suggests treaty designers can entice reluctant states to the bargaining table by incorporating flexibility provisions (sunset clauses, escape hatches, withdrawal provisions, etc.) as stopgaps against future change (Koremenos, 2005). The shorter the length of time between periods of renegotiation, the fewer gaps between terms and incentives, thus diminishing the threat of noncompliance. Of course, this entails another set of tradeoffs for institutional designers: the more continuous the bargaining process, the less like "law" it looks, and thus the fewer advantages it offers as a focal point. The optimal agreement proposal, if it exists, must balance between these two extremes.

The difficulty of cooperation in cyberspace is driven in large part by reciprocal insecurity – in part, a classic security dilemma -- but also, this paper argues, by uncertainty over the future. Formal agreements lock in behavioral obligations; states that are optimistic about strategic opportunities over the horizon are understandably reluctant to sign on. These expectations -- whether accurate or not -- can derail accord. Beliefs and uncertainty about corollary technological unknowns such as artificial intelligence, quantum computing, and the "internet of things" may only exacerbate these pressures. This paper advances the idea that anticipatory bargaining is likeliest when fear of the future is maximized. If true, a great irony is that the world may in some cases be made a safer place when anxiety is widespread, and more dangerous when we feel most comfortable.

Works Cited

"A Quick Look at the Technical Results of Starfish Prime," DTIC 1962, http://www.dtic.mil/dtic/tr/fulltext/u2/a955411.pdf

Abbott, Kenneth W. (1993). “Trust But Verify: The Production of Information in Arms Control Treaties and Other International Agreements”. eng. Cornell International Law Journal 26, pp. 1–58.

Axelrod, Robert (1984). The Evolution of Cooperation. en. Google-Books-ID: NJZBCGbNs98C. Basic Books.

Bendor, Jonathan (1993). “Uncertainty and the Evolution of Cooperation”. en. Journal of Conflict Resolution 37(4), pp. 709–734.

Biddle, Stephen (2006). Military Power: Explaining Victory and Defeat in Modern Battle. English. unknown edition. Princeton University Press: Princeton, NJ.

Brown, Fredric (2005). Chemical Warfare: A Study in Restraints. English. 1 edition. Routledge: New Brunswick, N.J.

Bunn, George (1999). “The status of Norms against nuclear testing”. The Nonproliferation Review 6(2), pp. 20–32.

Carpenter, Charli (2016). “Rethinking the Political / -Science- / Fiction Nexus: Global Policy

Making and the Campaign to Stop Killer Robots”. en. Perspectives on Politics 14(1), pp. 53–69.

Clarke, Richard A. and Robert Knake (2011). Cyber War: The Next Threat to National Security and What to Do About It. English. Reprint edition. Ecco: New York.

Coats, Daniel R. (2018). Worldwide Threat Assessment of the U.S. Intelligence Community. en-US. Tech. rep. Director of National Intelligence.

Cohen, Eliot A. (1996). “A Revolution in Warfare”. en-US. Foreign Affairs(March/April 1996). Convention (II) with Respect to the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land. The Hague, 29 July 1899, https://ihl-databases.icrc.org/ihl/INTRO/ 150?OpenDocument

Covault, Craig. “China’s ASAT Test Will Intensify U.S.-Chinese Faceoff in Space” (2007), Aviation Week and Space Technology.

Downs, George W., David M. Rocke, and Randolph M. Siverson (1985). “Arms Races and Cooperation”. en. World Politics 38(1), pp. 118–146.

Downs, George W. and David M. Rocke (1990). Tacit Bargaining, Arms Races, and Arms Control.

English. First edition edition. University of Michigan Press: Ann Arbor.

Fairbanks, Charles H. and Abram N. Shulsky (1987). “From “Arms Control” to Arms Reductions:

The Historical Experience”. The Washington Quarterly 10(3), pp. 59–73.

Fearon, James D. (1998). “Bargaining, Enforcement, and International Cooperation”. International Organization 52(2), pp. 269–305.

Fearon, James D. (2011). “Arming and Arms Races”. Annual Meetings of the American Political Science Association, Washington, DC.

Fortna, Virginia Page (2004). Peace Time: Cease-Fire Agreements and the Durability of Peace by Fortna, Virginia Page (2004) Paperback. Princeton University Press.

Glaser, Charles (2000). “The Causes and Consequences of Arms Races”. Annual Review of Political Science 3, pp. 251–276.

Grynaviski, Eric (2014). Constructive Illusions: Misperceiving the Origins of International Cooperation. English. 1 edition. Cornell University Press: Ithaca.

Guzman, Andrew T. (2010). How International Law Works: A Rational Choice Theory. English. 1 edition. Oxford University Press: Oxford.

Hadfield, Gillian K. and Barry R. Weingast (2012). “What Is Law? A Coordination Model of the Characteristics of Legal Order”. en. Journal of Legal Analysis 4(2), pp. 471–514.

Healey, Jason. “China Is a Cyber Victim, Too.” Foreign Policy (blog). Accessed January 20, 2019. https://foreignpolicy.com/2013/04/16/china-is-a-cyber-victim-too/.

Healey, Jason and Karl Grindal, eds. (2013). A Fierce Domain: Conflict in Cyberspace, 1986 to 2012. English. Cyber Conflict Studies Association: Vienna, VA.

Hollingham, Richard (2018). The Cold War nuke that fried satellites. en.

Hostbeck, Lars (2015). “Space Weapons’ Concepts and their International Security Implications”. en. In: Handbook of Space Security. Springer, New York, NY, pp. 955–983.

Horowitz, Michael C. (2010). The Diffusion of Military Power: Causes and Consequences for International Politics. English. Princeton University Press: Princeton, N.J.

Horowitz, Michael C. (2016). “Public opinion and the politics of the killer robots debate”. en. Research & Politics 3(1), p. 2053168015627183.

Hostbeck, Lars (2015). “Space Weapons’ Concepts and their International Security Implications”. en. In: Handbook of Space Security. Springer, New York, NY, pp. 955–983.

Ikenberry, G. John (2000). After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order After Major Wars. English. Princeton University Press: Princeton.

Jervis, Robert (1976). Perception and Misperception in International Politics. English. New edition with a went from a delayed paperback to simultaneous edition. Princeton University Press: Princeton, New Jersey.

Jervis, Robert (1978). “Cooperation Under the Security Dilemma”. World Politics 30(2), pp. 167–214. Jervis, Robert (1993). “Arms Control, Stability, and Causes of War”. Political Science Quarterly 108(2), pp. 239–253.

Kramer, Franklin, Stuart H. Starr, and Larry Wentz, eds. (2009). Cyberpower and National Security. English. 1 edition. Potomac Books: Washington, D.C.

Keohane, Robert O. (1984). After Hegemony: Cooperation and Discord in the World Political Economy. English. With a New preface by the author edition. Princeton University Press: Princeton, N.J.

Keohane, Robert O. (1999). “International Relations and International Law: Interests, Reputation, Institutions”. Proceedings of the Annual Meeting (American Society of International Law), pp. 375–379.

Krisch, Nico (2005). “International Law in Times of Hegemony: Unequal Power and the Shaping of the International Legal Order”. en. European Journal of International Law 16(3), pp. 369–408.

Koremenos, Barbara (2005). “Contracting around International Uncertainty”. American Political Science Review 99(4), pp. 549–565.

Kydd, Andrew (2000). “Arms Races and Arms Control: Modeling the Hawk Perspective”. American Journal of Political Science 44(2), pp. 228–244.

Lanius, Roger (2017). The Establishment of the Outer Space Treaty. en.

Libicki, Martin C. (2009). Cyberdeterrence and Cyberwar. English. RAND Corporation: Santa Monica, CA.

Long, Austin (2016). A Cyber SIOP? Operational Considerations for Strategic Offensive Cyber Planning. en. SSRN Scholarly Paper ID 2836204. Rochester, NY: Social Science Research Network.

Main, Steven J. (2015). “Gas on the Eastern Front During the First World War (1915–1917)”. The Journal of Slavic Military Studies 28(1), pp. 99–132.

Majeski, Stephen J. (2004). “Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations”. International Studies Quarterly 48(2), pp. 455–470.

Maurer, Tim (2011). “Cyber Norm Emergence at the United Nations—An Analysis of the UN’s Activities Regarding Cyber-security”. Science, Technology, and Public Policy Program, Belfer Center.

Mazanec, Brian M. (2015). The Evolution of Cyber War: International Norms for Emerging-Technology Weapons. English. Potomac Books: Lincoln.

Mercer, Jonathan (2010). “Emotional Beliefs”. en. International Organization 64(1), pp. 1–31.

Milgrom, Paul R., Douglass C. North, and Barry R. Weingast* (1990). “The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs”. en. Economics & Politics 2(1), pp. 1–23.

Miller, Steven E. (1984). “Politics over Promise: Domestic Impediments to Arms Control”. International Security 8(4), pp. 67–90.

Mitchell, Billy (1925). Winged Defense: The Development and Possibilities of Modern Air Power–Economic and Military. English. First edition. Fire Ant Books: Tuscaloosa, AL.

Montgomery, Evan Braden (2006). “Breaking out of the Security Dilemma: Realism, Reassurance, and the Problem of Uncertainty”. International Security 31(2), pp. 151–185.

Moravcsik, Andrew (1997). “Taking Preferences Seriously: A Liberal Theory of International Politics”. en. International Organization 51(4), pp. 513–553.

Morgan, Steve (ed.) (2016). Hackerpocalypse: A Cybercrime Revelation

Morrow, James D. (1991). “Electoral and Congressional Incentives and Arms Control”. en. Journal of Conflict Resolution 35(2), pp. 245–265.

“Navy Hits Satellite With Heat-Seeking Missile,” (2008), Space.com.

North, Douglas (1990). Institutions, institutional change and economic performance. Cambridge University Press, Cambridge

Olson, Mancur (1965). The Logic of Collective Action. en. Google-Books-ID: jv8wTarzmsQC. Harvard University Press.

Operation Hardtack I Report (1958), DNA6038F ADA136819, pgs. 259-260, http://www.dtic.mil/dtic/tr/ fulltext/u2/a136819.pdf

Perlroth, Nicole, and David E. Sanger. “Cyberattacks Put Russian Fingers on the Switch at Power Plants, U.S. Says.” The New York Times, October 3, 2018, sec. U.S. https://www.nytimes.com/2018/03/15/us/politics/russia-cyberattacks.html.

Plait, Phil (2012). The 50th anniversary of Starfish Prime: the nuke that shook the world.

Price, Richard (1995). “A Genealogy of the Chemical Weapons Taboo”. International Organization 49(1), pp. 73–103.

Price, Richard (1998). “Reversing the Gun Sights: Transnational Civil Society Targets Land Mines”. en. International Organization 52(3), pp. 613–644.

Portree, David S. F. (1999). Orbital debris : a chronology. en. Tech. rep. United States, National Aeronautics and Space Administration.

Posen, Barry R. (1986). The Sources of Military Doctrine: France, Britain, and Germany Between the World Wars. English. N/A edition. Cornell University Press: Ithaca.

Puffer, Raymond L. (1985). "The Death of a Satellite." Air Force Flight Test Center Moments in History, http://web.archive.org/web/20031218130538/www .edwards.af.mil/moments/docs_html/85-09-13.html

Putnam, Robert D. (1988). "Diplomacy and domestic politics: the logic of two-level games". en. International Organization 42(3), pp. 427–460.

Rutherford, Kenneth R. (2000). "The Evolving Arms Control Agenda: Implications of the Role of NGOs in Banning Antipersonnel Landmines". World Politics 53(1), pp. 74–114.

Romano, James A. Jr. et al. (2007). Chemical Warfare Agents: Chemistry, Pharmacology, Toxicology, and Therapeutics, Second Edition. en. Google-Books-ID: MGcnAIu4vyIC. CRC Press.

Schelling, Thomas C. and Morton H. Halperin (1961). Strategy and arms control. en.

Google-Books-ID: hbwGAAAAMAAJ. Twentieth Century Fund.

Schindler, D. and J. Toman (1988) The Laws of Armed Conflicts, Martinus Nihjoff Publisher, pp.202-204. https://ihl-databases.icrc.org/ihl/INTRO/160?OpenDocument

Schmitt, Michael N. and Liis Vihul (2017). International Cyber Law Politicized: The UN GGE's

Failure to Advance Cyber Norms.

Snyder, Jack (1989). The Ideology of the Offensive: Military Decision Making and the Disasters of 1914. English. First Edition. Cornell University Press: Ithaca.

Stares, Paul B. (1985). The Militarization of Space: U.S. Policy, 1945-1984. English. First Edition. Cornell Univ Pr: Ithaca, N.Y.

Tannenwald, Nina (1999). "The Nuclear Taboo: The United States and the Normative Basis of

Nuclear Non-Use". International Organization 53(3), pp. 433–468.

United States Department of Defense Cyber Strategy, 2015. https://www.defense.gov/Portals/1/features/2015/0415_cyber-strategy/Final_2015_DoD_CYBER_STRATEGY_for_web.pdf

UN General Assembly A/RES/53/70 (1999), https://www.ccdcoe.org/sites/default/files/documents/UN-981204-ITIS.pdf

UN General Assembly A/60/202:2 (2005), https://digitallibrary.un.org/record/555369/files/A_60_202-EN.pdf?version=1

UN General Assembly Resolution 2222 (XXI) (1967), http://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/outerspacetreaty.html

Valeriano, Brandon, and Ryan C. Maness. Cyber War versus Cyber Realities: Cyber Conflict in the International System. 1 edition. Oxford ; New York: Oxford University Press, 2015.

Zanders, Jean Pascal (2003). "International Norms Against Chemical and Biological Warfare: An Ambiguous Legacy". Journal of Conflict and Security Law 8(2), pp. 391–410.

‘No More Free Bugs’: The History, Organization, and Implications of the Market for Software Vulnerabilities

Ryan Ellis[8]

Abstract

The paper examines the creation and organization of the market for software vulnerabilities. In 1995, Netscape launched a then-novel idea: a program that paid users that discovered flaws in the most recent version of their Netscape Navigator web browser. Over the past two decade, “bug bounty” programs, as they are known, have become commonplace: Google, Microsoft, Facebook, and hundreds of other companies now purchase flaws from thousands of individuals across the globe. The paper follows two related lines of inquiry: (1) It charts the invention of the market, tracing how vulnerabilities were transformed into commodities; and (2) it considers the implications of the creation of the market for labor, exploring the legal, technical, and economic conditions that surround the business of identifying and selling newly-discovered vulnerabilities. The work-in-progress papers argues that market emerged from the twin efforts of security researchers to recast or redefine their contributions to the maintenance and repair of platforms as work and, at the same time, it sprang from the efforts of software vendors and online services to blunt what they saw as the risks that accompanied non-market forms of vulnerability disclosure (particularly what is known as “full disclosure”). That is, the market emerged from the intersection of efforts from below (workers moving to exert control over their field activity) with efforts from above (vendors seeking to enclose non-market forms of activity within a predictable logic of commercial exchange). Additionally, the paper examines in detail the labor practices that define the contemporary market for bugs. Reviewing a mix of qualitative and quantitative data, the paper argues that the market for flaws is defined by precarity: the economic, legal, and technical outlines of the market leave labor on unstable ground. As non-market forms of circulation give way to commercial transactions, individuals identifying and disclosing new vulnerabilities remain vital, marginalized, and, above all else, vulnerable.

[8] The author would like to thank Columbia University SIPA and the Carnegie Corporation for their generous research support.

Cliff Stoll discovered something interesting. Stoll, a systems manager at Lawrence Berkeley Lab working in the 1980s, discovered what appeared to be a serious ongoing case of espionage. Recounting his experience in the classic, *the Cuckoo's Egg*, he wondered—in a dilemma that would be common to hackers and security researchers—what should he do when discovering a new flaw? Stoll writes:

> Suppose I find a computer security problem, and its widespread. Perhaps I should keep my mouth shut, and hope that nobody else figures it out. Fat chance. Or Perhaps I should tell the world. Post
>
> a notice on lots of electronic bulletin boards saying, 'Hey you can break into any Unix computer by…' That would at least wake up the people who manage the systems. Maybe even prod them into action. Or should I create a virus, one that takes advantage of this security hole? If there was a trusted clearinghouse, I could report them. They, in turn, could figure out a patch for the problem, and see that the systems are fixed.[9]

How should newly discovered and previously unknown flaws be disclosed? Stoll optimistically hoped that a trusted third party might appear to help coordinate disclosure. He could hardly have foreseen what the coming decades would bring. In the ensuing years, heated debates about vulnerability disclosure would fill message boards, mailing lists, and the pages of publications specializing in computer security. Later, the issue would become a topic of high-politics, international intrigue, and popular conversation. In 2013, Edward Snowden revealed that the National Security Agency had recently spent over $25 million dollars to acquire previously unknown and undisclosed vulnerabilities (what are often called zero-days or 0-days) for espionage.[10] *Time* magazine, hardly a specialized or esoteric outlet, would spotlight questions of vulnerability disclosure on its cover with the headline "World War Zero: How Hackers Fight to Steal Your Secrets."[11] The story, like other contemporaneous popular accounts that followed the Snowden disclosure and the ensuing discussions and debates, highlighted the secret and seemingly lucrative market for vulnerabilities.[12] Nation states, it appeared, were purchasing zero-days as part of an ongoing effort to create what writers would breathlessly if inaccurately describe as new and powerful "cyberweapons."[13] As *Time* writer and novelist Lev Grossman wrote: "Cyberwar isn't

[9] Cliff Stoll. *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage*. New York: Pocket Books, 1990. 291. See also Robert Chesney. "Cybersecurity in 1989: Looking Back at Cliff Stoll's Classic the Cuckoo's Egg." *Lawfare*. Oct. 13, 2015

[10] Brian Fung. "The NSA Hacks Other Countries by Buying Millions of Dollars' Worth of Computer Vulnerabilities." *The Washington Post*. Aug, 31, 2013. Available Online: <https://www.washingtonpost.com/news/the-switch/wp/2013/08/31/the-nsa-hacks-other-countries-by-buying-millions-of-dollars-worth-of-computer-vulnerabilities/>.

[11] Lev Grossman. "World War Zero: How Hackers Fight to Steal Your Secrets." *Time*. July 10, 2014.

[12] For example, see: Bruce Schneier. "Should U.S. Hackers Fix Cybersecurity Holes or Sell Them? *The Atlantic*. May 19, 2014. Available Online: <http://www.theatlantic.com/technology/print/2014/05/should-hackers-fix-cybersecuirty-holes-or-exploit-them/371197/>; Chris Bryant. "Companies Eye Lucrative Zero-Days Market." *Financial Times*. Jan. 14, 2014; Nicole Perlroth and David E. Sanger. "Nation's Buying as Hackers Sell Flaws in Computer Code." *The New York Times*. July 13, 2013.

[13] Ibid.

the future; it's already here. It's business as usual. In this war, the battlefield is everywhere, bugs are weapons…"[14]

The following paper turns away from the popular conversation regarding the market for flaws to focus on the mundane tasks of discovering and selling flaws. It examines not the shadowy and underground market for flaws populated by defense contractors, governments, and criminals, but the routine, widespread, and transparent market for flaws: the purchase of bugs by software vendors from security researchers. In the decades since Stoll penned his musings on the topic, vulnerability disclosure has been transformed into a thriving market. What was once an experiment, then a curiosity adopted by a few companies, is now widespread and common: each year hundreds of companies purchase thousands of flaws from thousands of different individual security researchers.[15] The market is not trivial—it accounts for millions of dollars in transactions annually.[16] A range of companies now run programs to purchase flaws—Google, Facebook, Twitter, Apple, and other companies that dominate the digital economy operate bug bounty programs.[17] But, so do somewhat more surprising players: United Airlines launched its own initiative to purchase flaws in 2015.[18] A year later, the Department of Defense announced its own bug bounty program—dubbed "Hack the Pentagon."[19]

The following paper examines the vulnerability market as a historical and political artifact. It examines two related threads of analysis: First, it considers the invention of the market. Rather than seeing the creation of the market for flaws as inevitable or natural, the paper explores the organizational and ideological shifts and transformations that undergird the creation of the market. This working paper traces how a form of non-market activity—vulnerability disclosure—came to be partially annexed by the market. For years, hackers discovered new flaws in commercial and non-commercial systems and disclosed them through non-market circuits of exchange. Why and how then did this form of activity become commercialized beginning in the mid-1990s? As the sociologists Richard Swedberg and Mark Granovetter observe (following Viviana A. Zelizer), "[t]here is nothing 'natural' about the fact that something has a price; a price, like everything else in the economy, has to be socially constructed."[20] Taking the market as a historical artifact reveals the key role that particular organizations and individuals played in institutionalizing the market for

[14] ibid.

[15] For a partial listing of programs, see: HackerOne, <https://hackerone.com/>; BugCrowd, "Reward Programs," <https://bugcrowd.com/programs/reward>; Bugsheet, "Bug Bounties & Disclosure Programs," <http://bugsheet.com/directory>. Additional details about the scope of the market are described in detail below.

[16] Google alone paid out over $3 million in bounties or rewards in 2016. Eduardo Vela Nava. "Vulnerability Rewards Program: 2016 Year in Review." *Google Security Blog*. Jan. 30, 2017. Available Online: <https://security.googleblog.com/2017/01/vulnerability-rewards-program-2016-year.html>

[17] See: Vela Nava. "Vulnerability Rewards Program: 2016 Year in Review."; Joey Tyson. "Facebook Bug Bounty: $5 Million Paid in Five Years." Oct. 12, 2016. Available Online: <https://www.facebook.com/notes/facebook-bug-bounty/facebook-bug-bounty-5-million-paid-in-5-years/1419385021409053/>; Twitter. "Policy." Available Online: <https://hackerone.com/twitter>.

[18] Greg Kumparak. "United Airlines Will Give You Up to A Million Miles for Finding Security Bugs." *TechCrunch*. May 14, 2015. Available Online: <https://techcrunch.com/2015/05/14/united-airlines-will-give-you-up-to-a-million-miles-for-finding-security-bugs/>.

[19] Lisa Ferdinando. "Carter Announces 'Hack the Pentagon' Program Results." U.S. Department of Defense. June 17, 2016. Available Online: <https://www.defense.gov/News/Article/Article/802828/carter-announces-hack-the-pentagon-program-results/>.

[20] Richard Swedberg and Mark Granovetter. "Introduction." *The Sociology of Economic Life.* San Francisco: Westview Press, 1992. 21.

flaws and it emphasizes the contingent nature of the market: it emerged not due to an invisible or inevitable historical force, but due to the congealing of mundane concerns and squabbles, parochial interests, and, eventually, the entrepreneurial zeal of key participants that proselytized for the market. Second, the paper turns to explore the labor dynamics of the market. Here, the paper considers: What has commercialization meant for labor? How is power organized within the market? Key researchers hoped that the creation of the market for flaws would, in part, blunt the risks and hazards that researchers faced under non-market forms of disclosure. The mantra of "No More Free Bugs"—a plea for the commercialization of vulnerability disclosure—was wrapped in a desire for protections for researchers. Yet, the work of identifying and selling bugs is and remains precarious. The technical, legal, and economic outlines of the market put labor on an unstable footing. Vendors have significant power to dictate the terms of the market—defining the price for bugs, determining what flaws will and will not qualify as commodities, and determining to some degree the legal protections that will or will not be extended to security research. Despite the headlines trumpeting researchers that sell flaws for significant sums and tout security research as a lucrative filed, the reality is the market for flaws is significantly stratified. A small core set of researchers find and sell a large number of flaws at comparatively high average and aggregate prices, while a vast pool of researchers are infrequent sellers trading their commodities for significantly lower prices. Ultimately, the work of selling flaws, like other forms of digital labor, is defined by precarity: it is at once vital and insecure.

The Invention of the Market: Creating a Valuable Flaw

The market for flaws grew out of the twin efforts of software vendors and online services to control what they saw as the risks that accompanied non-market forms of vulnerability disclosure (particularly what is known as "full disclosure") and, at the same time, it sprang from the efforts of security researchers to recast or redefine their contributions to the maintenance and repair of platforms as work. That is, the market emerged from the intersection of efforts from above (vendors seeking to enclose non-market forms of activity within a predictable logic of commercial exchange) with efforts from below (workers moving to exert control over their field activity). It was these complementary organizational and ideological shifts that initially supported the broad adoption of the market as a model for managing vulnerability disclosure. More recently, key firms—namely the companies Bugcrowd and HackerOne—helped actively diffuse and institutionalize the market.

The creation of the market proceeded in a tentative fashion. At first, commercialization was little more than a public relations stunt—a way for an embarrassed company (Netscape) to spin negative headlines at a moment when bad press was particularly worrying. In 1995, Netscape launched the first high-profile bug bounty program.[21] At the time, Netscape was a run-away success with both users and investors. In August, Netscape went public and stunned investors. The company doubled its initial public offering share price on its first day of trading, opening at $28, spiking at $75

[21] Joan E. Rigdon. "Netscape is Putting a Price on the Head of Any Big Bug Found in Web Browser." *The Wall Street Journal.* Oct. 11, 1995; Dow Jones & Company. "Netscape Unveils 'Bounty' Program for Navigator 2.0." *Dow Jones News Service.* Oct. 10, 1995.

during the day, and closing at $58.25 by the end of the day.[22] The financial press produced reams of praise the next day marveling at the year-and-a-half old company's success.[23] Netscape's web-browser, Netscape Navigator 1.0, was dominate: It accounted for an estimated 70-80% of the browser market.[24] In the weeks after Netscape's IPO, however, a string of negative press reports surfaced: independent researchers—hackers—discovered a series of significant flaws in Navigator.[25] In August, a French student at Ecole Polytechnique discovered a flaw in Navigator—a way of cracking the encryption scheme used by the browser—that made headlines just days after Netscape's public offering.[26] Then, in September, two graduate students at the University of California, Berkeley—David Wagner and Ian Goldberg—discovered a flaw in Navigator that undermined the security settings of the browser.[27] They posted their findings publicly on the mailing list alt.cypherpunks and the news quickly spread to the front page of the *New York Times.*[28] Wagner was surprised about the attention that their discovery generated, remarking that "It was just kind of a preliminary heads-up we were giving to the cypherpunks people about a project we were in the middle of and still working on."[29] Goldberg was not generous to Netscape in his public comments, offering a tart observation that perhaps Netscape's security features were being oversold and lulling users into a false sense of confidence.[30] He noted that Navigator's security was "not as good as people thought, which is probably worse than no security."[31] Netscape pledged to fix the flaw as soon as possible. But, Netscape was back in the news days later. Another flaw had been found and reported via the same mailing list. The *Wall Street Journal* reported that the bug yet again called into question the integrity of Netscape's browser and linking the discovery to larger concerns about privacy and security on the Internet.[32] Press reports about the series of flaws called into question the security of Navigator and, by extension, punctured some of the hype surrounding Netscape. The reports wondered openly if the series of flaws would undermine the growth of commerce on the web and they wondered what the disclosures might mean for the future of Netscape.

The negative press stories came at a crucial moment for Netscape. Netscape would report first

[22] Scott Reeves. "Netscape's IPO Sends Its Stock into Orbit and Stuns the Market." *Dow Jones News Service.* August 10, 1995; Molly Baker. "Technology Investors Fall Head Over Heels for Their New Love—Little Stock Called Netscape is Lofted to the Heavens in a Frenzy of Trading." *Wall Street Journal.* August, 10, 1995.

[23] Ibid.

[24] David A. Kaplan. "Nothing by Net." *Newsweek.* Dec. 25, 1995; Jared Sandberg. "Netscape Acknowledges Encryption Flaw." *Wall Street Journal.* Sept. 20, 1995; Jared Sandberg. "Sun and Netscape are Forming Alliance Against Microsoft on Internet Standard." *The Wall Street Journal.* Dec. 4, 1995.

[25] See: Susan Moran. "Netscape Security Flaw Bodes Ill for Commerce." *Reuters News.* Sept. 19, 1995; Aaron Zitner. "Netscape Flaw Seen Setback for Business." *Boston Globe.* Sept. 20, 1995; Kevin Maney and Robyn Meredith. "Risky Business on the Internet: Few Feel Safe Making On-Line Transactions." *USA Today.* Sept. 20, 1995. Jared Sandberg. "Netscape Offers Reassurances on Data Safety." *The Wall Street Journal.* Sept. 20, 1995; Jared Sandberg. "Netscape's Internet Software Contains Flaw that Jeopardizes Security of Data." *Wall Street Journal.* Sept. 19, 1995.

[26] Mark Tran. "Hacker Takes the Gloss off Netscape's Floatation Success." Aug. 18 1995; Sandberg. "Netscape's Internet Software Contains Flaw that Jeopardizes Security of Data.".

[27] Zitner. "Netscape Flaw Seen Setback for Business."

[28] Zitner. "Netscape Flaw Seen Setback for Business."

[29] Qtd. in Zitner. "Netscape Flaw Seen Setback for Business."

[30] Sandberg. "Netscape's Internet Software Contains Flaw that Jeopardizes Security of Data.".

[31] Qtd. in Sandberg. "Netscape's Internet Software Contains Flaw that Jeopardizes Security of Data."

[32] Jared Sandberg. "Netscape Software for Cursing the Internet is Found to Have Another Flaw." *The Wall Street Journal.* Sept. 25, 1995.

quarter earnings in late-October, a highly-anticipated moment for the newly-public company that had not yet shown a profit (and that left many investors puzzled as to what, exactly the company did).[33] Additionally, Netscape was about to launch a new version its browser in a few days—Netscape Navigator 2.0.[34] This was a significant release. Navigator 2.0 was a high-profile collaboration with Sun Microsystems.[35] The browser would integrate the then new programming language, Java, and was touted as a significant step forward.[36] Netscape faced what at the time was seen as an existential threat from Microsoft. Contemporaneous with reports about flaws in Navigator, Microsoft unveiled its new browser, Microsoft Internet Explorer, bundled with the release of Windows95.[37] The clash between Netscape and Microsoft was bitter and it would eventually form a core part of the U.S. anti-trust case against Microsoft (which charged, in part, that the bundling of Microsoft's web-browser with the operating system thwarted competition). The early fall of 1995, then, was a vital moment for Netscape. The string of press reports charging that the new company's software was in someway defective was a significant problem. Invariably, Navigator 2.0 would also ship with its share of unknown flaws. As Sun CTO Eric Schmidt stated, "we expect people to find bugs."[38]

Netscape tried to put the negative press stories behind it. On Oct. 10th, it announced a novel, experimental, new plan. It launched what it called a "bugs bounty" program.[39] It would buy flaws from researchers that found new flaws in Navigator 2.0 and reported them directly to Netscape. Participants would earn cash or items from Netscape's general store in exchange for vulnerabilities.[40] Mike Homer, Vice President of Marketing, described Netscape's thinking: "By rewarding users for quickly identifying and reporting bugs back to us, this program will encourage an extensive, open review of Netscape Navigator 2.0 and will help us to continue to create products of the highest quality."[41]

The broader context of Netscape's decision to launch the first bug bounty program is important. Netscape was confronting and attempting to control a particular form of non-market vulnerability disclosure. The string of public reports about flaws impacting Navigator left Netscape reeling. They had to quickly—and publicly in the press—respond to reports that their products were defective. Additionally, they had to update Navigator on the fly and push out security upgrades that would fix the identified flaws. Netscape was wrestling with the challenges of public or "full disclosure." By the mid-1990s, a new ideology or rationale for vulnerability disclosure took root—"full disclosure." By this point, a subset of hackers became disenchanted with reporting

[33] Reeves. "Netscape's IPO Sends Its Stock into Orbit and Stuns the Market."; Mathew Ingram. "Netscape Fortunes Still a Mystery." *The Globe and Mail.* Dec. 1, 1995. Reuters. "Netscape Posts its First Profit; Stock Surges." *The New York Times.* Oct. 25, 1995.

[34] Jared Sandberg. "Netscape to Unveil New Set of Software for Electronic-Publishing." *The Wall Street Journal.* Sept. 15, 1995; Moran. "Netscape Security Flaw Bodes Ill for Commerce.";

[35] Michael Putzel. "Netscape Reshapes Web Landscape." *The Boston Globe.* Oct. 20, 1995.

[36] Jared Sndberg. "Sun and Netscape are Forming Alliance Against Microsoft on Internet Standard." *The Wall Street Journal.* Dec. 4, 1995.

[37] Tran. "Hacker Takes the Gloss off Netscape's Floatation Success."; Baker. "Technology Investors Fall Head Over Heels for Their New Love."

[38] Qtd. in Joan E. Rigdon. "Netscape Begins to Pay Bounty for Bugs." *The Asian Wall Street Journal.* Oct. 12, 1995.

[39] Rigdon. "Netscape is Putting a Price on the Head of Any Big Bug Found in Web Browser."; Netscape. "Netscape Announces 'Netscape Bugs Bounty' with Release of Netscape Navigator 2.0 Beta." *PR Newswire.* Oct. 10, 1995.

[40] Ibid.

[41] Netscape. "Netscape Announces 'Netscape Bugs Bounty' with Release of Netscape Navigator 2.0 Beta."

vulnerabilities privately to impacted vendors. The complaints were legion: vendors, it was said, were slow to respond to reports; they rarely gave researchers proper credit for identifying new flaws and helping to fix them; and, at worst vendors retaliated to reports from well-meaning researchers with legal threats. Releasing flaws publicly—rather than directly to the vendor—came to be seen as some as an antidote to these (and other) maladies. In releasing flaws publicly, vendors could not delay patching; credit would not be an issue; and the public would be informed immediately about bugs in the products they relied upon. A number of fora catered to full disclosure. Mailing lists, including "Full Disclosure" and "Bugtraq," served as outlets for researchers to disclose previously unknown and unreported vulnerabilities. Security conferences, including Black Hat and DEF CON, offered platforms for researchers to demonstrate novel flaws before an appreciative audience. Vendors took a dim view of full disclosure: Microsoft described it as "information anarchy" and compared full disclosure to shouting fire in a crowed theater.[42] Reporting bugs publicly without forewarning to the impacted vendor not only was embarrassing, it allowed for exploits and attacks to proliferate in the wild before a patch or update could be released. Netscape's "bugs bounty" program was transparently a way of short-circuiting full disclosure. For rewards of merchandise or, for serious flaws, payouts of $1,000, Netscape tried to pull vulnerability disclosure into the predictable logic of the market. No longer would Netscape learn about a new flaw from a public message board or, worse, a reporter calling from the *New York Times*. Now, it was hoped, the market would provide an incentive to yoke researchers into confidential agreements and discrete transactions.

Netscape's innovation would eventually become adopted—but only in time. Initially, the notion of commodifying bugs was taken-up by security companies that were looking to distinguish themselves from their peers. iDefense launched its Vulnerability Contributor Program in 2002 and TippingPoint launched Zero Day Initiative (ZDI) in 2005.[43] These companies were operating according to a logic that was distinct from Netscape and, later, other vendors. They were not buying up flaws that impacted their systems or products; they were not trying to mitigate the perceived negative effects of full disclosure. On the contrary, they were tying to find an edge that would separate themselves from other security companies offering various forms of anti-virus protection and security services. Here, these companies sought to encourage hackers to provide new vulnerabilities directly to these third-party companies. They would, in turn, sell security services to their clients that were based, in part, on the promise that they could and would protect them from exploits and flaws that no other security vendor knew about. Both iDefense and ZDI would also disclose the flaws they purchased from researchers directly to the impacted companies at some point, but they would first offer an exclusive window of protection to their customers. It was based on this gap—in the moment when ZDI or iDefense knew about a flaw but the impacted vendor and other security companies did not—that ZDI and iDefense sought to market and define their offerings. Though the business rationale was different, the launching of these companies was

[42] Kevin Poulsen. "Microsoft Reveals Anti-Disclosure Plan." *Security Focus*. Nov. 9, 2001. Available Online: <http://www.securityfocus.com/news/281>; Scott Culp. "It's Time to End Information Anarchy." *Tech Net*. Oct. 2001. Archived Copy Available Online: <http://www.angelfire.com/ky/microsfot/timeToEnd.html>.

[43] See: "IDEFENSE PAYING $$$ FOR VULNS." Full Disclosure (Mailing List). Aug 7, 2002. Available Online: <http://seclists.org/fulldisclosure/2002/Aug/168>; Brian Gorenc. "Zero Day Initiative: Life Begins at 10." Trend Micro. July 30, 2015. Available Online: <http://blog.trendmicro.com/zero-day-initiative-life-begins-at-10/>; Michelle Delio. "Bug Finders: Should they be Paid?" *Wired News*. Aug. 9, 2002. Archived Version Available Online: <https://web.archive.org/web/20051110165218/http://wired-vig.wired.com:80/news/culture/0,1284,54450,00.html>.

important for the institutionalization of the market. They followed Netscape and helped reinforce the idea that vulnerabilities were commodities and, by extension, that hunting for bugs was a form of work.

The institutionalization of the market occurred in earnest when other software vendors began to adopt bug bounty programs. Netscape's PR stunt would become a taken-for-granted way of managing vulnerability disclosure and bugs would become commodities. In 2004, the Mozilla Foundation launched a high-profile bug bounty program for their open-source browser, Mozilla Thunderbird, and other software.[44] Mozilla was the direct successor to Netscape; it was spun out with support from American Online's Netscape division (American Online agreed to acquire Netscape in 1998) as an open-source browser that would carry on Netscape's legacy.[45] Mozilla's bounty program offered $500 payouts for bugs.[46] Mozilla's program did not technically require non-disclosure of the vulnerability; but the program terms were structured to keep the information to a tight circle through access controls. Mozilla asked and requested that members of the security group and private mailing list that discussed security vulnerabilities not discuss to non-members of the Mozilla's security group or post the information elsewhere on publicly accessible venues.[47]

The broad-based adoption of bug bounty programs, however, would start in earnest in 2010 with the launch of Google's vulnerability rewards program for Chromium (the program would expand in the following years to cover Google's web properties, Android, and other offerings).[48] Google announced the program as an experimental way to encourage outside researches to contribute to the ongoing development of Chromium.[49] In unveiling the new initiative, project lead Chris Evens signaled that the project was inspired by and modeled on Mozilla's ongoing successful program.[50] The initial prices that Google quoted—$500 as a base price—was also directly modeled on Mozilla's payment structure (and acknowledged as such).[51] Evens made it clear that this program was different from the periodic contests that Google and others had held to reveal new vulnerabilities and exploits, declaring that "[t]his is not a competition, but rather an ongoing reward program."[52] This was not a contest, it was designed to be an ongoing, transparent, and durable market. The terms also made it clear that payouts were conditional on discretion. The announcement included a set of questions and answers in order to define the guidelines of the program. One exchange was telling:

[44] Mozilla Foundation. "Mozilla Foundation Announces Security Bug Bounty Program." *Mozilla Press Center.* Aug. 2, 2004. Available Online: <https://blog.mozilla.org/press/2004/08/mozilla-foundation-announces-security-bug-bounty-program/>.

[45] Craig Bicknell. "Mozilla Stomps Ahead Under AOL." *Wired.* Nov. 24, 1998. Available Online: <https://www.wired.com/1998/11/mozilla-stomps-ahead-under-aol/>.

[46] Mozilla Foundation. "Mozilla Foundation Announces Security Bug Bounty Program."

[47] Mozilla Foundation. "Handling Mozilla Security Bugs." Version 1.1. Available Online: <https://www.mozilla.org/en-US/about/governance/policies/security-group/bugs/>; Mozilla Foundation. "Mozilla Security Bug Bounty FAQ." Archived FAQ. Available Online: <https://www-archive.mozilla.org/security/bug-bounty-faq.html#already-published>.

[48] Chris Evans. "Encouraging More Chromium Security Research." *Chromium Blog.* Jan. 28, 2010. Available Online: <https://blog.chromium.org/2010/01/encouraging-more-chromium-security.html>.

[49] ibid.

[50] ibid.

[51] ibid.

[52] ibid.

> Q) Will bugs disclosed publicly without giving Chromium developers an opportunity to fix them first still qualify?
>
> A) We encourage responsible disclosure. Note that we believe responsible disclosure is a two-way street; it's our job to fix serious bugs within a reasonable time frame.[53]

Google, like Mozilla and Netscape, were attempting to use the market to make the reporting of bugs more predictable and discreet (and to encourage more eyes to focus on possible flaws). Responsible disclosure or coordinated disclosure as it is often described requires submitting a vulnerability to the vendor and providing them with an exclusive window (60-days, for example) to fix the bug. After the specified date passes, the researcher is free to disclose the bug publicly, regardless if the vendor has patched the bug or not. Roughly ten months into the Chromium experiment, Google announced that the experiment was a success.[54] The program had pulled in high-quality reports. As a result, Google announced that it would expand its rewards program to include its web properties, including the domains google.com and youtube.com.[55] Now, the terms were explicit: bugs that were publicly disclosed or even shared in private domains outside of the confines of the program would not qualify for a payment.[56]

The following year, Facebook followed Google's lead, announcing its own bug bounty program. Like Google, Facebook explicitly followed Mozilla's lead—offering $500 initially for reported bugs as a base. Additionally, only bugs reported to Facebook privately would be considered for payment.[57] The Internet publication *The Register* greeted the news with a pointed question, subtitling their write-up of Facebook's new plan with the subhead: "Microsoft, Oracle, you listening?"[58] In 2013, Microsoft, often described as a "holdout" in paying for bugs announced their own bounty program.[59]

There is a significant gap between Mozilla's announcement in 2004 and Google's subsequent adoption in six years later. Why, then, did Google and then other companies eventually come to

[53] ibid.

[54] Chris Evans, Neel Mehta, Adam Mein, Matt Moore, and Michal Zalewski. "Rewarding Web Application Security Research." *Google Security Blog.* November 1, 2010. Available Online: <https://security.googleblog.com/2010/11/rewarding-web-application-security.html>

[55] ibid.

[56] ibid.

[57] Elinor Mills. "Facebook Launches Bug Bounty Program." *CNET.* July 29, 2011. Available Online: <https://www.cnet.com/news/facebook-launches-bug-bounty-program/>; Dan Goodin. "Facebook Dangles Cash Rewards for Bug Rewards." *The Register.* July 29, 2011. Available Online: <https://www.theregister.co.uk/2011/07/29/facebook_bug_bounties/>. See: Facebook. "Security Bug Bounty." Internet Archive. Archive date: Aug. 8, 2012. Available Online: <https://web.archive.org/web/20120808072456/http://www.facebook.com/whitehat/bounty.

[58] Goodin. "Facebook Dangles Cash Rewards for Bug Rewards."

[59] Andy Greenberg. "Microsoft Finally Offers to Pay Hackers for Security Bugs with $100,000 Bounty." *Forbes.* June 19, 2013. Available Online: <http:// www.forbes.com/sites/andygreenberg/2013/06/19/microsoft-finally-offers-to-pay-hackers-for-security-bugs-with-100000-bounty/>; Brian Krebs. "Microsoft to Offer Standing Bug Bounty Program." *Krebs On Security.* June 19, 2013. Available Online: <https://krebsonsecurity.com/2013/06/microsoft-to-offer-standing-bug-bounty/>; Matt Miller and David Ross. "New Bounty Program Details." *Microsoft TechNet.* June 19, 2013. Available Online: <https://blogs.technet.microsoft.com/srd/2013/06/19/new-bounty-program-details/>.

adopt bug bounty programs? In other words, how and why did the market for vulnerability spread? Netscape, Mozilla, ZDI, and iDefense legitimized the notion of treating bugs as commodities and full disclosure remained a thorn in the side of vendors: it was seen as a perennial and recurrent problem. Other models or options for managing or mitigating the perceived harms of full disclosure were, however, readily available (and were deployed): punitive legal pressure, enhanced software development, and the socialization of responsible disclosure through professional organizations each offered the possibility of blunting the challenges of full disclosure. Indeed, in the early-2000s, Microsoft would lead an industry coalition that tired to develop a responsible disclosure standard to counter full disclosure (it stalled).[60] Google and then Facebook's adoption and public announcement of their bug bounty programs further legitimized the market. These were large, publicly visible, companies that launched their programs to significant fanfare. In just three short years, in the time between Google's initial experimental announcement and Microsoft's launching of a bounty program, the market for flaws was understood in starkly different terms. When Google launched its program it was still something of a novelty; when Microsoft launched their program just a few years later it was viewed in the press and elsewhere as a long-overdue and inevitable move. *Finally*, reports noted, Microsoft was launching a bounty program. This shift in reception or attitude—from viewing bounty programs as an interesting lark to something approximating a duty demonstrates in part the power that large companies have to legitimate and help diffuse institutional models. Once Google and Facebook signed on, bounty programs stopped being interesting adjuncts to standard security practices and started being seen as expected. Large firms can help spark institutional isomorphism—diffusion of similar ways of doing things—through coercive power, relations with partners and professionalization, and, most importantly, through mimeses: organizations look to other successful similar (or aspirational) organizations as a model for behavior.[61] Despite the relative paucity of programs at the time, once Google and other key player backed the institutional model—the market—others joined suit.

But, other changes—shifts in how hackers or researchers understood their activities—also pushed companies to implement more formal bounty programs. While pressure from above helped create markets, pressure from below, from those that sought out, discovered, and reported flaws, also helped to create the market. Researchers began to demand the adoption of the market. In 2009, two well-known computer security researchers and hackers—Dino Dai Zovi and Alex Sotirov—appeared on stage at CanSecWest holding a crude sign.[62] Scrawled across the improvised cardboard placard in block letters was a new mantra: "no more free bugs."[63] It was a radical protest. CanSecWest is an annual conference that specializes in applied security. A recurring highlight of the conference is the public presentation of novel exploits and newly discovered vulnerabilities that impact operating systems, web browsers, other software, and mobile devices.[64] The protest was hatched by Dai Zovi, Sotirov, and their colleague Charlie Miller (who joined them onstage).

[60] Kevin Poulsen. "Microsoft Reveals Anti-Disclosure Plan." *Security Focus.* Nov. 9, 2001. Available Online: <http://www.securityfocus.com/news/281>.

[61] See: Paul J. DiMaggio and Walter W. Powell. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *The New Institutionalism in Organizational Analysis.* Walter W. Powell and Pau J. DiMaggio, eds. Chicago: University of Chicago Press, 1991. 63-82.

[62] Dennis Fisher. "No More Free Bugs for Software Vendors." *Threat Post.* March 23, 2009. Available Online: <https://threatpost.com/no-more-free-bugs-software-vendors-032309/72484/>. Security Focus. "No More Bugs for Free, Researchers Say." *Security Focus.* March 24, 2009. Available Online: <http://www.securityfocus.com/brief/933>.

[63] Fisher. "No More Free Bugs for Software Vendors."

[64] See, CanSecWest. "History." Available Online: <https://www.cansecwest.com/csw09archive.html>.

They were fed up. For decades, researchers had been identifying and disclosing previously unknown flaws in commercial software and hardware through non-market forms of exchange. As Dai Zovi noted at the time, "reporting vulnerabilities for free without any legal agreements in place is risky volunteer work."[65] Miller, Dai Zovi, and Sotirov were agitating for change: from now on they would no longer give away their discoveries for free. In their view, vendors had been "freeloading" off of security research for too long or, worse, using legal threats to silence security research.[66] If vendors wanted access to their bugs, they were going to have to pay. This was an important ideological shift. In 2001, the security researcher operating under the name of "Rain Forest Puppy" (RFP) outlined a full disclosure policy that sought to manage the interaction between vendors and hackers (it called for vendors to respond within five days to a reported bug or face public disclosure).[67] The policy was drafted with input from, among others, Chris Wysopal from @Stake.[68] The policy is an artifact of the disclosure debates of the 1990s and early-2000s. It views paying for bugs as anathema to security research. It noted that "[m]onetary compensation…is highly discouraged."[69] Dai Zovi, Sotirov, and Miller broke with this notion. Now, hackers understood their activities as work. Miller, Dai Zovi, and Sotirov viewed hunting for bugs as labor with value that should be compensated and they treated bugs as commodities. They hoped that the commercialization of vulnerability disclosure would provide a measure of security—that it would provide legal stability recognition, and fair compensation for important work. While Google and Facebook may have initially modeled their offerings on Mozilla and Netscape's earlier efforts, researchers also helped to encourage isomorphism: they began to demand that companies start offering bounties.

The final piece that accounts for the spread of the market for bugs is the rise of entrepreneurial platforms that worked to aid the start-up of new bounty programs. Two companies in particular HackerOne and BugCrowd, actively worked to facilitate the diffusion of bounty programs.[70] These companies specialize in the management and operation of bug bounty programs on behalf of other companies: they publicly promoted the benefits of bug bounty programs, actively sought new clients to launch new bounty programs, and, importantly, they offered they tools, experience, and platform to launch programs on behalf of interested companies and organizations. The companies were evangelists for commercialization and they drew from some key industry players. HackerOne was founded and led by group of Dutch hackers and featured key figures that had helped design, launch, and run bounty programs at Google, Facebook, and Microsoft respectively: Facebook's former head of product security, Alex Rice was a company co-founder and CTO; Katie Moussouris was the architect of Microsoft's bug bounty programs and served as HackerOne's Chief Policy Officer during its launch; Chris Evans, who started and ran Google's vulnerability rewards

[65] Dino Dai Zovi. "No More Free Bugs." *Trail of Bits.* March 22, 2009. Archived Page Available Online: <https://web.archive.org/web/20130207203323/http://blog.trailofbits.com:80/2009/03/22/no-more-free-bugs/>.
[66] Ibid.
[67] Rain Forest Puppy. "Full Disclosure Policy (RFPolicy) v2.0." Archived Version. Accessed Oct. 23, 2001. Available Online: <https://web.archive.org/web/20011023233527/http://www.wiretrip.net:80/rfp/policy.html>; Kim Zetter. "Three Minutes with Rain Forest Puppy." *PC World.* Sept. 28, 2001. Available Online: <https://web.archive.org/web/20120105001011/http://www.pcworld.com/article/63944/three_minutes_with_rain_forest_puppy.html>.
[68] Rain Forest Puppy. "Full Disclosure Policy (RFPolicy) v2.0."
[69] ibid.
[70] For example, see: HackerOne. "The Hacker-Powered Security Report 2017." 2017. Available Online: <https://www.hackerone.com/resources/hacker-powered-security-report>; BugCrowd. *2017 State of Bug Bounty Report.* 2017. Available Online: <https://www.bugcrowd.com/resource/2017-state-of-bug-bounty/>.

program, was a company advisor.[71] The company was backed with venture capital and had a straightforward mission: to aid companies in starting and operating bug bounty programs.[72] By 2017, they were supporting bounty programs from several hundred companies and organizations, including Twitter, Uber, Adobe, Yahoo!, the Department of Defense, Starbucks, and many more.[73] The diverse slate of participating organizations reflects both the diffusion of the market for bugs and the key role that HackerOne plays in facilitating the spread of the market across different organizational sectors. BugCrowd, started in 2012, operated with a mostly similar business model, running programs for Master Card, Sophos, Fiat Chrysler, and hundreds of others.[74] These companies worked to make the market-model easy to use both for hackers, giving them centralized platforms to submit reports, and companies, providing the infrastructure that undergirds bounty programs.

The market for vulnerabilities started as a PR stunt and gradually became a taken-for-granted way of managing vulnerabilities. The invention of the market was simply that: an invention. There was nothing inevitable about the rise of the market. Indeed, during the mid-1990s, it certainly did not appear that a thriving market for bugs was on the horizon. Netscape pioneered a model and ideology that did not appear to be taken up widely by others. But in time, key companies—Google, Facebook, and Microsoft—would adopt bug bounty programs and legitimize the market for flaws. Researchers, too, played a central role in creating the market. They redefined their efforts as labor, rather than a form of community service or hobby, and pushed companies to adopt bounty programs. Platforms—drawing talent from key early bounty programs—further helped diffuse the market model by actively served as apostles for bounties and offering ready-made tools to launch and operate programs across different organizational sectors. The market was created by a set of reinforcing organizational and ideological shifts: companies turned to bounties as a way of combating full disclosure and profiting off of the work of independent researchers; researchers turned to the market as a way of mitigating the insecurities and vagaries of non-market disclosure. Now that the market is firmly institutionalized, it can be difficult to see how radical of an invention it was. But it was something of revolution: bugs were transformed into commodities, hackers into independent contractors, and commercial transactions into a routine way of managing vulnerability disclosure. It was a future that would have been difficult to see in 1995 when Netscape launched its "Bugs Bounty" program as a way of combating bad press in advance of introducing Navigator 2.0. In his 2014 keynote address at Black Hat, Dan Geer observed:

> Vulnerability finding is a job. It has been a job for something like eight years now, give or take. For a good long while, you could do vulnerability finding as a hobby and get paid in bragging rights, but finding vulnerabilities got to be too hard to do as a hobby in your spare time—you needed to work it like a job and get paid like a job.[75]

[71] Nicole Perlroth. "HackerOne Connects Hackers with Companies, and Hopes for a Win-Win." June 7, 2015. *The New York Times.* June 7, 2015.

[72] Perlroth. "HackerOne Connects Hackers with Companies, and Hopes for a Win-Win."; HackerOne. "The Hacker-Powered Security Report 2017."

[73] HackerOne. "The Hacker-Powered Security Report 2017."

[74] See: BugCrowd. "Programs." Available Online: <https://bugcrowd.com/programs>.

[75] Dan Geer. "Cybersecurity as Realpolitik." Black Hat. Keynote. Delivery Draft. Aug. 6, 2014. Available Online: <http://geer.tinho.net/geer.blackhat.6viii14.txt>.

Geer's comment is fitting: vulnerability finding became a job.

Precarious Work: Technical, Legal, and Economic Insecurity

Researchers hoped that the creation of a market for vulnerabilities would stabilize and secure the work of identifying and disclosing bugs. They hoped that by moving into the market, the recurring challenges that hackers faced—a lack of recognition, legal threats, an absence of compensation—would be mitigated. A closer look at the conditions of labor within the market, however, tells a more complicated story. The technical, legal, and economic outlines of the market make discovering and selling bugs precarious work. Headlines in the popular press routinely spotlight the high-prices that vulnerabilities fetch on the market and report on the large amounts that companies are paying out for flaws.[76] For most researchers selling flaws, however, the reality of the market is different. Technical and legal uncertainty makes the work insecure. Additionally, the market is significantly stratified: A small number of researchers sell a large number of flaws at high prices, well the larger majority of hackers are infrequent, low-paid, sellers. As a result, firms wield an enormous amount of power in dictating the terms of the market, leaving sellers in a disadvantageous position. Here, as elsewhere, digital labor sits on a precarious foundation.

A Perishable Good: The Technical Foundations of the Market

Vulnerabilities are perishable goods.[77] They, like fruit and produce, risk spoiling before they can be sold at market. At their core, the market for flaws is a market for secrets. What is being sold is not just technical details—a flaw in code—but an information asymmetry. Selling bugs is selling excusive information—a secret that, as far the seller and buyer can be sure, nobody else knows. Vendors buy flaws in order to fix them before they can be exploited by others. The terms of bug bounty programs, in nearly all cases, make it clear that they are purchasing information that has not been previously disclosed. A market for secrets is based on exclusivity. Once knowledge about a particular flaw, a particular secret, is publicly available, the price of the commodity shrinks to (near) zero.

Flaws, however, are open to rediscovery. There is nothing to prevent a researcher from finding and disclosing a flaw that another researcher hoped or was preparing to sell. Independent rediscovery is possible. Pinning down an exact rate of likely rediscovery is quite difficult. The degree of likelihood varies due to a number of factors, including the pool of researchers examining

[76] For example: Russell Brandom. "Apple is Launching an Invitation-Only Bug Bounty Program." *The Verge.* Aug. 4, 2016. Available Online: <https://www.theverge.com/2016/8/4/12380036/apple-bug-bounty-program-vulnerability-security>; Liam Tung. "Android Bugs Made Up 10 Percent of Google's $2m Bounty Payouts—In Just Five Months." *ZDNet.* Jan. 29, 2016. Available Online: <http://www.zdnet.com/article/android-bugs-made-up-10-percent-of-googles-2m-bounty-payouts-in-just-five-months/>.
[77] See: Charlie Miller "The Legitimate Vulnerability Market: Inside the Secretive World of 0-day Exploit Sales." Workshop on the Economics of Information Security. 2007. Available Online: <http://www.econinfosec.org/archive/weis2007/papers/29.pdf>.

a piece of software. All academic studies conclude that rediscovery is possible, however.[78] Trey Herr, Bruce Schneier, and Christopher Morris recently estimated that rediscovery rates hover near 12%.[79] Lilly Albon and colleagues at RAND, working with a different data set and a different methodology, discovered a lower figure of rediscovery—roughly 6%.[80] Earlier work by Andy Ozment wound up noting that rediscovery was likely to occur in roughly 8% of cases.[81] But across data sets, it is clear that the possibility of rediscovery is non-trivial. Importantly, hackers *believe* that rediscovery is possible.[82] This creates a ticking-clock for researchers. Being first to market is all that matters. Once a flaw is discovered there is a race to sell it first; an often-told secret is no secret at all.

Additionally, code itself is not static. It changes over time with new releases and updates. Bugs disappear not only due to discovery and mitigation, but also due to the regular churn of code. This, too, puts some pressure on researchers hoping to sell a flaw. If the next update of the browser, website, or other targeted software incidentally does away with the bug, the commodity is spoiled.

Take together, sellers face a constant pressure to sell their wares before they might spoil. This makes the work of researchers insecure. They cannot be sure if their discovery will be valuable in a few days, weeks, or months. Whatever barging power they might have in shopping a sale or haggling over price and terms is undermined in part by the technical features of the market.

Legal Uncertainty: Safe Harbors, Terms of Service, and Risk

The legal context of vulnerability discovery and disclosure also introduces insecurity into the work of finding and selling bugs. Security research often runs afoul of both terms of service (TOS) and end-user license agreements (EULAs). A number of laws make security researcher precarious. Most notably, the Computer Fraud and Abuse Act (CFAA) and the Digital Millennium Copyright Act (DMCA) create legal barriers to conducting security research.[83] The CFAA prohibits "unauthorized access"—a vague legal standard that is difficult to apply consistently or without significant controversy. What exactly counts as unauthorized access is hard to define.[84] This calls into questions standard techniques for testing for vulnerabilities and creates substantial legal uncertainty around the discovery and disclosure of flaws.[85] The DMCA also creates legal uncertainty for researchers. It prohibits the circumvention of certain security protections (technical

[78] For an overview of the existing rediscovery literature, see: Trey Herr, Bruce Schneier, and Christopher Morris. "Taking Stock: Estimating Vulnerability Rediscovery." SSRN. March 7, 2017. Available Online: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2928758>.

[79] Trey Herr, Bruce Schneier, and Christopher Morris. "Taking Stock: Estimating Vulnerability Rediscovery."

[80] Lillian Albon and Andy Bogart, "Zero Days, Thousands of Nights." RAND. 2017. Available Online: <https://www.rand.org/content/dam/rand/pubs/research_reports/RR1700/ RR1751/RAND_RR1751.pdf>.

[81] Andy Ozment. "The Likelihood of Vulnerability Rediscovery and the Social Utility of Vulnerability Hunting." 2005. Available Online: <http://www.infosecon.net/workshop/pdf/10.pdf>.

[82] Miller "The Legitimate Vulnerability Market."

[83] For a general overview, see: Electronic Frontier Foundation (EFF). "A Grey Hat Guide." EFF. Available Online: <https://www.eff.org/pages/grey-hat-guide>.

[84] See: Orin S. Kerr. "Vagueness Challenges to the Computer Fraud and Abuse Act." 94 *Minnesota Law Review* 1561 (2010); GWU Legal Studies Research Paper No. 482; GWU Law School Public Law Research Paper No. 482. Dec. 22, 2009. SSRN. Dec. 22, 2009. Available Online: <https://ssrn.com/abstract=1527187>.

[85] For a cautionary tale, see the case of Weev: Hanni Fakhoury. "Weev's Case Flawed from Beginning to End." EFF. July 3, 2013. Available Online: <https://www.eff.org/deeplinks/2013/07/weevs-case-flawed-beginning-end>.

protection measures) that are designed to prevent the unauthorized use or modification of software.[86] As has been widely-noted, the DMCA's anti-circumvention provision calls into question research efforts.[87] The Library of Congress, however, has carved out a "good faith" exemption from the DMCA for security research. This narrow exception, as Amit Elazari's recent work notes, still provides little comfort for security research: research that is exempt form the DMCA must still comply with the CFAA.[88]

Researchers hoped that the growth of the market would provide a safe-harbor for security research. Yet, as Elazari's work demonstrates, bug bounty program terms do not always live up to this promise.[89] The legal protections outlined in bounty program terms are not standardized across different platforms (different programs offer different language and legal protections), nor are they always harmonized with the general contractual terms (TOS and EULAs) that pair with the bounty program. That is, in certain cases, the bug bounty terms conflict or at the very least remain ambiguous with how they interact with the website or software vendor's larger set of legal policies.[90] Bug bounty programs must explicitly authorize the particular types of access that researchers will need to conduct security research in order to obviate legal jeopardy. In the U.S., the federal government is working to develop standard legal language that be adopted by firms or other organizations regarding vulnerability disclosure.[91]

The broader legal context puts firms in a prime position to dictate the boundaries of acceptable security research and leaves researchers in an at times uncertain legal position. The CFAA and the DMCA's broad scope generally is a barrier to security research (even with narrow exemptions under the DMCA). Firms have an enormous amount of power to define the scope of legal protection: the contractual terms that they define under TOS and EULAs carry significant legal weight. The legal language that accompanies bounty programs can thwart or possibly chill research: in some cases the language is unintentionally ambiguous—leaving researchers to accept uncertain legal risk in order to find and disclose bugs; in others, it is contradictory—at odds with the tools and techniques that researchers need to use to actually conduct their research. Here, the legal outlines of the market join the technical features in putting labor in a precarious position.

Economic Stratification: Infrequent Sellers, Low Prices, and Single-Buyer Markets

The labor market for bugs is stratified: a small core set of sellers earn a disproportionate share of total payouts, sell the majority of all bugs, garner higher prices-per bug compared to program averages; and sell to a large number of different buyers.[92] For most workers, the market for bugs is defined by low-prices, infrequent sales, and single-buyer markets. A review of public data

[86] *Universal City Studios, Inc. v. Corley*, 273 F.3d 429 (2d Cir. 2001)

[87] See: Stan Adams. "Security Research and the DMCA: The Copyright Office Streamlines the Exemption Process." Center for Democracy and Technology (CDT). Nov. 14, 2017. Available Online: <https://cdt.org/blog/security-research-and-the-dmca-the-copyright-office-streamlines-the-exemption-process/>.

[88] Amit Elazari. "Hacking the Law: The Legal Terms of Bug Bounty Programs Explored." Working Paper. 2017.

[89] Ibid.

[90] ibid.

[91] ibid.

[92] See: Ryan Ellis, Keman Huang, Michael Siegel, Katie Moussouris, and James Houghton. "Fixing a Hole: The Labor Market for Bugs." *New Solutions for Cybersecurity*. Howard Shrobe, David L. Shrier, and Alex Pentland, eds. MIT Press. With. 122-147.

collected from the HackerOne platform and anonymized data shared from Facebook's rewards program (see Table 1) reveal a striking, but preliminary, picture of economic stratification. The HackerOne data was collected from 61 different bounty programs over nearly two years. The data includes 650 unique sellers, over 2,000 individual transactions, and over $1.2 million in sales. The Facebook dataset spans slightly less than four years and includes 725 sellers, over 1,900 transactions, and $3.6 million in sales.

A closer look at the data offers a window into the market. Most sellers are infrequent participants: 78% of HackerOne hackers sell three or fewer bugs (see Chart 1); and in Facebook's program 84% of hackers make three or fewer sales (see Chart 2). Given the lack of frequent or repeated sales, it is unsurprising that most hackers on the HackerOne platform are tethered to one or two programs: 65% of all hackers only sell to one outlet; 89% sell to three or less outlets (see Chart 3). It is not clear if the lack of frequency and lack of mobility is due to the relative novelty of the programs—a disparity that will smooth over time—or if it will remain a persistent feature of the market. A small cadre of hackers make the majority of sales and earn a significant slice of all payouts. The top 5% of all hackers working on HackerOne supported bounty programs make nearly a quarter of all sales and earn over 40% of all money paid out across the programs (see Table 2). The Facebook data is even more striking. Here, the top 5% sell 36% of all bugs and earn over 50% of all payouts (see Table 3). These sellers are high-volume and high-paid. In the HackerOne dataset, the top 5% average roughly 16 sales, $16 thousand in career earnings, and $999 per bug; the average participant makes three sales, earns nearly $2 thousand in sales, and averages $542 per sale. The Facebook data tells the same more or less the same story: the top 5% average 20 sales, career earnings of $50 thousand, and $2,496 per sale; the average seller in Facebook's program makes three sales, earns roughly $5,000, and collect a per-bug average of $1,810.

The market for bugs approaches a monopsony market. Rather than thinking of the market as a mix of many buyers and many sellers, it is better conceived as a series of parallel single-buyer markets. Flaws are not readily salable across programs: each vendor is in the main buying up flaws that only impact their platform or software.[93] These flaws are specific: A hacker cannot sell a flaw in the Department of Defense's website, for example, to United Airlines. This structure, combined, with the technical and legal context of the market puts vendors in a significant positon of power. They can dictate which flaws qualify as commodities; they can determine the scope of legal protections afforded sellers; and they can set prices. A small number of high-volume, mobile sellers (selling across different programs), may have power to push back against terms that are deemed unfavorable, but most sellers find themselves in a market where they face winner-take-all competition, legal uncertainty, and little economic bargaining power.

Precarity and Digital Labor: Tentative Conclusions and Further Study

This in-progress paper offers initial impressions. Additional data, including added historical detail on market formation, qualitative data capturing the perspective of sellers active in the market, and an expanded and larger quantitative data set may lead to significant revision and refinement. Yet,

[93] There are notable exceptions: the Internet Bug Bounty program, run by a consortium of companies, explicitly seeks out bugs in programming languages (such as Python and Ruby) and infrastructure technologies that have cross-platform impacts. See: Internet Bug Bounty. Available Online: <https://internetbugbounty.org/>.

this initial stage of observation lends itself to some tentative insights regarding both the invention of the market and labor dynamics associated with the market. The market is an attempt to displace the perceived negative effects of non-market forms of exchange. Full disclosure presented vendors with recurring challenges and commercialization offered a way of mitigating these challenges by folding hackers into predictable patterns of disclosure. The institutionalization of the market was both an organizational and ideological transformation. Over a period of little more than a decade, flaws were transformed into commodities and hacking into work. This transformation was pushed first by Netscape and a few other organizations that experimented—and legitimized—the purchasing of flaws. Later, Google, Facebook, and Microsoft laid the ground for widespread adoption, both by virtue of their prominence and, importantly, by spinning off key personnel to oversee new firms that would actively cultivate and support bounty programs across hundreds of new businesses. Hackers played a key role in diffusing the model as well—by 2009 a small vocal minority sought to claim vulnerability disclosure as work and called for firms to treat it as such. The market was forged both by efforts from above and below: through the interaction of powerful organizations and key hackers (or, as they would become, workers). Here we see the comingling of organizational behavior and ideology—and the ways in which multiple mechanisms pushed for institutional isomorphism. Additional data will clarify the ways in which these different mechanisms interacted to create the market.

The market for flaws mirrors other forms of digital labor. Workers split into a well-compensated core and a large pool of infrastructural workers that are, in the main, lack security of employment, legal protection, or bargaining power.[94] The invention of the market, to some, promised stability and security. A provisional review of the conditions of labor, though, emphasizes precarity. The technical, legal, and economic outlines of the market mirror other trends in digital work: Labor is vital and insecure. Identifying and disclosing bugs is important: it is essential work that stabilizes and ensures the security of the digital world. But, labor has few protections or assurances. The market is defined by competition between sellers and little competition between buyers. The perishability of the commodity pits hackers against one another, while the economic organization of the market—the dominance of single-market-buyers—ensures that there are limited opportunities to shop discoveries. Firms purchasing bugs have the power to define price and the scope of legal protections. The market, of course, need not be organized in this fashion. It is possible to imagine other forms—companies like ZDI dominating rather than largely being displaced by firms buying up flaws in their own products; legal protections and the extension of safe harbor as a precondition of bounty programs—and as the market model continues to unspool these sorts of issues, questions about equitable treatment, legal protection, and fair compensation, can be revisited. Markets are, indeed, historical and political artifacts. They are, at bottom, social constructions, subject to not only invention by remaking and reinvention as well.

[94] On precarity and digital labor, see: See: Ursula Huws. *Labor in the Global Digital Economy*. New York: Monthly Review Press, 2014; Trebor Scholz (ed.). *Digital Labor: The Internet as Playground and Factory*. Routledge: New York, 2013; Nick Dyer-Whiteford. *Cyber-Proletariat: Global Labor in the Digital Vortex*. London: Pluto Press, 2015; Sara Constance Kingsley, Mary L. Gray and Siddharth Suri. "Accounting for Market Frictions and Power Asymmetries in Online Labor Markets." *Policy and Internet* 7.4 (2015): 383–400.

Tables and Charts

Table 1: Data at a Glance

	HackerOne Dataset	*Facebook Dataset*
Programs	61	1
Sellers	650	725
Total Sales	2,177	1968
Total Payments	$1,180,018	$3,562,684
Average Payment	$542.04	$1,810.31
Dates	11/2013 – 10/2015	6/2011- 4/2015

Table 2: HackerOne Dataset

	Top 1%	*Top 5%*	*Top 10%*	*Top 20%*	*Top 30%*
Number of Sellers	6	32	65	130	195
Number of Sales *(Percentage of Total Sales)*	161 (7.4%)	508 (23.33%)	777 (35.69%)	1181 (54.25%)	1435 (65.92%)
Earnings *(Percentage of Total Payments)*	$190,267 (16.12%)	$507,515 (43.01%)	$695,744 (58.96%)	$907,714.25 (76.92%)	$1,003,954.25 (85.08%)
Average Number of Sales Per Seller *(HackerOne Average: 3.34)*	26.83	15.88	11.95	9.08	7.36
Average Career Earnings Per Seller *(HackerOne Average: $1,815.41))*	$31,711.17	$15,859.84	$10,703.75	$6,982.42	$5,148.48
Average Number of Customers	4.83	4.09	3.68	3.36	3.05
Average Value Per Sale *(HackerOne Average: $542.04)*	$1,181.78	$999.05	$895.42	$768.60	$699.62

Table 3: Facebook Dataset

	Top 1%	*Top 5%*	*Top 10%*	*Top 20%*	*Top 30%*
Number of Sellers	7	36	72	145	217
Number of Sales *(Percentage of Total Sales)*	274 (13.9%)	715 (36.33%)	873 (44.36%)	1158 (58.84%)	1336 (67.89%)
Earnings *(Percentage of Total Payments)*	$899,184 (25.2%)	$1,784,984 (50.10%)	$2,248,384 (63.11%)	$2,731,884 (76.96%)	$3,014,034 (84.6%)
Average Number of Sales Per Seller *(Facebook Average: 2.71)*	39.14	19.86	12.13	7.99	6.15
Average Career Earnings Per Seller *(Facebook Average: $4,914.04)*	$128,455	$49,583	$31,228	$19,104	$13,890
Average Value Per Sale *(Facebook Average: $1,810.31)*	$3,281.69	$2,496.48	$2,575.47	$2,359.14	$2,2256.01

Chart 1: Percentage of Sellers with (N) Sales: Aggregate HackerOne Dataset

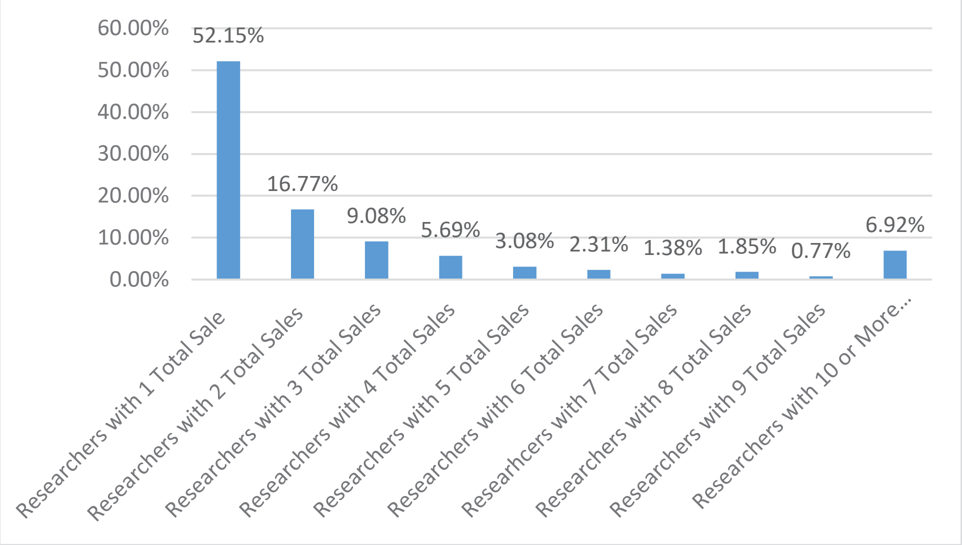


Chart 2: Percentage of Sellers with (N) Sales: Facebook

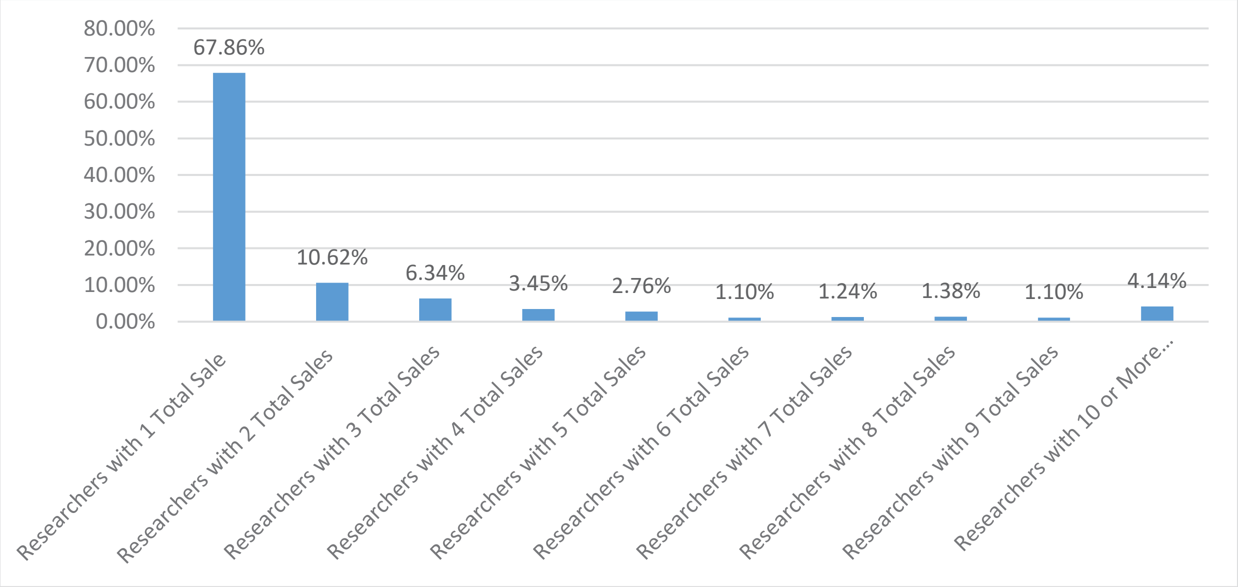
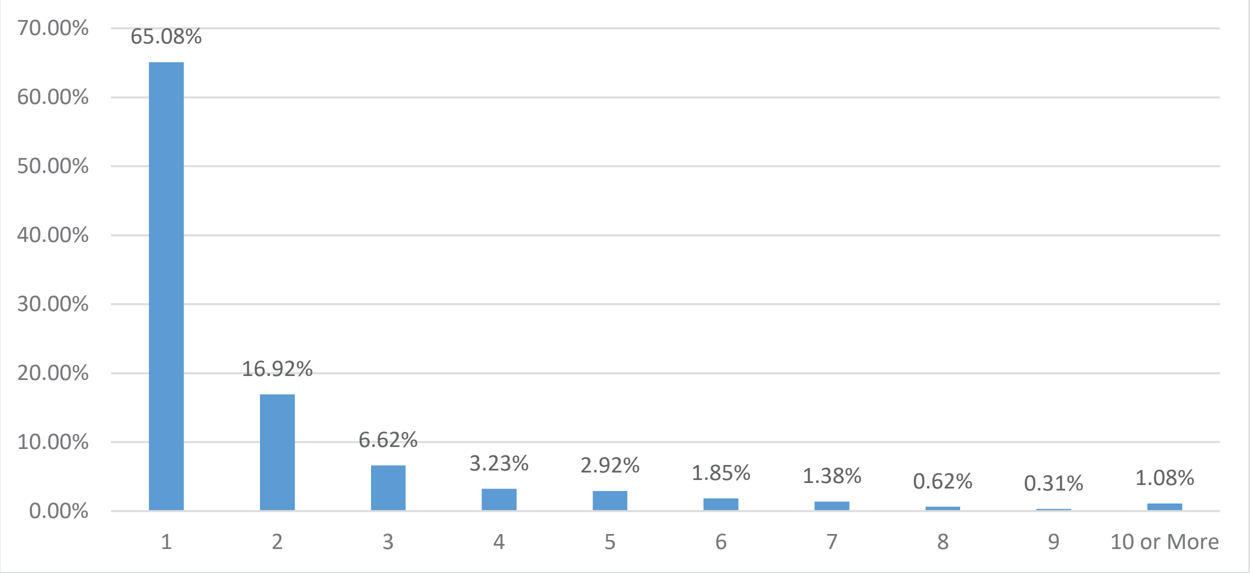


Chart 3: Percentage of Sellers Participating in N Different Programs: HackerOne Dataset



Can Laws Deter Cyber Attacks?[*]

Nadiya Kostyuk[†]

December 1, 2018

PRELIMINARY DRAFT

PLEASE DO NOT DISTRIBUTE

Abstract

Do increases in state legal institutions deter or provoke cyber attacks? Recent research has seen a proliferation of cyber deterrence studies, but it has told us relatively little about the effectiveness of state legal measures adopted with the purpose of punishing online culprits. Moreover, most research on this topic is qualitative and country-specific. This project will address this gap in the literature by quantitatively analyzing original panel data of cyber-crime laws and of large daily distributed denial-of-service (DDoS) attacks across all countries from 2013 to 2016. Using time-series analysis, this article will demonstrate that laws deter DDoS attacks against countries that are able to attribute the origin of such operations and when these countries have a working MLAT with the country in which the perpetrator resides (if the attacks originated abroad). My findings shed some light on the role that legal state cyber capacity plays in deterrence and my findings have important implications for policy-makers.

Keyoword: legal cyber capacity, distributed denial-of-service (DDoS) attacks, deterrence.

[*] This research was possible thanks to the Columbia University's School of International and Public Affairs and Carnegie Research Support.

[†] Doctoral Candidate, University of Michigan, Ann Arbor; nadiya@umich.edu

During the last few decades, distributed denial-of-service attacks[1] have become a policy tool, domestically and internationally. Activists have been using DDoS attacks to signal their protest against governments, as it took place during the popular uprisings against the authoritarian regimes during the Arab spring in 2011 (Coleman, Hacker and Whistleblower, 2014). Hacktivists or patriotic hackers also use their tools to support their government's actions or political stance during the time of conflict or tension. These actors launch cyber attacks on an adversary either without direction from the government – as in the case of the group of teenagers that sabotaged Hezbollah's website in Lebanon (Allen and Demchak, 2003, p. 52), or the American hackers who attacked jihadist websites (Di Justo, 2002) – or with direction – as in the case of the patriotic hackers that executed cyber operations against Estonia and Georgia, in 2007 and 2008, respectively. Governments have also been using DDoS attacks to coerce their opponents. For instance, in 2009, viewed as a Russian attempt to pressure the Kyrgyz president to deny U.S. access to a key airbase, Kyrgyzstan suffered DDoS attacks on two of its primary internet service providers (ISPs).

While these attacks continue to grow in size and scope, countries attempt to address their shortcomings in cybersecurity regulations by updating old legislation or writing new ones. For instance, the United Kingdom's Police and Justice Act of 2006, which amended Section 3 of the country's Computer Misuse Act of 1990, set ten years in prison as the maximum penalty for DDoS attacks.[2] According to §1030 of the U.S. Computer Fraud and Abuse Act, DDoS attacks constitute a federal crime.[3] While these laws are meant to deter and punish future culprits, it remains unknown how effective they, in fact, are and whether *an increase in state legal institutions deters*

[1] DDoS attacks "use multiple computers and Internet connections to flood the targeted resource." This definition is taken from the webopedia website: http://www.webopedia.com/TERM/D/DDoS_attack.html.

[2] For the full text of the law, please visit: http://www.legislation.gov.uk/ukpga/1990/18/section/3

[3] For the full text of the law, please visit: https://www.gpo.gov/fdsys/pkg/USCODE-2010-title18/html/USCODE-2010-title18-partI-chap47-sec1030.htm

or provokes cyber attacks.

The complex relationship between cyber attacks and institutions is understudied in political science, largely because of the challenges of attribution, which limit researchers' abilities to collect data on cyber attacks. Most research on this topic is rather limited, qualitative, and country-specific. A few individual studies have shown that under-developed cyber institutions create vulnerabilities that hackers eagerly exploit – to sell personal information, steal industrial secrets, or even disrupt military operations. In Nigeria, for instance, a lack of proper state cybersecurity laws and policy caused an unprecedented rise in computer-related financial crimes (Moses-Òkè, 2012). Focusing on the United States, Romanosky, Telang and Acquisti 2011 demonstrate the effectiveness of data breach disclosure (security breach notification) laws in reducing identity theft.

This project attempts to address this issue by presenting the research design that uses original panel data of large daily distributed denial-of-service (DDoS) attacks across all countries from 2013 to 2016 and of laws that recognize such attacks as illegal. Using time-series analysis, I expect to demonstrate that laws deter DDoS attacks against countries that are able to attribute the origin of such operations and when these countries have a working MLAT with the country in which the perpetrator resides (if the attacks originated abroad). While the creation of cyber-crime laws and DDoS attacks have been widely observed during the last decade, this is the first study that attempts to explain the deterrent effect of laws and regulations on cyber operations. These findings will have direct implications for policy-makers.

The article proceeds as follows. Before introducing my theory, I discuss different types of cyber operations and explain why I have decided to focus on DDoS attacks (Section 1). I also briefly introduce and define legal cyber capacity in Section 1. Section 2 introduces the argument and explains why laws should deter cyber operations. Section 3 introduces new datasets and Section 4 presents my empirical strategy. I end with a discussion of the expected results, before suggesting avenues for future research (Section 5).

1 Paper Focus

1.1 Cyber Operations

The literature distinguishes four types of cyber incidents based on the objectives they attempt to achieve (Valeriano and Maness, 2015). First is *information operations* or *propaganda*, which seeks to influence public opinion by "trolling" online comments pages and establishing forums and websites to promote certain messages.[4] Second is *espionage* whose main goal is to collect intelligence on an adversary via the online environment. Third is *degradation*, which includes the use of malicious code to inflict damage or compromise infrastructure and military objects. Last is *disruption* – "low-cost, low-pain initiatives that harass a target to influence their decision calculus" (Valeriano and Maness, 2018, 225) – includes efforts to use proxies to block access to websites, among others.

This article focuses on disruption-type attacks since their impact is easy to observe and measure.[5] Specifically, I focus on the distributed-denial-of-service (DDoS) attacks that use multiple computers and internet connections to flood the targeted resources. A DDoS attack is analogous to a group of people crowding the entry door to a shop and not letting legitimate parties enter the shop, disrupting normal operations. As a result of DDoS attacks in 2007, for instance, the Estonian government and business entities were not able to properly operate for about three weeks.

[4] During the last few years, scholars have been intensively studying these efforts and demonstrated that China (King, Pan and Roberts, 2013, 2017) and Russia (Sanovich et al., 2015) have been two leading governments in this regard.

[5] Since the main focus of propaganda and espionage campaigns is the long-term goal of influencing public opinion, which is often hard to measure, these attacks lie beyond the scope of this paper. Similarly, the degradation operations are often quite complex and involve an espionage component on its initial stages.

These attacks are commonly used a decade later as they are easy to execute, readily available to purchase for as low as $70 USD per day (Goncharov, 2012), and can serve as a tool of censorship (Deibert and Rohozinski, 2010; Villeneuve and Crete-Nishihata, 2011; King, Pan and Roberts, 2013; MacKinnon, 2013), contention (Asal et al., 2016), and revenge. So-called *Internet-of-things* devices – home routers, web-cameras, digital video recorders, and other everyday appliances that have internet capabilities built into them – allow the scale of DDoS attacks to grow. These devices, often designed to be as inexpensive and easy-to-use as possible, are frequently exploited as bots.[6]

1.2 Legal Cyber Capacity

Out of the various components of *state cyber capacity*, this project focuses on *legal* measures. A legal framework sets the minimum standards agencies should follow and provides guidance on how to deal with cyber-crime. The binding nature of these laws makes their implementation longer than the establishment of national cybersecurity agendas, for instance, but this implementation should also be at a low cost. Moreover, the legal nature of such documents signals a country's intent to recognize certain actions as illegal and its readiness to punish those who engage in such actions. For instance, Mitchell L. Frost served a 30-month prison sentence for a series of DDoS attacks that he launched against the websites of the former major of New York, Rudy Giuliani, as well as the University of Akron's network, where he was a student. Legal measures also aim to deter future would-be perpetrators.

[6] A *bot* is a software application that runs automated tasks – executes DDoS attacks, for instance – over the internet.

2 Deterrence in Cyberspace

Deterrence is the act of preventing an actor from doing something by making them believe that the cost will exceed the benefit (Schelling, 1966). Classical deterrence theory rests primarily on two main mechanisms: a credible threat of punishment for an action (*deterrence by punishment*); and denial of gains from an action (*deterrence by denial*). Scholars agree that the challenge of attribution in the online environment makes deterrence by punishment difficult (Borghard and Lonergan, 2017; Gartzke, 2013; Libicki, 2009; Lindsay and Gartzke, 2015; Nye Jr, 2017). The usage of proxies provides perpetrators with an endless number of ways to re-route cyber attacks and mask their true identities. In some cases, it is possible to attribute cyber attacks to the actual machine that was used to execute these attacks. Such attribution however does not necessarily point to the attack initiator (e.g., the usage of a public computer). Even though the expert cyber community has recently agreed that the technical evidence "was advanced enough both to trace and attribute attacks" (Soldatov and Borogan, 2017),[7] no country has accepted such standards as its official state policy.

Attribution is also an issue for *deterrence by entanglement* and *deterrence by normative taboos* (Nye Jr, 2017).[8] The former implies that the existing interdependences between two countries make a successful cyber attack costly for both the attacker and the target. The latter is similar to the naming-and-shaming IR norm. *Deterrence by denial* is indifferent to the problem of attribution since its main emphasis is defense. President Barack Obama's administration practiced such deterrence by protecting the objects of the

[7] Specifically, they mentioned, "If an attack could be attributed to a hacking group with a known history of attacking similar targets and this group's attacks consistently worked to benefit one particular country, this constituted enough evidence to determine that the attacks were backed and directed by the state of that beneficiary country" (Soldatov and Borogan, 2017).

[8] Snyder Glenn 1961 consider "broad deterrence," which includes two other political mechanisms: entanglement and norms.

U.S. critical infrastructure. For instance, the 2013 Presidential Policy Directive 21 recommends federal agencies and critical infrastructure owners and operators to work together to minimize cyber risks and strengthen resilience to attacks. Apart from protecting its own systems, indirect ways to increase a country's defense are to sponsor research and development, implement standards in computer network defense, provide incentives for the private sector to protect is own infrastructure, encourage information-sharing between public and private sectors and to subsidize the education of computer security professionals (Libicki, 2009).

Legal Capacity. I view legal regulations as deterrence by punishment because their goal is to create fear of punishment that perpetrators face in the case of violating these regulations. Two conditions should be considered in this scenario – the attack origin and a country's ability to attribute cyber attacks. If the attack originated within Country A's[9] borders and Country A has the capacity to attribute this attack, then we should observe a decline in such attacks once Country A implements the new legislation (*a 'deterrent' case*). If the country lacks the ability to attribute cyber attacks, the origin becomes irrelevant. Since Country A does not know where the attack originated, establishing laws that punish responsible perpetrators will not deter Actor X.[10] Thus Actor X will continue executing attacks at the same (*'no effect'*) or higher rate (e.g., an increase in scope, sophistication and scale of DDoS attacks) (*an 'inflammatory' case*), facing no fear of punishment.

The effect of laws on DDoS attacks can also have both effects – inflammatory and deterrent. It can inflame attacks at intermediate levels but deter them at extremes (*an 'inverted U' case*). This situation is typical when DDoS attacks originate outside of Country A and Country A has the ca-

[9] Country A is a proxy for any country that is being attacked, paying no particular attention to whether the target of such attacks is of a private or public nature.

[10] Actor X is a proxy for any attacker (state or non-state).

pacity to prove it. In this case, we might see a rise in DDoS attacks at first and then a decline in such operations due to an increase in international cooperation between Country A and Country B, where Actor X[11] resides, possibly through Mutual Legal Assistance Treaties (MLAT)[12] or other tools. This cooperation can lead to a decrease in attacks. For instance, after the Mandiant report pointed to the building where China's Unit 61398 resided (Westby, 2013), which was responsible for hacking into the networks of Westinghouse Electric, a U.S. Steel Corporation, the two countries signed an agreement that forbade the theft of intellectual property. Some claim that as a result of this agreement, China's industrial espionage fell from 60 network compromises in February 2013 to less than 10 by May 2016 (Segal, 2016).

Lastly, laws can lead to a decrease in DDoS attacks at intermediate levels and to an increase at extremes (*a 'U-shape' case*). This scenario takes place when Country A has the capacity to attribute domestic cyber operations, causing their decline. But Country A does not have an MLAT treaty with Country B where Actor X resides, or this treaty is not effective, causing a rise in attacks that originate abroad. While it is possible to observe the latter two scenarios – 'U-shape' and 'inverted U' – using the empirical strategy described in Section 4, testing these scenarios would require focusing on international cooperation on cyber crime. All these scenarios demonstrate that *laws serve as deterrent in the countries that are able to attribute the origin of DDoS attacks and have a working MLAT with the country where a perpetrator resides* (if the attacks originated abroad).

- **Hypothesis:** Laws serve as deterrent in the countries that are able to attribute the origin of DDoS attacks and have a working MLAT

[11] For simplicity, I use Actor X – the attacker – to mean any state or non-state actor/s in this scenario.

[12] *A mutual legal assistance treaty (MLAT)* is an agreement between two or more countries for the purpose of gathering and exchanging information in an effort to enforce public or criminal laws. From: https://en.wikipedia.org/wiki/Mutual_legal_assistance_treaty.

with the country where a perpetrator resides (if the attacks originated abroad).

3 Data

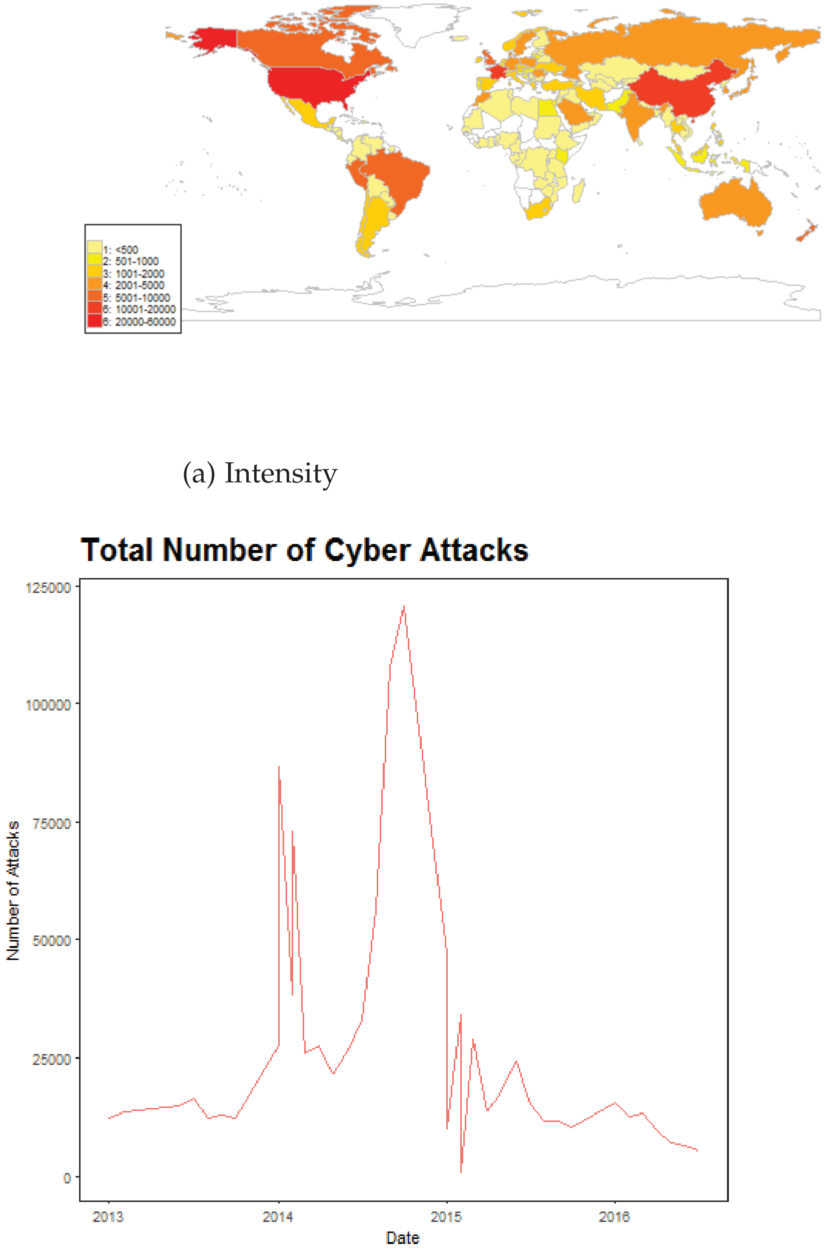
In this section, I present new cross-national time-series data to study the relationship between DDoS attacks and state legal cyber capacity. For my DDoS attacks data, I used Arbor Networks' data from June 1, 2013 to August 22, 2016. I am in the process of compiling a cross-national time-series dataset on laws that recognize DDoS attacks as illegal during this time period. Lastly, I measure a country's ability to attribute the origin by cyber operations by a presence or absence of a state military cyber unit from 2013 to 2016.

3.1 Distributed Denial-of-Service Attacks

To test my hypotheses, I compiled a new global dataset on the destination of daily distributed denial-of-service (DDoS) attacks across 172 countries from June 1, 2013 to August 22, 2016, which contains 225,757 unique events. I focus on DDoS attacks because they are prevalent – as this dataset shows – but have been understudied empirically.[13] Figure 1a displays the intensity of DDoS attacks per country during the studied period and Figure 1b displays the total distribution of DDoS attacks during the studied period. Data from Arbor Networks contain two types of DDoS attacks – the top 2% in size of the reported attacks and attacks that are associated with an unusually high amount of internet traffic per given country. This dataset is unique as Arbor Networks utilizes anonymous attack traffic data between countries and network outage reports.

[13] The only existing dataset in political science on cyber attacks, the *Dyadic Cyber Incident and Dispute Data*, includes only major cyber incidents and disputes (Valeriano and Maness, 2014, 2018). Kostyuk and Zhukov 2017 uses low-level attacks to study cyber coercion during the conflicts in Ukraine and Syria.

Figure 1: DDoS Attacks per Country (June 2013-August 2016)



(a) Intensity

(b) Distribution

3.2 Laws

Since I am interested in testing my main hypothesis as it relates to my theory, I am in the process of compiling a dataset that records which countries have adopted laws that recognize DDoS attacks as illegal. This dataset includes laws and regulations that broadly deal with cybersecurity and cybercrime and more specifically cover unauthorized access, interference, and interception of computers, systems, data protection, breach notification, and certification requirements. In my dataset, I plan to include only those laws that contain specific clauses that address DDoS attacks or could be applied to DDoS attacks.[14] For instance, the Computer Fraud and Abuse Act, or 18 U.S.C. §1030, is the primary federal law that applies to most DDoS-related attacks. The 2001 District of Alaska case *United States v. Dennis* demonstrates such applicability. In this case, a former computer systems administrator in Alaska pled guilty to one misdemeanor count for launching three e-mail based DDoS attacks against a server at the U.S. District Court in New York. He was charged under 18 U.S.C. §1030(a)(5) with "interfering with a government-owned communications system."

My independent variable, *Law*, is a binary variable that captures whether a country has a laws that recognizes the use of DDoS attacks as illegal for a given month. Most of these laws are available in English. To make sure that my list is complete and includes documents in native languages, I use a multi-lingual team of researchers to help me collect these data. I also contacted in-country scholars, known to be working on cybersecurity issues, and asked them to comment on the completeness of my sample. Such a multi-stage approach to data collection stresses the importance of cultural awareness. Often the titles of documents published in a native language might not imply their relevance to *cybersecurity*. In addition, some countries (for instance Russia) use the term *information security* instead of

[14] For instance, extortion against online gambling sites and online business, may fall under 18 U.S.C. §1030(a)(7), which covers extortionate threats.

cybersecurity. I recorded the dates when laws were enacted, amended, and drafted.[15]

3.3 Ability to Attribute

I measure the country's ability to attribute by te presence of a cyber military unit, which requires significant investment in resources and expertise. Attributing cyber attacks is difficult, but it is also easier when a specialized cyber unit is able to coordinate efforts by in-house cybersecurity experts, private tech companies[16] that have technical expertise, and intelligence agencies that have political expertise to identity which adversary might have the incentive to carry out a cyber attack (Yarhi-Milo, 2014). Without the technical expertise of cybersecurity specialists, it would be difficult for the target state to have a complete understanding of the nature and scope of cyber attacks. Without the political expertise of the intelligence community, it would be difficult for the target state to understand the political motives driving the attacks. A dedicated military cyber unit is key to promote intra-governmental cooperation and sometimes private-public partnerships in order to address cyber threats.

My second independent variable, *Attribution*, is a binary variable that captures whether a country has a cyber army or a cyber unit within its military for a given month.[17] Similarly to the legal regulations, I construct my data on military cyber units from technical reports, national cybersecurity

[15] Even though the difference between being "enacted, amended, and drafted" is important, I treat all these actions as a response to DDoS attacks.

[16] These private tech companies include *FireEye, Kaspersky Lab*, among others. They have been used to investigate the Democratic National Convention (DNC) hack that testifies to their capabilities. Recently, the U.S. government has paid private contractors $460 million USD to assist the U.S. Cyber Command in developing and supplying "cyber weapons" in cooperation with U.S. intelligence agencies and in providing technical support to the Cyber Command in planning, organising and coordinating defensive and offensive military activities.

[17] This measure is one of the ways I am thinking about attribution. I am currently in the process of creating other measures that can capture a country's ability to attribute cyber operations.

agenda and defense documents, publications by governmental agencies and inter-govermental organizations, newspaper articles, among others. I also use the approach outlined above to address reporting bias.

4 Empirical Strategy

The adoption of laws that recognize DDoS attacks as illegal can have one of the following effects. First, it can inflame attacks. Second, by raising the costs of attacks and fear of punishment, laws can deter DDoS attacks. Third, it can have both types of effects, inflaming DDoS attacks at intermediate levels but deterring them at extremes – an 'inverted U' – or decreasing DDoS attacks at intermediate levels and increasing at extremes – a 'U-shape'. Lastly, it can have no effect.

To decide which of these effects are working, I plan to fit the following model:

$$\begin{split}\text{DDoS}_{i,t} = \beta_1 \text{Law}_{i,t-1} + \beta_2 \text{Law}^2_{i,t-1} + \beta_3 \text{Attribution}_{i,t-1} + \\ \beta_4 \text{MLAT}_{i,t-1} + X_{i,t} + \gamma_t + u_{i,t} \quad (1)\end{split}$$

where DDoS_{it} is the number of DDoS attacks a given country i suffered in a given month t, Law_{it-1} a binary variable that states whether a country i had a law that recognizes DDoS attacks as illegal in a previous month $(t-1)$. $Attribution_{i,t-1}$ is a dummy variable whether a country had a military cyber unit in a previous month $(t-1)$. $MLAT_{i,t-1}$ is a binary variable that states whether a country had a MLAT treaty with the country where DDoS attacks originated in a previous month $(t-1)$. I plan to use time fixed effects γ_t, representing common shocks over time. $X_{i,t}$ is a vector of plausible confounders that I plan to control for. First, I plan control for GDP per capita and take a natural logarithm of it to account for data skewness. This is to account for the possibility that richer countries should attract more DDoS attacks because hackers can extort more from individuals and companies from those countries. Second, I control for the number of internet users as a percentage of each country's population. A country with a high proportion of internet users could attract more DDoS attacks because there are more available targets for such attacks.[18] To make estimates maximally comparable, I report standardized coefficients (i.e. impact of a standard deviation increase in laws on standard deviation changes in DDoS attacks).

I am interested in how the β_1 and β_2 coefficients vary across countries. The relationship between DDoS attacks and laws is strictly inflammatory if $\beta_1 > 0, \beta_2 \geq 0$, with increases in laws followed by linear ($\beta_2 = 0$) or exponential ($\beta_2 > 0$) increases in DDoS attacks. If $\beta_1 < 0, \beta_2 \leq 0$, the relationship is negative ('deterrence'), with DDoS attacks declining after adoption of the law. If $\beta_1 > 0, \beta_2 < 0$, the relationship is 'inverted-U'-shaped, where increases in laws are correlated with increases in DDoS attacks, but the rate of increase gradually declines and reverses post law adoption. Finally, $\beta_1 < 0, \beta_2 > 0$ indicates the opposite, 'U-shaped' relationship.

In addition, I plan to pay attention to how the β_3 and β_4 coefficients vary across countries and to include interaction effects between *Law* and *Attribution*, *Law* and *MLAT*, and an interaction between all three variables.

Robustness Checks. DDoS attacks are not homogeneous in their nature. They differ in their level of sophistication, which usually depends on how much they cost or how difficult it is to execute them. I view the speed and/or size of the attack as a proxy for its complexity – the larger and/or longer an attack is, the more resources it takes to execute it. The level of sophistication can determine how much damage the attack can cause and what target it can effect. The fear of prosecution can deter future would-be perpetrators to execute attacks large in size and duration against the countries that have enabled laws and MLAT agreements. Specifically, I plan to include different thresholds for attack *size* and *duration*: (1) >50%,

[18] I do not include a model with country fixed effect as many of my independent variables are either time-invariant or change very slowly over time.

(2) 75%-95%, and (3) top 5% for large-scale attacks. Alternatively, I could have subset my data by a target to check the robustness of my findings. My data limitation does not allow me to run these results.

5 Discussion and Implications

If this study confirms my hypothesis, my findings will demonstrate that besides creating laws that punish perpetrators the country needs to have the ability to attribute cyber operations for these laws to be effective. If the origin of cyber attacks lies beyond national borders, a country needs a working Mutual Legal Assistance Treaty to punish culprits and deter would-be perpetrators. Even when these two conditions are met, the question comes to whether the damage from DDoS attacks was significant enough that the government is willing to spent its limited resources to punish the guilty ones instead of using them to deal with other national security issues. If it is the latter case, we might not observe a drop in low-level (or possibly in high-level) DDoS attacks in the countries that have MLATs and can attribute cyber attacks. For this reason, future research should consider individual country-level characteristics that can affect leaders' decision to prioritize other issues over the damage from DDoS attacks. The findings from this research will contribute to the literature that investigates the role of state legal cyber capability on deterrence and will have major implications for policy-makers.

References

Allen, Patrick D and Chris C Demchak. 2003. "The Palestinian-Israeli Cyberwar." *Military Review* 83(2):52.

Asal, Victor, Jacob Mauslein, Amanda Murdie, Joseph Young, Ken Cousins and Chris Bronk. 2016. "Repression, Education, and Politically Motivated Cyberattacks." *Journal of Global Security Studies* 1(3):235–247.

Borghard, Erica D and Shawn W Lonergan. 2017. "The Logic of Coercion in Cyberspace." *Security Studies* 26(3):452–481.

Coleman, Gabriella, Hoaxer Hacker and Spy Whistleblower. 2014. "The many faces of anonymous." *New York: Vero* .

Deibert, Ronald and Rafal Rohozinski. 2010. "Liberation vs. control: The future of cyberspace." *Journal of Democracy* 21(4):43–57.

Di Justo, Patrick. 2002. "How Al-Qaida site was hijacked." *Wired News* 10:54455–2.

Gartzke, Erik. 2013. "The myth of cyberwar: bringing war in cyberspace back down to Earth." *International Security* 38(2):41–73.

Goncharov, Max. 2012. "Russian underground 101." *Trend Micro incorporated research paper* p. 51.

King, Gary, Jennifer Pan and Margaret E Roberts. 2013. "How censorship in China allows government criticism but silences collective expression." *American Political Science Review* 107(2):326–343.

King, Gary, Jennifer Pan and Margaret E Roberts. 2017. "How the Chinese government fabricates social media posts for strategic distraction, not engaged argument." *American Political Science Review* 111(3):484–501.

Kostyuk, Nadiya and Yuri M. Zhukov. 2017. "Invisible Digital Front: Can cyber attacks shape battlefield events?" *Journal of Conflict Resolution* p. forthcoming.

Libicki, Martin C. 2009. *Cyberdeterrence and cyberwar*. Rand Corporation.

Lindsay, Jon R and Erik Gartzke. 2015. "Coercion through Cyberspace: The Stability-Instability Paradox Revisited." *Typescript, University of California, San Diego* .

MacKinnon, Rebecca. 2013. *Consent of the networked: The worldwide struggle for Internet freedom*. Basic Books (AZ).

Moses-Òkè, Roseline Obada. 2012. "Cyber capacity without cyber security: A case study of Nigeria's national policy for information technology (NPFIT)." *The Journal of Philosophy, Science & Law* 12(1):1–14.

Nye Jr, Joseph S. 2017. "Deterrence and Dissuasion in Cyberspace." *International Security* 41(3):44–71.

Romanosky, Sasha, Rahul Telang and Alessandro Acquisti. 2011. "Do data breach disclosure laws reduce identity theft?" *Journal of Policy Analysis and Management* 30(2):256–286.

Sanovich, Sergey, Denis Stukal, Duncan Penfold-Brown and Joshua Tucker. 2015. Turning the Virtual Tables: Government Strategies for Addressing Online Opposition with an Application to Russia. In *Annual Conference of the International Society of New Institutional Economics*.

Schelling, Thomas C. 1966. "Arms and influence." *New Haven: Yale* .

Segal, Adam. 2016. "The U.S.-China Cyber Espionage Deal One Year Later." *Council on Foreign Relations* .

Snyder Glenn, H. 1961. "Deterrence and Defense: Toward a Theory of National Security.".

Soldatov, Andrei and Irina Borogan. 2017. *The Red Web: The Struggle Between Russia's Digital Dictators and the New Online Revolutionaries*. Public Affairs.

Valeriano, Brandon, Jensen Bejnamin and Ryan C Maness. 2018. *Cyber Strategy: The Evolving Character of Power and Coercion*. Oxford University Press.

Valeriano, Brandon and Ryan C Maness. 2014. "The dynamics of cyber conflict between rival antagonists, 2001–11." *Journal of Peace Research* 51(3):347–360.

Valeriano, Brandon and Ryan C Maness. 2015. "Cyber Hype versus Cyber Reality: Restraint and Norms in Cyber Conflict.".

Villeneuve, Nart and Masashi Crete-Nishihata. 2011. "Control and Resistance: Attacks on Burmese Opposition Media." *Access Contested, ONI [OpenNet Initiative] Access: Denied, Controlled, Contested: http://access. opennet. net/contested/chapters/, geprüft am* 14:2012.

Westby, Jody. 2013. "Mandiant Report on Chinese Hackers is Not News But Its Approach Is." *Forbes Magazine* .

Yarhi-Milo, Keren. 2014. *Knowing the Adversary: Leaders, Intelligence, and Assessment of Intentions in International Relations*. Princeton University Press.

Blind Spots:
Tracking targeted threats to Civil Society in Reporting by the Infosec Industry

Lennart Maschmeyer[1]

Abstract

This paper focuses on an understudied issue in cyber conflict studies: targeted digital threats to civil society and the role of the infosec industry. It contributes an original dataset based on a systematic analysis of 364 threat reports by over 60 firms in the infosec industry. The study has two objectives: (1) determining how the industry reports the issue; and (2) triangulating global patterns in threat activity targeting civil society by tracing attack vectors, attribution, target type and the geographic distribution of activity. Its key findings are: 1) threats to civil society are not a priority in infosec reporting; 2) infosec reporting suggests a relative decline in the proportion of threat activity targeting civil society compared to other target types; 3) the majority of attacks rely on social engineering rather than sophisticated exploits; 4) authoritarian regimes are the main threat actors, with most activity attributed to China and Russia; 5) these regimes use ICTs to stifle dissent not only domestically, but across borders and within 'safe havens'; and 6) most reported activity takes places in East Asia and North America, while South America and Africa constitute blank spots on the threat map. The paper concludes with a discussion of the implications of these findings and a set of policy recommendations.

This paper highlights a neglected issue in cyber conflict: targeted digital threats[2] to civil society. Existing debates focus mainly on interstate or state vs. non-state interactions, identifying a power asymmetry in favor of small yet nimble threat actors. Cyber conflict is painted as a high-level, high-stakes game of cat and mouse where shadowy threat actors, many of them assumed to be state-backed, continue to outsmart and exploit large and powerful organizations and governments.

[1] The author would like to thank Columbia University SIPA and the Carnegie Corporation for their generous research support.

[2] Defined as "persistent attempts to compromise and infiltrate the networked devices and infrastructure of specific individuals, groups, organizations, and communities." (Citizen Lab 2014, 5)

In cyberspace, the offense is assumed to have an advantage, threatening critical infrastructure, governments agencies and the private sector. This threat landscape evolves rapidly and remains unpredictable, creating an environment of insecurity where cyber attacks routinely pierce through sovereign borders.

Where some see threats, others see opportunity. The security vacuum in cyberspace has facilitated the emergence of private security services providing protection against these new threats. Over the past years, a new business sector of information security services has grown into a multibillion-dollar business, offering clients protection and information on the threat landscape, called 'threat intelligence'. However, these firms also offer some of this information to the public at large, regularly publishing free reports on targeted threats they discovered. This public threat reporting has become the largest publicly available body of data on targeted digital threats, but what is reported is driven mainly by commercial interests.

Civil Society Organizations (CSOs)[3] are typically poor and understaffed, hence they provide little promise of business. Yet, as a body of reporting by independent research centers and human rights organizations has shown, CSOs face many the same threats as the resourceful actors that are in the spotlight of most reporting and academic writing on cyber conflict. The largest body of reporting on such threats is produced by the University of Toronto's Citizen Lab, as well as nonprofit organizations such as AccessNow, the Electronic Frontier Organization and Human Rights Watch. Foremost, the Citizen Lab's "Communities @ Risk" report provides an in-depth study of ten CSOs tracking targeted digital threats they face as well as how these threats impact the organizations over a duration of four years. It shows that several of the groups studies were in fact targeted by the same threat actors, using the same TTP as cyber campaigns previously highlighted in public infosec reports focusing on corporations and governments (Crete-Nishihata et al. 2014, 18). Research of this kind provides invaluable data, but producing high quality research requires time, resources and manpower—and as independent and nonprofit organizations, these institutions face constraints in all of these areas. Meanwhile, more information on the threat landscape faced by civil society is urgently needed. It is an urgent task because a functioning civil society is a vital part of democracy. Hence, the global reach of targeted digital threats to CSOs poses a risk to democracy and human rights worldwide.

This working paper provides a systematic analysis of the data provided by infosec reporting to assess its utility as a source of data and triangulate the threat landscape faced by CSOs. It contributes a dataset of the publicly available body of threat reporting, comprised of 364 reports by over 60 different firms. The main part of the paper proceeds with an analysis of this data to answer two main research questions. First, what is being reported by the industry? Second, based on this data, what threat landscape do CSOs face? This analysis in turn serves two main purposes. First, determining how the issue is prioritized by the industry, and thus evaluating the limitations of the data itself. Second, triangulating the threat landscape faced by CSOs based on trends and patterns evident in the data.

[3] These are defined here as independent organizations and groups who engage in non-violent political activity or research, and are aligned with the values of liberal democracy and human rights. This definition builds on White's more abstract definition of civil society as "an intermediate associational realm between state and family populated by organisations which are separate from the state, enjoy autonomy in relation to the state and are formed voluntarily by members of society to protect or extend their interests or values" (White 1994, 379)

To develop this analysis, the paper formulates a set of expectations concerning the threat landscape faced by CSOs based on the body of existing research. It will focus on research by Citizen Lab, an independent research center at the University of Toronto, because it has produced by far the largest and most rigorous body of work on the topic. It then employs descriptive statistics to confirm whether patterns evident in the data confirm or contradict expectations derived from existing research.

The key findings of this study are as follows. First, threats to civil society are not a high priority in infosec reporting. Although a sizable number of reports mentions the issue (20%), only a fraction of reports (4%) is dedicated primarily to campaigns targeting CSOs and low response rates to interview requests. Second, cunning social engineering rather than sophisticated technology is the key tactic used to gain access to systems in most attacks: spear phishing (highly customized and targeted phishing emails) are by far the most frequent attack vector. Third, most threats to CSOs come from authoritarian regimes. There is no confirmed case of a liberal democratic regime targeting civil society—domestically or abroad. Fourth, authoritarian regimes and their proxies are using ICTs to stifle not only domestic dissent, but also to project power abroad and target activists abroad. Threat actors from China and Russia are responsible for the majority of threats against civil society, trailed by Iran. Whereas Chinese actors focus mostly on domestic opposition groups, their Russian counterparts mostly project power extraterritorially, targeting CSOs in other states—as illustrated by the ongoing controversy about its meddling in the 2016 US Presidential Election. Fifth, most reported activity takes place in East Asia and North America, while South America and Africa constitute a blind spot on the threat map. This pattern stands in marked contrast to proportion of Internet users in the global south and threat activity in these regions documented by independent research. Sixth, infosec reporting suggests a relative decline in the proportion of targeted threats to CSOs compared to other target types. in contrast to countervailing trends in a proxy category: threats to small businesses.

I. Premise: Cyber Conflict and Civil Society

There is a widespread assumption among security scholars that cyber conflict is marked by a power asymmetry in favor of smaller, nimble threats actors. Large-scale actors such as corporations and government institutions are seen as vulnerable to attack by nimble and highly skilled threat groups, in some cases even individuals (Kramer 2009, 4–5; Nye 2010, 3; Lynn III 2010; Betz and Stevens 2011, 9; Brantly 2014, 474; NATO 2016; Nye 2011, 21; Perritt 1997, 427; Kello 2013, 35; Miller, Brickey, and Conti 2012; Tsagourias 2012; Libicki 2009, 43–52; Rattray 2009, 272). Size becomes a vulnerability: the larger a network, the more potential points of entry. It is near impossible to prevent intrusions since every piece of software contains vulnerabilities that can be exploited. Since anyone with internet-connected computer and the appropriate skillset can, in principle, break into systems, the proverbial teenager in a basement could bring an entire nation to its knees—as Miller, Brickey and Conti put it, “A Seventeen Year Old can Command an Army” (2012).

Despite their superior resources, powerful organizations continue to be bested by threat actors who require few resources besides their key asset: superior skills in deception and exploitation of vulnerabilities. A recent example is the 2017 breach at Equifax, a company with a market valuation of over 10B US Dollars, which led to the theft of personal information of over 140m Americans

(Siegel 2017). The follow-up investigation by the infosec firm FireEye revealed that the intrusion and data exfiltration was carried out from only about 35 different IP-Addresses (Goodin 2017). Even if each IP was used by multiple individuals, rather than obfuscating the origins of the attack, it was still a rather small group of people who, by virtue of a specific skillset rather than resource endowment, were able to compromise a company with close to 10,000 employees and exfiltrate personal information of half of the population of the United States.

The unfavorable correlation between network size and vulnerabilities, often exacerbated by an imbalance in skills, has facilitated the rise of a new industry: firms offering specialized services for detection and defense as well as forward-looking threat intelligence. The rapidly growing information security, short 'infosec', sector promises its clients to offset some of this imbalance by providing corporations with the skills and expertise needed to detect intrusions, defend networks and predict potential aggressors. FireEye, a leader in this sector, and contracted by Equifax to contain the fallout from its massive 2017 breach, underlines the resulting sense of insecurity in its corporate brochure: "Technology is outpacing our ability to secure it. Despite substantial spending on legacy security products, advanced attackers are bypassing these defenses at will and spreading unchallenged" (FireEye 2016a). Obviously, threat inflation is in the interest of companies promising to alleviate it. Accordingly, the CEO of SecureWorks, one of its competitors, reminds its clients that it is "not a matter of 'if', it is a reality of 'when' they [corporations] will be hacked…Most organizations are not resourced to effectively protect their IT environments and confidential data." (Cote n.d.) Unfortunately, however, the unabating stream of security breaches across major corporations suggests this seemingly hyperbolic statement is in fact not far from actual reality.

Leading infosec firms promise their clients not only a better understanding of the problem, but also a solution in the form of actively defending networks, hunting down intruders and, in some cases, even taking offensive actions against perpetrators (Cf. Kurtz 2013)—the still controversial practice of 'hacking back'. As such, these companies fulfil increasingly political roles as providers of security and protectors of property. Moreover, they are increasingly engaging with and competing with nation-states as the most advanced threat actors are typically state-backed. FireEye's brochure, for example, highlights the company's capacity to "provide nation-grade intelligence" (FireEye 2016a). This emerging role of infosec companies as intelligence brokers, increasingly competing with national intelligence agencies and potentially compromising operations, is unprecedented and raises important questions about the potential responsibilities that come with growing influence.

Kaspersky researcher Juan Andrés Guerrero-Saade highlights the perils of providing services to clients with questionable intentions, or the inverse risk of uncovering a legitimate espionage operation by a friendly intelligence agency that is perceived as aggression by the latter (Guerrero-Saade 2015). The solution, he suggests, requires "solid backchannel relationships" of infosec firms with intelligence agencies to avoid enraging the latter by compromising covert operations (Guerrero-Saade 2015, 7). Regardless of whether one agrees that this argument for responsibility is valid, the fact that it has become a point of debate underlines the growing capacities of the infosec sector, with leading firms now competing with some of the most powerful actors in cyber conflict. If this empowerment creates potential responsibilities towards the strong in cyber conflict, however, what about the weak?

The weak suffer what they must? The power asymmetry faced by civil society.

Considering this question highlights an inverse power asymmetry that is largely missing from the established picture of cyber conflict as portrayed in academic work: the precarious state of civil society. The power asymmetry between civil society and state actors is nothing new. However, today's technology provides states and state-associated threat groups with multipurpose, cheap and highly scalable, means of subversion and surveillance. This development poses a significant, in some cases even existential threat, to civil society organizations around the globe.

Citizen Lab research documents multiple cases involving the use of the same TTP used against powerful private sector entities and government agencies in campaigns targeting civil society organizations (Cf. Crete-Nishihata et al. 2014; Deibert and Rohozinski 2009; Scott-Railton et al. 2015).[4] While corporations often lack the expertise and skills to detect and fend off these threats, they do at least have the resources to hire the services of leading infosec firms that allow them to prevail. Typically priced at a premium level, these services are mostly inaccessible to struggling nonprofits (Citizen Lab 2014, 25). This problem is exacerbated by the fact that some of the most advanced threat actors today are associated with, or directly sponsored by, authoritarian regimes.

Rather than empowering liberal reform movements and activists to challenge authoritarian rule, as widely assumed, the increasing use of ICTs has provided these regimes with new and effective tools of surveillance and repression. Moreover, even regimes who lack the capabilities to develop sophisticated tools in-house have access to surveillance and espionage toolsets on par with some of the world's leading intelligence agencies on the commercial spyware market. Two of the leading firms in this space, Hacking Team and NSO Group have been used extensively across the globe (Bi. Marczak et al. 2014; Scott-Railton, Marczak, Guarneri, et al. 2017; B. Marczak and Scott-Railton 2016; Zscaler 2015). Contrary to prevailing expectations, authoritarian regimes have shown "now only resilience, but a capacity for resurgence" (Deibert 2015b, 64). The growing use of ICTs for targeted digital threats to civil society thus threatens not only these individual groups, but may in fact provide new avenues to crush dissent, stifle liberal reforms and undermine democracy itself.

In the past, tight-knit communities had a key advantage against state agencies: their internal recognition made infiltration difficult. While a new face (and thus potential government agent) in a tight-knit group of activists or an ethnic minority group is easy to spot, a compromised computer or phone is not. Advanced threat actors have perfected the art of compromising systems or software to monitor what a user is doing, to read their emails, and steal their data while everything appears normal. This information can be used to identify members of civil society, disseminate embarrassing information, reveal their location to law enforcement and provide evidence used in later prosecution. The malware required to do this is often delivered via a phishing email that tricks the user into clicking on an attachment or link that installs a piece of malware in the background. Some more advanced pieces of malware are even able to use a computer's phone to stealthily record any conversations in the vicinity, turning someone's own laptop into a listening device (CyberX 2017). While their smaller size and close personal linkages among members constituted a potential advantage against traditional surveillance and subversion operations, civil society is at a severe disadvantage against targeted digital threats. Compromised devices and services are hard to detect without adequate skills and resources, especially if one's adversary is a nation-state

[4] A full database of reports is available at https://citizenlab.ca/category/research/targeted-threats/

backed intelligence or law enforcement agency employing advanced espionage toolsets and exploiting yet unknown vulnerabilities in these devices or systems.

What sets targeted digital threats apart from more traditional means of surveillance and subversion is the ease with which tools, techniques and procedures developed for one set of targets can often be adapted to target a different set of targets, or simply trained on a wider group of actors. Once an effective social engineering toolset or an exploit for a zero-day vulnerability in a popular software has been developed for an international espionage campaign, for example, these same tools can be used against civil society groups at home or abroad at little or no additional cost, and are being used in practice. A 2014 report by Citizen Lab involving ten NGOs showed they were targeted by the same China-based APT as were large Fortune 500 corporations in the West. As authoritarian regimes increasingly rely on targeted, offensive means of stifling dissent and suppressing opposition, Ron Deibert warned in 2015 that “authoritarian systems of rule are showing not only resilience, but a capacity for resurgence. Far from being made obsolete by the Internet, authoritarian regimes are now actively shaping cyberspace to their own strategic advantage” (Deibert 2015b, 64). This point is echoed in a report by the Center for Long Term Cybersecurity at Berkeley, noting how growing interconnectivity offers an “unprecedented opportunity for governments and other adversaries to pursue attacks against PVOs” (CLTC 2018, 3). The lack of effective defenses against these advanced tools, which typically remain undetected by standard antivirus programs and require a dedicated effort by a skilled expert to identify and investigate, makes CSOs inviting targets.

CSOs not only lack defenses, however, they also face potentially more severe effects of cyber operations. Whereas businesses are at risk of financial loss, targeted individuals are at risk of personal safety or, in extreme cases, even death. An in-depth study of the effects of targeted threats on CSOs by Citizen Lab shows that in the most serious cases victims would “experience physical intimidation, abuse, detention, or imprisonment by authorities that stems in whole or in part from surreptitiously monitored communications” (Citizen Lab 2014, 24). Another less immediate, yet potentially more consequential long-term effect is the degradation of communication not only within the organizations themselves, but among the communities they aim to support (Citizen Lab 2014, 25; B. Cf. also Marczak, Scott-Railton, and Marquis-Bore 2014). The capacity to communicate effectively, and public discourse in general, is essential for the existence of civil society, however.

In short, the lack of resources and skilled labor and the personal vulnerability of members of CSOs combined with the resourcefulness and growing experience of threat actors as well as the relative ease with which tools used for international espionage can be adapted for political repression makes CSOs highly vulnerable to targeted threats. Since a functioning civil society is vital for the establishment, maintenance and survival of democracy (Castells 2008), targeted threats to civil society threaten to cripple not only the specific organizations targeted, but constitute a significant threat to democracy itself.

II. Research design

The significance of the problem is, unfortunately, not reflected by scholarly attention to the topic. In fact, a search on the University of Toronto's library service on relevant terms[5] only produced one relevant result: the Citizen Lab *Communities @ Risk* report mentioned further above. While there are several organizations producing dedicated research on threats to civil society, the University of Toronto's Citizen Lab, founded and directed by Prof. Ronald Deibert, is the only academic research centre focusing on this issue. Corresponding to this situation, the author is aware of only one peer-reviewed publication on the topic by other academic researchers (Le Blond, Uritesc, and Gilbert 2014). The Electronic Frontier Foundation, AccessNow, Privacy International and Human Rights Watch, among others, also produce high quality reports on threats to civil society and do important advocacy work. These reports are usually not peer reviewed, however, and neither do they typically provide broad, quantitative data on the extent of the problem. The only broad comparison the author is aware of is a 2012 study by AccessNow (2012), which has already become somewhat outdated due to the rapid evolution of targeted threats.

Therefore, this paper will focus on the body of research produced by Citizen Lab to formulate a set expectations on 1) infosec engagement with threats to CSOs and 2) patterns and trends among threat campaigns to verify in the data analysis. Over the past decade, Citizen Lab has produced a collection of in-depth case studies tracing targeted threats to civil society and their effects. These studies provide the best available source of data on the types of threats faced by CSOs, their evolution as well as some of the impacts on targeted organizations because Citizen Lab reporting constitutes not only the largest body of work on the topic, but also adhere to rigorous academic standards.

Expectations

E1: Targeted threats to CSOs are a low priority in infosec threat reporting.

Threats to CSOs are assumed to be low priority in reporting due to the two key business interests driving publication. Based on an inside perspective of an infosec researcher, reports are published 1) as a marketing tool to increase sales of the product or service offered by the company authoring the report and 2) to gain prestige and recognition by demonstrating analytical capabilities. Kaspersky researcher Juan Andres Guerrero-Saade suggests that, "the intended purpose [of reporting] is a PR-coup to both attract new customers for closed-release intelligence reports as well as garner brand recognition and industry respect for formidable findings" (Guerrero-Saade 2015, 4). Since CSOs are typically small and cash-strapped organizations, they do not constitute attractive clients. However, since these organizations are assumed to be regularly targeted by state-sponsored threat groups, investigating these threats may reveal sophisticated and hitherto unseen tools, techniques and procedures (TTP). Establishing the level of prioritization constitutes a challenge due to the lack of data for comparison—there is no data on the overall proportion of reporting on different target types and gathering this data would go beyond the scope of this project. However, this expectation can be verified or rejected with reasonable confidence based on inferences drawn from three different types of proxy data points:

[5] University of Toronto libraries, searches for the strings: "targeted threats" and "civil society", "cybersecurity and civil society" "cyber attack civil society", "cybersecurity non-government" (search date: 1/12/2018).

First, the proportion of threat reports with a primary focus on threats to CSOs. The lower the relative proportion of reporting prioritizing threats to CSOs, the lower the prioritization of the issue. Second, divergence between the proportion of threats to CSOs and threats to small business. Since small businesses and CSOs share the same vulnerability profile for targeted threats, it is reasonable to assume roughly similar proportions of threats targeting both types of organizations. In particular, since targeted threats are typically state-sponsored or proxies for states rather than cyber crime groups interested in profit alone. Third, interview responses by infosec representatives, and their overall response rate to interviews. Interview responses indicating low prioritization would constitute the clearest form of evidence verifying this assumption. In their absence, however, inferences can be drawn on the response rate to interview requests.

E2: Threat actors use the minimum level of technical sophistication necessary to achieve objectives, relying mostly on social engineering to achieve their objectives.

The use of the lowest necessary level of sophistication and reliance on social engineering has been a key pattern in past campaigns targeting civil society *(Crete-Nishihata et al. 2018; Hulcoop et al. 2016; Crete-Nishihata et al. 2014, 21).* This expectation would be confirmed if it becomes evident that the majority of online threats to CSOs reported relies on social engineering as the main attack vector.

E3: Most threats to CSOs are perpetuated by authoritarian regimes or their proxies.

Contrary to prevalent expectations of ICTs empowering liberal reform movements and civil society groups, authoritarian regimes have in fact been increasingly adept at using these tools to their strategic advantage, resurging and actively shaping cyberspace to their advantage (Deibert 2015b, 64, 2013, chap. 1; CLTC 2018). Therefore, we would expect most threat activity reported to be taking place in, or being attributed to, authoritarian regimes. This expectation would be confirmed by a majority of CSOs targeted within, and anti-CSO campaigns attributed to, authoritarian regimes.

E4: State-sponsored threat groups and their proxies are targeting CSOs across borders.

Prior research has illustrated that "targeted digital threats extend the 'reach' of the state (or other threat actors) beyond borders and into 'safe havens'" (Crete-Nishihata et al. 2014, 21). This expectation would be confirmed by evidence of extraterritorial campaigns of state-sponsored or other threat actors targeting CSOs.

E5: A significant proportion of threat activity occurs in the Global South.

The majority of Internet users today hails from the Global South, and some of the fastest growing online populations live in autocratic or authoritarian regimes (Deibert 2015a, 12).

Meanwhile, existing research documents multiple threat campaigns targeting CSOs in the global south (Crete-Nishihata et al. 2018; B. Marczak et al. 2017; Scott-Railton, Marczak, Guarneri, et al. 2017; Scott-Railton, Marczak, Bahr, et al. 2017). This expectation would be confirmed by a majority of threat activity taking place, or originating in, the global South.

E6: The proportion of targeted threats to civil society has been increasing relative to other target types.

While we lack broad quantitative data on the evolution of threats to civil society at the global level, the infosec industry provides regular updates on a useful proxy: small enterprises. This target category is a useful proxy for threats to civil society because they share two key properties. Although there are clearly significant differences between smaller businesses and civil society groups concerning the motivation for intrusions (criminal intent or economic espionage versus political repression) and, accordingly, the type of information sought, both actors share a key property that makes them increasingly attractive targets: a lack of effective cyber security skills and relevant resources (Jay 2017; McLean 2017) (Symantec 2013, 4; Jay 2017; McLean 2017). Accordingly, attacks against small businesses have been steadily on the rise. Symantec data shows that between 2011 and 2015 the proportion of targeted threats to small enterprises has more than doubled, increasing from 18% to 43% (Symantec 2015). This increased share of small businesses is significant since most CSOs share the vulnerability profile of these actors. Hence, it is reasonable to expect threats to CSOs to increase in a proportionate amount compared to small enterprises as their typically weak defenses provide the same opportunities for potential intruders (Haight 2015). This expectation would be confirmed by an overall increase in the proportion of targeted threats to CSOs relative to the proportion of other reported threat activity.

Methods and data

This study takes a mixed method approach. The key methods used are qualitative text analysis and coding, descriptive statistics and structured interviews. The source of data are 364 industry reports on targeted threats as well as one interview transcript[6].

The paper builds an original dataset based on analysis of these reports in order to derive descriptive statistics on key patterns among targeted threats to CSOs and gauge industry engagement with the issue. The overall strategy employed is one of triangulation, comparing and integrating insights from different types of data to seek for convergence and confirm expectations (Creswell 2014, 15, 201–10). The objective is determining the validity of patterns evident in public reporting compared to the findings of independent academic research (Citizen Lab Reports). Doing so allows determining how useful and reliable the data from public reporting is by identifying its shortcomings, and thus establishes the degree of generalizability of the findings drawn from analysis of this data.

There were three selection criteria for reports to be considered for analysis: 1) they must have been authored by an infosec firm, 2) they must focus on targeted threats and 3) they must have been

[6] This low number of interview participation in itself becomes a data point, as discussed further below.

published, either as full threat reports or as threat research blog entries on the company website. The period considered was the beginning of reporting (2010) until June 2017. The bulk of the reports were collected from the GitHub 'APTNotes' Repository (https://github.com/aptnotes/data) that is curated and maintained by members of the infosec community. To ensure all available reports were included, this was complemented by a search through the web archives of the individual firms to add any reports missing from the collection.

The analysis proceeded as follows. All reports were coded with the basic categories of company, year of publication, type of report and mention of threats to CSOs. Reports with a focus on CSOs were identified via keyword search for relevant terms.[7] Those reports were then analyzed in more detail and carefully coded along the following rubrics: type of CSO targeted, region, country, attack vector, attribution, threat actor and notification of victims.[8] which groups in which regions and which countries are targeted by actors from which countries. Finally, the results of this analysis were used to derive descriptive statistics. This analysis was followed up by interview requests sent to representatives of five leading threat intelligence firms. These structured interviews involved stringent privacy protection measures that were approved by the University of Toronto's Research Ethics Board. The aim of the interviews was to garner additional data on the prioritization of threats to civil society in the infosec industry, and the incentives shaping public reporting. As mentioned at the outset, multiple requests sent to several representatives at each firm ultimately resulted in only a single interview. While the low participation rate underlines the expected low prioritization of the issue, the data from this single interview can merely provide anecdotal evidence and is only used as such. Apart from this limitation, there are two additional limitations concerning reporting data that need to be addressed.

First, data garnered from a public report will be both incomplete, as more detailed findings are typically reserved for paying clients and likely biased towards the solution offered by the company in question. As a threat intelligence researcher from a leading firm stated in an interview, the ratio between internal, private reporting and public reporting is something around 300:1 (Threat Intelligence Researcher 2018). Public reporting this provides only a fraction of the whole picture. That being said, the same researcher highlighted that a core criterion for publication of a report is the discovery a unique tool, target type, tactic of strategy. Assuming that this is a general incentive, each report thus presents a unique case in its own way, making qualitative analysis the most useful tool of analysis (George and Bennett 2005, 138). Moreover, the threat can be assumed to be presented in the direst terms in order to maximize the sense of insecurity among potential clients and thus the demand for security products and services. However, since incentives driving publication are universal across the industry, overall the data gained from these reports likely still provides a representative overview of the threat landscape perceived by these firms. In short, while it is important to keep in mind the limitations of the data gained from threat reporting, they provide the best available source of broad data on the evolution and patterns in targeted threats against civil society.

Second, the lack of conformity in naming of threat groups and the competition for media attention may result in multiple reports on what is in fact one operation. For example, reports on the threat groups involved in meddling in the US elections have proliferated over the past year, with overlaps

[7] The strings searched for were: "NGO", "non-government", "non-profit", "nonprofit", "dissident", "opposition", "journalist", "activist", "civil society", "think tank".

[8] For a more detailed description of the criteria and coding of each rubric, please see appendix 1.

between reports and rival naming schemes that make disentangling individual campaigns and objectives non-trivial. This situation is exacerbated by the practice of using proxy groups to obfuscate origins of campaigns that is a common aspect of Chinese and Russian operations (Galeotti 2016; FireEye 2013). Since the objective of this practice is obfuscating origins, the lack of data on the extent of this practice means conclusions about overall patterns in threat activity based on country profiles are preliminary. Yet this does not necessarily render findings invalid. While some of the campaigns reported that were counted towards China's record engagement with domestic actors may be incarnations of a larger campaign orchestrated by a central agency, the fact that the required proxy groups have been established and are being maintained nonetheless reflects significant investment into cyber operations. Moreover, it reveals the political significance that the actors behind these campaigns ascribe to them. If these were not important, why take such elaborate measures to obfuscate their origins in order to minimize the risk for repercussions?

Based on these points, and taking account the limitations of the data used, the approach of triangulation employed here will still provide useful insights into the overall patterns in threats faced by civil society. Moreover, the comparison of expectations derived from the findings of academic, case-based research to patterns observed in industry reporting will specify the shortcomings of the data itself. Since Citizen Lab reporting has followed a consistent methodology, and constitutes a cumulative research programme, this is a valid comparison. However, the ambitions of this analysis remain exploratory in scope, hence the focus on confirming/rejecting expectations. Findings generated from this analysis will be essential to further develop this research agenda and formulate (causal) hypotheses. The conclusion will identify potential avenues for such research.

III. Findings

Findings on the overall proportion of reporting on threats to CSOs, interview response rates and the relative decline in reporting on threats to CSOs support the expectation that these threats are not prioritized (E1). The relative decline on reported threats to CSOs challenges the assumption that these threats are increasing due to the vulnerability of the target type and growing capabilities of state actors (E2). Based on the data from reporting, E2 needs to be rejected, yet considering additional data on threats to small businesses, this finding supports the overall lack of prioritization suggests a relative decline in attention by the industry rather than decrease in threat activity in the real world. This point will be addressed in the discussion.

As illustrated in Fig. 1, analysis showed that 62 out of the 364 reports analyzed discuss some form of targeted threats to CSOs.

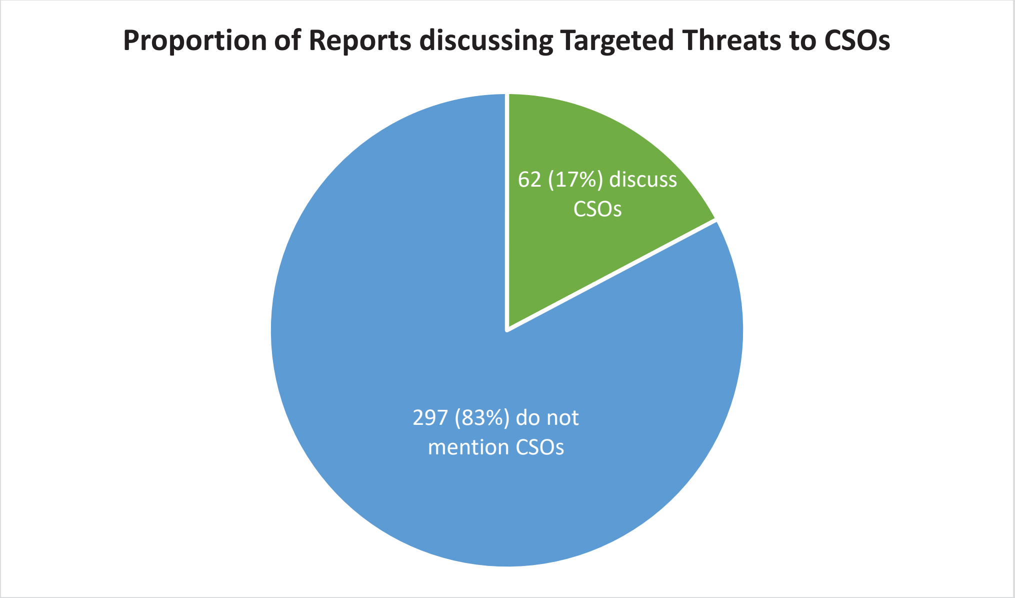


Fig. 1: CSO reporting overall.

This is a sizeable proportion of reporting, especially considering that these organizations are not typically attractive potential clients for infosec firms. Based on this incentive alone, one would not assume resources spent on investigating and reporting on threats to CSOs apart from passing mentions. However, a deeper look reveals that only a small fraction of reporting puts its primary focus on targeted threats to CSOs: as Fig. 2 illustrates, only 13 reports, or 4% of total reporting, place their primary focus on civil society. The remaining 47 reports either place a secondary focus on CSOs (23 reports, 6% of total reporting), discussing them very briefly in part of a larger analysis, or mention CSOs only in passing (24 reports, 7% of total reporting).

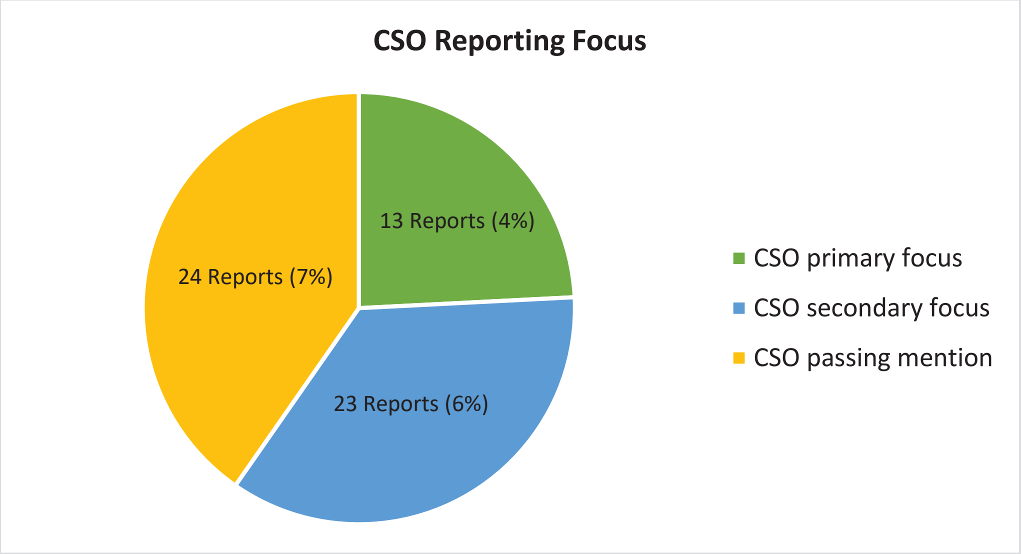


Fig. 2: CSO reporting focus.

The finding that only 4% of reporting puts a primary focus on threats to CSOs supports the assumption that these threats are not prioritized. This assumption is further supported by the low response rate to multiple interview requests sent out to the six leading firms.

To determine how the industry engages with this issue, interview requests were sent to representatives at the six leading infosec firms (FireEye, CrowdStrike, Kaspersky, Secureworks, Symantec, and Trendmicro) in September 2017. Two of the interview requests initially got a positive response, but the representatives did not reply to multiple attempts at setting up an interview. Subsequently, on January 16, 2018 a short questionnaire was sent out (see Appendix 3) to these firms' general contact form to gather data on their engagement with the issue in question. None of the firms replied. Ultimately, out of 19 requests sent in total, only one researcher from one of these firms agreed to an anonymous interview, reflecting a response rate of around 5% -- compared to an average response rate to interview requests of 36% (Yang, Wang, and Su 2006). The low response rate thus provides further support for the assumption that threats to CSOs are not a priority in reporting.

This interview in turn provided anecdotal data for the potential incentives behind prioritization of threats to CSOs. While stating that reporting is not prioritized according to target categories, the researcher also highlighted the benefits of investigating threats to CSOs due to the frequency of attacks and the fact that government-sponsored threat actors may use new tools and techniques against them as a proof of concept, as a test before deploying these TTPs against higher value targets (Threat Intelligence Researcher 2018)—aligning with the expectation that CSOs face the same threats as large corporations or governments.

Finally, findings on the relative proportion of reporting on threats to CSOs are also congruent with the expected low prioritization of threats to CSOs in reporting. As outlined in the previous section, CSOs share the same vulnerability profile as small businesses. Since the proportion of targeted threats to small businesses has been increasing, relative to medium and large-size businesses, it was thus expected that threats to CSOs would also increase relative to other threat activity. Contrary to this expectation, however, this proportion has declined significantly. The overall number of reports published grew sharply from 13 in 2011 to 88 in 2014 (an almost sevenfold increase), the relative proportion of reports discussing threats to CSOs declined from 31% (4 out of 13) in 2011 to 18% (16 out of 88) in 2014. Since 2014, the absolute number of reports on CSOs published has decreased, and their relative proportion declined even further (12% in 2015, 15% in 2016).

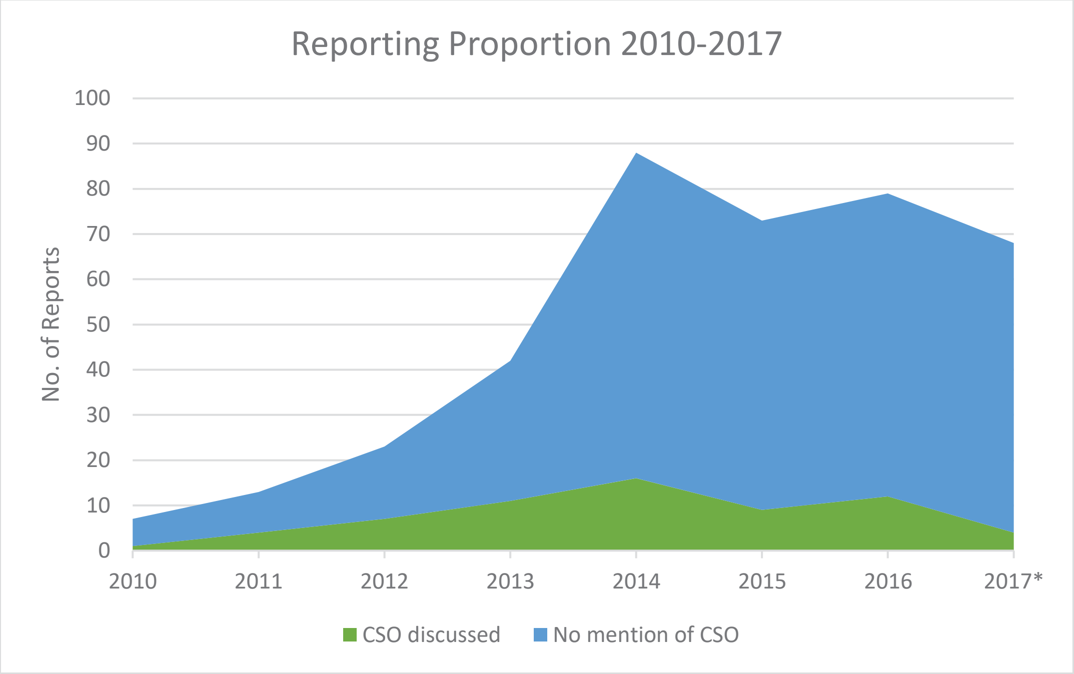


Fig. 3: Reporting trends.
* = Yearly estimate, extrapolated from Q1+Q2 data.

Finally, findings on notification of victims of attacks provides further evidence to buttress E1. Only 2 out of 62 reports indicate that victims were notified: the Bellingcat report noted above (ThreatConnect 2016) and TrendMicro's report on Operation Pawnstorm (APT28) targeting dissidents of the Russian government (TrendMicro 2014). The Comfoo report by SecureWorks suggests 'many' of compromised victims were notified (SecureWorks 2013), meaning it is likely some of the think tanks targeted were notified. The remaining 59 reports (95%) do not indicate whether victims were notified. One can only hope that the actual rate of notification is significantly higher than the rate of notification disclosure.

Combined, findings on the overall proportion of reporting on CSOs, interview response rates and reporting trends thus support E1. Meanwhile, E2 is not supported by the findings.

E3: Attack vectors

Findings on attack vectors verify the expected reliance on social engineering and low (minimum necessary) sophistication (E3). As illustrated in Fig. 3, low-sophistication operations relying on effective social engineering techniques, rather than technically sophisticated threats, pose the main risk to civil society actors.

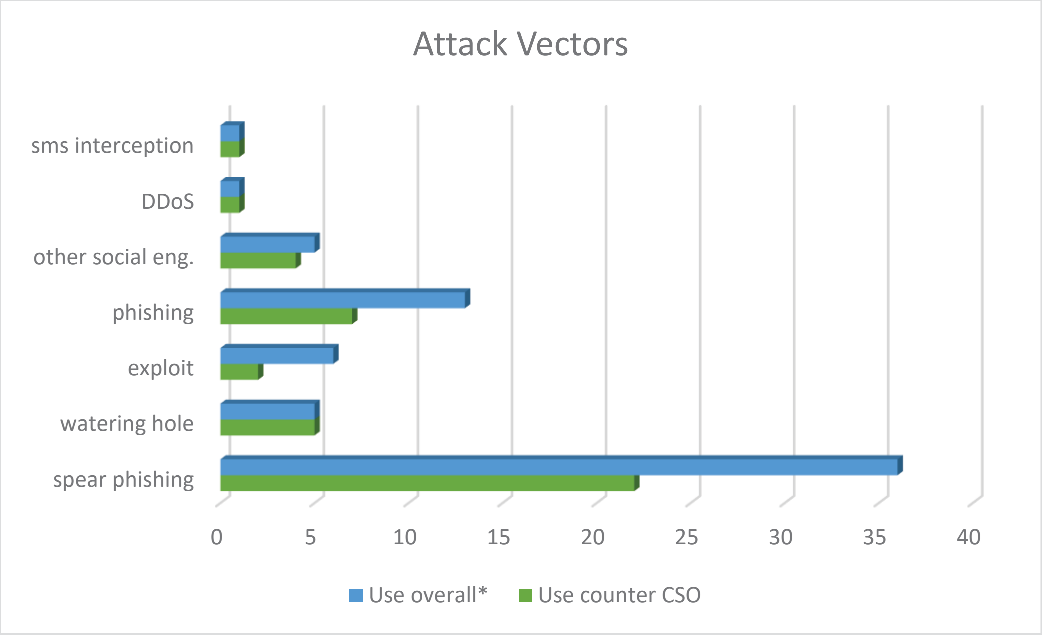


Fig. 4: Attack vectors
* preliminary findings, coding incomplete.

Over half of the operations targeting CSOs relied on some form of social engineering. 38 out of 62 these operations (55%) relied on spear phishing, phishing, a watering hole[9] or some other form of social engineering[10] as the initial attack vector. These proportions clearly verify E3, and correspond to overall trends in the digital threat landscape, marked by an increasing shift towards social engineering (ProofPoint 2017).

Exploits, where vulnerabilities in software or hardware used allows intruders direct access to systems without the knowledge of users, in contrast were used only in two operations (3%). One of the operations in question were Greedy Wonk, which exploited a vulnerability in Adobe Flash to redirect visitors to the websites of the following institutions to a malicious link: the Peter G. Peterson Institute for International Economics (Think Tank, Washington D.C.), the American Research Center in Egypt (independent research center in Cairo, Egypt) and the Smith Richardson Foundation (Think Tank in Westport, US) (FireEye 2014). The other campaign is the Aurora compromise, also known as Hydraq, which exploited a zero-day vulnerability in Microsoft Internet Explorer to gain access to Gmail accounts—targets ranged from high-profile corporate actors to human rights activists in China (CA ISBU 2010). One of the most devious operations was the 2016 compromise of the email server used by journalists of the Bellingcat collective, reporting among other things on the downing of flight ME17, likely obtained through interception of SMS-messages sent to the journalists phone as part of the two-factor authentication process (an additional safety measure) (ThreatConnect 2016). ThreatConnect attributed this campaign to APT 28, suggesting that, if necessary, these Russian actors are capable of considerably more sophisticated and elaborate tactics to compromise targets. The result was not only embarrassing, but potentially

[9] A watering hole attack refers to the compromise of a website frequented by targeted individuals (TrendMicro 2013).
[10] Usually this means the report in question did not further specify the technique used.

dangerous for the targeted journalist as highly personal information, including a scan of his passport, was then published online.

This final point highlights the divergence in effects between corporate actors and CSOs targeted by the same campaign. In addition to financial losses, targeted threats that compromise CSOs' systems put their members and constituents personally at risk, including of physical persecution by law enforcement or other agents of authoritarian regimes. Considering this situation, notification of victims of threat campaigns is crucial to their safety.

E4: Regime type

Authoritarian regimes were assumed to be the main perpetrators of targeted threats to CSOs and the findings clearly support this expectation: all reported activity is attributed to authoritarian regimes (as classified by Freedom House).

Before discussing attribution, this section will first present findings on overall geographic distribution of threat activity. Considering that two converging interests, the need for maintaining political neutrality and a lack of business opportunities in the victim community, run counter to publication of threats to civil society under repressive regimes, the amount of reporting on the topic is surprisingly large. In fact, as the detailed breakdown in Table 1 shows, with 44 out of 76 (58%), the lion's share of CSOs whose locations were reported are situated in repressive regimes.[11]

Tab. 1: Regime index of targeted CSOs

Country	**No. targeted CSOs**	**Freedom house score (higher = less free)**[12]
Palestine	1	N/A
Australia	1	1.0
Canada	1	1.0
Japan	3	1.0
United States	12	1.5
Mongolia	1	1.5
South Korea	3	2.0
Israel	2	2.0

[11] Characterized by a FreedomHouse score of 6 or higher, with regime types including 'one-party or military dictatorships, religious hierarchies, or autocrats' (Freedom House 2018)

[12] Source: Freedom House 2018 country scores, available at https://freedomhouse.org/report/freedom-world-2018-table-country-scores (last accessed January 19, 2018)

India	2	2.5
Ukraine	3	3.0
Pakistan	1	4.5
Myanmar	1	5.0
Jordan	1	5.0
Iran	4	6.0
Vietnam	1	6.0
Egypt	3	6.0
China	18	6.5
Russia	6	6.5
Syria	2	7.0
Tibet	10	7.0
Total	***76***	-

Displayed on a map, as in Fig. 6, this skewed distribution becomes even more starkly clear as more than a third of these organizations are located within a single territory: China has the questionable honor of being host to by far the most targeted CSOs, with 28 out of the 76 organizations whose location was specified in reporting, or 37% of total reported. Out of these, nearly half target Tibetan activists.[13]

[13] The map diagram used does not include an option to display results for the autonomous region of Tibet separately, hence they are displayed under China.

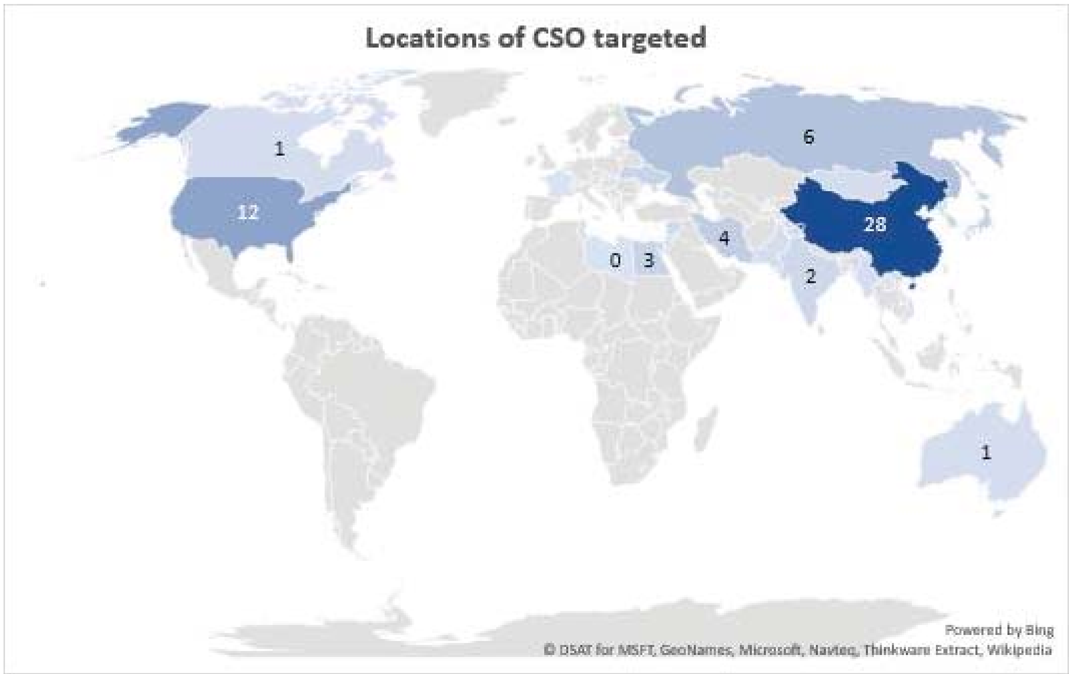

Locations of CSO targeted
1
6
12
28
4
0
3
2
1
Powered by Bing
© DSAT for MSFT, GeoNames, Microsoft, Navteq, Thinkware Extract, Wikipedia

Fig. 4: Locations of targeted CSOs

The concentration of other threats to CSOs in authoritarian states such as Russia, Iran and Egypt corresponds to the pattern of repressive regimes becoming increasingly adept at harnessing the opportunities provided by information technology to bolster their power documented in existing research (Ronald Deibert et al. 2010, 2008, 2012). However, the fact that the second most targeted group are civil society organizations located in the United States is likely to raise eyebrows. Where do these threats come from?

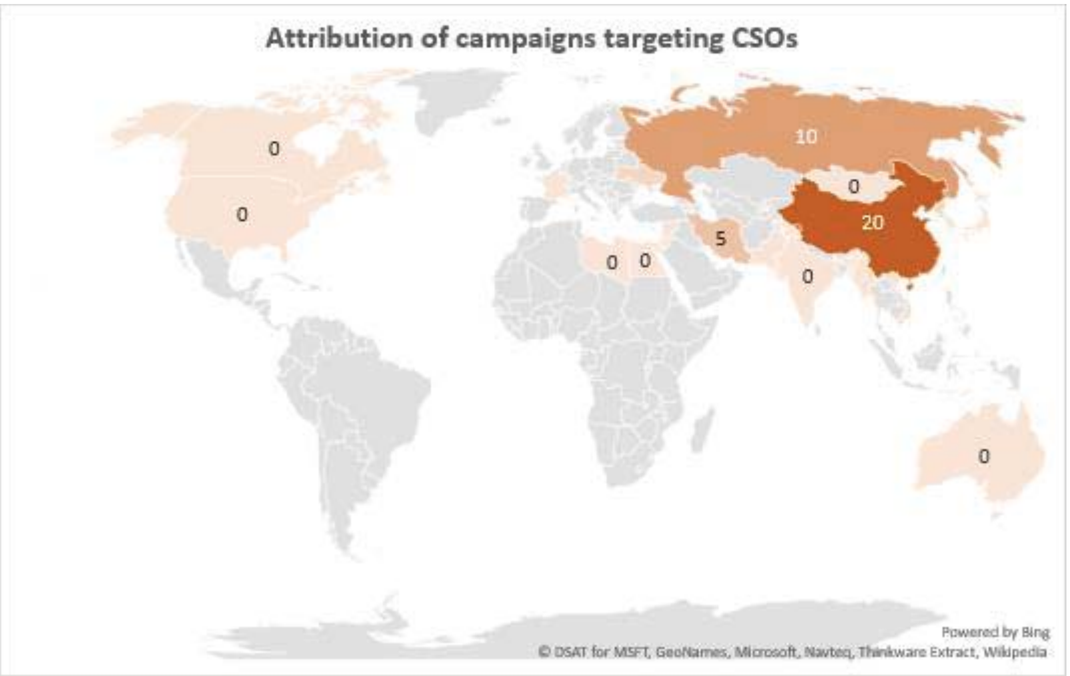

Attribution of campaigns targeting CSOs
0
0
10
0
20
5
0 0
0
0
Powered by Bing
© DSAT for MSFT, GeoNames, Microsoft, Navteq, Thinkware Extract, Wikipedia

Fig. 6 : Attribution of campaigns targeting CSOs

While the US is home to 16% of targeted CSOs (12 out of 76), based on attribution by the industry none of these attacks actually originate in the US. In fact, none of the anti-CSO campaigns discussed in threat reports are attributed to liberal regimes. Four out of the five states to whom anti-CSO campaigns are attributed are authoritarian: China (20 campaigns), Russia (10 campaigns), Iran (5 campaigns), North Korea (2 campaigns) and the Ukraine (2 campaigns). The Ukraine as a semi-liberal regime is an outlier here, and the contested political situation in the country makes attribution even more fraught than usual. The two campaigns in question are phishing campaigns, one targeting journalist of the Bellingcat group (ThreatConnect 2016) and the other targets, among others, Ukrainian journalists with the "Groundbait" espionage toolkit (ESET 2016), both attributed to threat actors operating from within Ukrainian territory. The fact that the Ukrainian intelligence service is heavily penetrated by Russian operatives ("30% of Ukrainian SBU Officers Were Russian FSB and GRU Agents -" 2014; Kuzio 2010) and the open territorial disputes in the Ukraine means attribution remains highly speculative. The targets of the Bellingcat campaign align with Russian objectives whereas the Groundbait campaign, targeting mainly separatist groups in East Ukraine, is congruent with Ukrainian interests—but each may also be a false flag operation by the other side. Importantly, however, if it were confirmed to be operated by an actor linked to the Ukrainian government, the Groundbait campaign would be the only use of a targeted threat by a liberal government against domestic civil society in the dataset.

Findings on reported attribution of targeted threats to CSOs thus provide sufficient evidence to verify E4 as all attributed threats were sponsored or perpetrated by authoritarian regimes.

E5: domestic and extraterritorial targeting patterns

Existing evidence suggests authoritarian regimes and their proxies are using ICTs not only to stifle dissent domestically, but to extend their reach across borders and target civil society in 'safe havens'. The findings provide strong support for this expectation.

For some repressive regimes the use of cyber means to silent dissenters and stifle opposition movements at home has become routine business. China by far leads the world in this regard: targeted threats attributed to actors associated with the Chinese government constitute a staggering 64% of all reported campaigns targeting domestic civil society, 16 out of 25 such campaigns reported overall. The only other states are Russia and Iran (3 campaigns each), while the uncertainty of the political situation in Ukraine prevents attribution to one side with reasonable confidence even if the origins can be pinned down geographically.

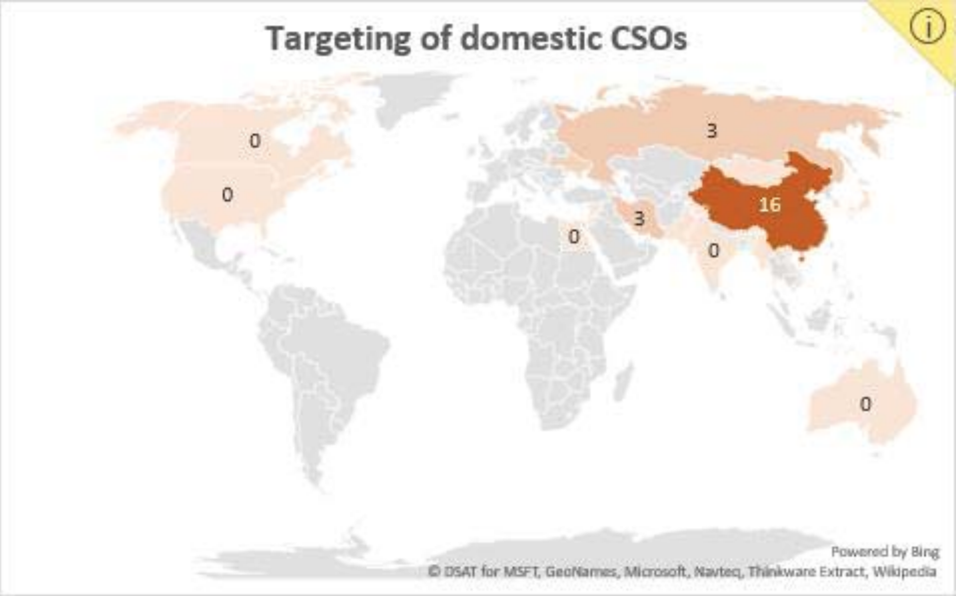

Targeting of domestic CSOs
0
0
3
16
3
0
0
0
Powered by Bing
© DSAT for MSFT, GeoNames, Microsoft, Navteq, Thinkware Extract, Wikipedia

Fig. 7: Domestic CSO targeting.

The uncertainty surrounding attribution is pervasive, highlighting a key strength of the use of cyber means in general and in specific against civil society: the ability to maintain plausible deniability. In order to avoid traceability, a more recent analysis suggests China has developed an organizational model that intentionally obscures state-sponsorship by distributing operations across a set of seemingly competing smaller actors who also engage in criminal activity alongside state-driven campaigns (FireEye 2013). Over the course of the investigation, the FireEye researchers discovered that “what we initially believed to be 11 different APT campaigns used the same malware tools, the same elements of code, binaries with the same timestamps, and signed binaries with the same digital certificates.” (FireEye 2013, 5). The lack of clear connections across the campaigns identified in the Lurid Downloader report can thus reasonably be assumed to be the product of active obfuscation parallel to the tactics uncovered by FireEye, underlining the threat faced by civil society as it gets targeted by such experienced, highly resourceful and persistent actors likely to be backed by government yet unlikely to be revealed as such.

Conversely to the use of tools, techniques and procedures from international espionage campaigns on domestic targets that characterizes Chinese modus operandi, a more recent trend is the use of cyber means to project power against civil society abroad. While China’s threat groups focus on economic and political espionage abroad and curtailing of civil society at home, Russian threat groups have emerged as world leading in targeting CSOs extraterritorially.

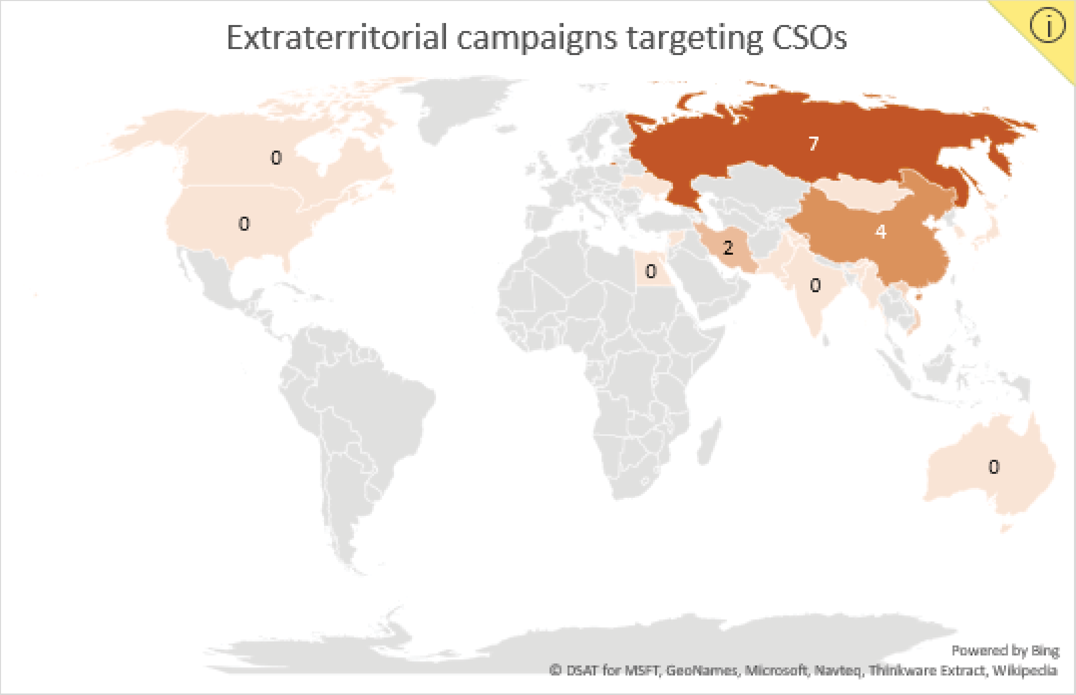

Extraterritorial campaigns targeting CSOs
0
0
7
4
2
0
0
0
Powered by Bing
© DSAT for MSFT, GeoNames, Microsoft, Navteq, Thinkware Extract, Wikipedia

Fig. 8: Extraterritorial targeting of CSOs.

As illustrated in Figure 4, Russian threat actors are behind 41%, or 7 out of 17, of all targeted extraterritorial cyber campaigns aiming to monitor, subvert and disrupt political opposition abroad. Targeting patterns correspond to Russian interests, targeting either its traditional sphere of influence in Central Asia and East Europe (4 out of 7) or its former Cold War rival, the United States (3 out of 7 campaigns). The emerging pattern of cyber operations against political opposition forces abroad underlines the extent to which the Russian government is using non-violent cyber means as effective tools of power politics.

The three Russian campaigns targeting the US all employ spear-phishing and are related to the now infamous threat actors linked to the meddling in the 2016 presidential election. Apart from the high-profile targets involved in the latter breach, these campaigns constitute a long-term, concerted assault on non-government organizations and think tanks in Washington as well as authors and journalists with an interest in Russia. For example, the threat actor behind the DNC compromise, known among other names as APT28 or Fancybear, ran a phishing campaign with the same technique as early as 2015 (SecureWorks 2016b). This campaign used the same fake google login page used, among others, to compromise John Podesta's account (SecureWorks 2016a). The third phishing campaign was run by APT29 / Cozybear via a compromised email server at Harvard University, commencing almost immediately with the election of President Trump and focusing on US non-governmental organizations and think tanks (Volexity 2016).

Yet half of the campaigns targeting CSOs in the United States have unknown origins as reports do not establish clear attribution. Although lacking explicit attribution, several of these campaigns bear significant hallmarks of Chinese origins. For example, the reports on the Voho campaign by RSA that targeted actors "involved in business and local governments in Washington, DC and Boston, Massachusetts, as well as organizations involved [in] the development and promotion of democratic process in non-permissive regions" (RSA 2012) lacks clear attribution yet clearly indicates Chinese origins since the campaign shared key characteristics with two infamous Chinese operations: Aurora and Gh0stNet.[14] Similarly, the Trojan Taidoor campaign reported by Symantec and TrendMicro (Symantec 2011; TrendMicro 2012) is not clearly attributed to China, but its targeting of 'think tanks engaged in US and Taiwanese affairs' as well as the infrastructure used for command and control of the malware used being located in Hong Kong (Symantec 2011) quite clearly indicate Chinese origins. Rather than concluding China is less engaged in targeting of civil society, these cases suggest Chinese threat actors may be better at hiding their tracks. If experience in the routine use of cyber means against domestic groups does give states an edge in cyber conflict, based on their record in domestic targeting Chinese threat actors should be among the most sophisticated. In line with this assumption, a 2016 FireEye report tracking Chinese activity following the agreement between Obama and Xi Jinping to curb cyber espionage suggests the observed decline in activity may in fact be the result of increased sophistication (FireEye 2016b).

Other campaigns targeting US CSOs whose origins remain cloudy are ShadyRAT (McAfee 2011), Scanbox II (PwC 2015) and Miniduke (Kaspersky 2015). Finally, the patchwork campaign targeting NGOs and government organizations across the US, Japan, China and the UK (Cymmetria 2016) is most likely Indian. Although neither Cymmetria nor Symantec (2016) attempt attribution, Kaspersky's report on the same actor only thinly veils the likely Indian origins: apart from titling the report "Dropping Elephant", Kaspersky indicates that the only traceable connections to the malware were made from India ("The Dropping Elephant Actor" n.d.).

To conclude, country profiles reveal China, Russia and Iran as the main actors targeting civil society. Whereas China mostly focuses on domestic targets, Russia is increasingly aggressive in targeting CSOs and political organizations abroad. However, Chinese threat actors tend to be more sophisticated and thus likely better at preventing attribution. Thus, China should not be prematurely dismissed as a threat to civil society.

E6: Global South

Although the majority of internet users now live in the global South, this prevalence is only partially reflected in reporting. While there is a wealth of reporting on China (37% of CSOs targeted are located in China, while 41% of attributed campaigns are tied to China), South America and Africa (except for two cases in Egypt) are blank spots on the map. Findings from reporting thus disconfirm the expected prevalence of threat activity in the global South. However, as the discussion will address, this finding in fact highlights potential distortions in the threat landscape painted based on reporting data due to the incentives driving publication.

To conclude, findings confirm E1, E3, E4 and E5 while E2 and E6 must be rejected based on analysis of threat reporting. The final section will now discuss the implications of these findings.

[14] Aurora refers to a highly sophisticated campaign linked to the Elderwood team that succeeded in breach Google and a range of high-profile targets in the US in 2009 (HBGary 2010; Security 2010b, 2010a). Gh0stNET in turn refers to the threat actor that may have started all threat reporting, a large-scale Chinese espionage campaign targeting not only the Tibetan exile government, but foreign ministries across the world ("Tracking GhostNet: Investigating a Cyber Espionage Network." 2009).

IV. Discussion

The analysis confirmed most of the expectations generated based on previous research (E1, E3, E4, E5), yet both the reported trends in targeted threats to CSOs and their geographical distribution (E2, E6) did not confirm expectations. Why?

There are two possible answers: either existing research got it wrong, or reporting provides an incomplete or distorted picture of the actual threat landscape. This discussion will focus on exploring the latter question.

First, contrary to expectations, the relative proportion of reporting on targeted threats to CSOs has declined since the beginning of threat reporting in 2011. Meanwhile, as discussed, the proportion of threats to small businesses, sharing the same vulnerabilities as CSOs has significantly increased. There are three possible explanations. First, CSOs as a target type may have some unknown properties that make them less attractive to threat actors compared to small businesses. Second, this divergence may reflect the relative growth of targeted threat activity by cyber criminals motivated by financial gains compared to targeting of political groups. Due to their lack of resources, CSOs are not attractive targets for criminals and have been typically targeted by state-sponsored or associated actors for their political activity. There is evidence to support this conclusion. For example, an Accenture survey of 254 companies found that in 2017 alone, the number of annual security breaches had increased by 27%, and the financial costs of such criminal activity had increased by 22% (Accenture 2017). Third, a significant proportion of threat activity targeting CSOs is not reported.

All three explanations are plausible and cannot be confirmed or rejected without additional research. Especially a survey of CSOs tracking breaches and effects, analogous to the Accenture study, would provide fascinating insights. The likelihood that CSOs have some property that makes them inattractive targets on the other hand appears very low, considering the amount of in-depth research on the topic that has already been done. In particular the Communities at Risk study investigated the conditions prevalent in the 10 CSOs included extensively. Based on the data that is available, moreover, the third explanation appears most likely—which leads to the second key point that has come out of the analysis.

Second, findings suggest that reporting is geographically skewed. Almost all reported activity occurs in North America, Europe, the Middle East and Asia. South America remains a blank spot, and based on the data available Egypt would appear to be the only country in Africa where targeted threats to CSOs are an issue. This conclusion is challenged by the wealth of research that has documented expansive online operations targeting CSOs in both regions.

For example, a set of recent investigations has revealed a widespread campaign of using spyware against human rights activists and NGOs in Mexico (Scott-Railton, Marczak, Bahr, et al. 2017; Scott-Railton, Marczak, Guarneri, et al. 2017) as well as a vast malware campaign targeting civil society across multiple South American states over seven years (Scott-Railton et al. 2015). Similarly, only last year a cyber operation targeting Ethiopian Dissidents was revealed (B. Marczak et al. 2017) while the CheckPoint Global Threat index placed four African countries among the highest risk for organizations to be targeted by cyber attacks (Checkpoint 2017). CSOs in these developing regions are likely to be even more vulnerable to targeted threats than in more developed regions since they will relatively fewer resources and less access to skilled personnel.

Moreover, without dedicated reporting, targeted threats are likely to be undetected as those tasked with protecting networks have their hands full with keeping systems operations. According to Neil Blasevic, ICT Manager at DefendDefenders[15], "we probably do encounter APT's but we just treat them as normal infections and don't identify them specifically, focusing on just cleaning them out" (Blazevic 2018). To alleviate this situation, Blazevich highlights the need for "more collaboration on monitoring and sharing malware samples [with the infosec sector] as well as training technologists and trainers who can do triage and collection…at the moment that isn't really happening." (Blazevic 2018).

These points highlight the persistent lack of comprehensive and representative data on both the scope and the scale of targeted digital threats to civil society at a global level. As the analysis in this paper has shown, public reporting by the infosec industry provides a rich, yet imperfect picture of the threat landscape. However, despite its imperfections, the findings underline the extent of the problem as threats to CSO are pervasive even where they are clearly not a priority. Hence, the core question emerging out of this study: how could this situation be improved?

Conclusion: The need for a coordinated research agenda and more public-private collaboration

This paper has contributed a systematic analysis of public reporting on targeted threats to civil society that illustrates a perilous power asymmetry between civil society and threat actors that has received too little attention in the literature on cyber conflict. Based on these findings, and the contributions of the participants of the 2018 Global Digital Futures workshop on this topic, this conclusion proposes four recommendations to policy-makers that will help improve the situation of civil society.

First, increasing funding for dedicated research on the topic. To alleviate the lack of comprehensive and representative data, the most obvious step is more academic research on threats to civil society. Doing also requires acknowledgement of the need for interdisciplinary research as this issue included both political, legal and technological aspects. Independent research centres such as Citizen Lab have demonstrated the effectiveness of this approach. The instruments and strategies used by authoritarian regimes to stifle dissent and suppress opposition are a well-established field of research, hence it is surprising the use of these tools in the digital realm has not received more attention. While this is a long-term aim, the reliability of existing data could be significantly improved in the short and medium-term through better understanding of the rationales driving public reporting by the infosec sector, the policy for exclusion or inclusion of targets, the attribution policy, notification policy and evolutions in the threat landscape over time. Interviews with management and analysts across the sector will be the best means of obtaining this data. The objective is to develop an overall understanding and variation across individual firms as well as over time. The findings presented in this paper will be a useful basis for follow-up interviews.

Second, a prioritization of civil society in regulatory and normative frameworks. While critical infrastructure has been the focus of attention of most policy-makers, civil society should be considered no less important due to its vital role for democracy.

[15] An NGO providing assistance against digital threats to CSOs in East Africa.

Third, raising the stakes by naming and shaming perpetrators and imposing costs to actors behind threat campaigns. A key example is the recent coordinated response to the NotPetya campaign (a malware campaign disrupting business operations in Ukraine and across multiple other states) by several Western states, naming and shaming perpetrators and retaliatory measures such as economic sanctions (Marsh 2018). It is too early to say whether these measures will have any measurable effect on cyber operations attributed to Russia, but raising the potential costs of hostile cyber operations adds an important variable to the decision process behind their use.

Fourth, stronger regulation of commercial spyware products. Ideally this would include both procurement legislation to constrain availability of these tools, reporting on government requests received as well as transparency reporting on the impact in order to constrain their deployment (Micek and Aydin 2017). Reporting requirements and initiatives for the oil and mining industries already provide a useful blueprint for effective ways to address this issue (HRW n.d.; Global Reporting Initiative 2016). Initiatives by the UK Government (HM Government 2016), the European Parliament ("The European Parliament Is Fighting to Strengthen the Rules for Surveillance Trade" 2017) and several US municipalities (Buttar 2016; ACLU n.d.) in this area are important first steps towards stronger regulation.[16]

Finally, promoting and facilitating contact and collaboration between the infosec sector and civil society. In an ideal world, the private sector might recognize a moral responsibility towards the weak resulting from their growing political role and power. As private actors develop their capabilities not only brokers of intelligence, but as defense contractors for targeted organizations, they have the resources to protect vulnerable groups. Yet in the real world, these are businesses bound by the need to turn a profit. Still, there is a case to be made for closer collaboration with civil society that would benefit both sides. As existing research highlights, pro-bono services to CSOs is likely to generate a payoff in the form of publishable research, the resulting media attention and potential new clients reached (Crete-Nishihata et al. 2014)—a point that was confirmed by the threat intelligence researcher interviewed for this project (Threat Intelligence Researcher 2018).

Increased attention to threats to CSOs in future reporting will be a welcome change and improve our understanding of the threat landscape they face. However, the data collected for this project provides a basis for further analysis as well. The findings of this exploratory study have largely confirmed expectations derived from independent reporting, while also highlighting the limitations and potential biases in the data derived from industry reporting. To put these findings on a firmer footing, and pave the way towards hypothesis testing, there are three key avenues for further research.

First, expanding the dataset to include the body of independent reporting on targeted threats to CSOs in order to determine specific blind spots and triangulate the actual threat landscape. Coding and analyzing these reports based on the same criteria as industry reports will enable a more detailed comparison of trends and patterns in the data. In particular, it will reveal the extent of specific divergences in geographical focus, targeting and attribution patterns. This analysis in turn will allow (1) determining the specific blind spots in industry reporting, and by taking these into

[16] The author is indebted to Peter Micek for pointing out these existing initiatives.

account, (2) allow triangulating the actual threat landscape faced by CSOs with greater accuracy by aggregating the data from both sources.

Second, extending the analysis to all industry reports to build a database of global threat activity and examine potential variations in these patterns based on different target types. This project has tracked targeting, attribution patterns and geographic distribution of threat activity in industry reports that address threats to CSOs due to the focus of the study and time constraints. Extending the analysis to all reporting, based on the same criteria used here, will provide not only an invaluable resource for future research on targeted threats overall, but also enable determining whether the patterns identified in geographical focus, attribution trends and reported threat activity are uniform across reporting, or differ depending on target type.

Third, case studies tracking the effects of targeted threats on CSOs. How do different threats affect different organizations? What are short, medium and long-term effects? Tracking effects of cyber operation will enable answering the underlying question whether CSOs are better or worse off in face of the new threats brought by the Information Age.[17] A historical case study of the threats faced by a CSO across several decades, and their impacts, would be an important and fascinating case study to further our understanding of the overall impact of the types of threats examined here.

[17] I am indebted to Jon Lindsay for this suggestion.

References

"30% of Ukrainian SBU Officers Were Russian FSB and GRU Agents -." 2014. Euromaidan Press. April 24, 2014. http://euromaidanpress.com/2014/04/24/30-of-ukrainian-sbu-officers-were-russian-fsb-and-gru-agents/.

Accenture. 2017. "COST OF CYBER CRIME STUDY." https://www.accenture.com/t20170926T072837Z__w__/us-en/_acnmedia/PDF-61/Accenture-2017-CostCyberCrimeStudy.pdf.

AccessNow. 2012. "Global Civil Society at Risk: An Overview of Some of the Major Cyber Threats Facing Civil Society." https://s3.amazonaws.com/access.3cdn.net/49632318adb472e369_yhm6ibn8c.pdf.

———. 2017. "The European Parliament Is Fighting to Strengthen the Rules for Surveillance Trade." *Access Now* (blog). December 8, 2017. https://www.accessnow.org/european-parliament-fighting-strengthen-rules-surveillance-trade/.

ACLU. 2017. "New Bill Holds NYPD Accountable for Surveillance Technology." American Civil Liberties Union. March 1, 2017. https://www.aclu.org/news/new-bill-holds-nypd-accountable-surveillance-technology.

Betz, David., and Tim. Stevens. 2011. *Cyberspace and the State : Toward a Strategy for Cyber-Power*. Abingdon: Routledge.

Blazevic, Neil. 2018. "RE: Introductions," January 12, 2018.

Brantly, Aaron F. 2014. "Cyber Actions by State Actors: Motivation and Utility." *International Journal of Intelligence and CounterIntelligence* 27 (3): 465–84. https://doi.org/10.1080/08850607.2014.900291.

Buttar, Shahid. 2016. "A California County Breaks New Ground for Surveillance Transparency." Electronic Frontier Foundation. June 15, 2016. https://www.eff.org/deeplinks/2016/06/california-county-breaks-new-ground-surveillance-transparency.

CA ISBU. 2010. "In-Depth Analysis of Hydraq."

Castells, Manuel. 2008. "The New Public Sphere: Global Civil Society, Communication Networks, and Global Governance." *The Annals of the American Academy of Political and Social Science* 616: 78–93.

Checkpoint. 2017. "Global Threat Impact Index 2017."

CLTC. 2018. "Center for Long-Term Cybersecurity Project on Protecting Politically Vulnerable Organizations Threat Landscape and Organizational Ecosystem." UC Berkeley.

Creswell, John W. 2014. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*.

Crete-Nishihata, Masashi, Jakub Dalek, Ronald Deibert, Seth Hardy, Katharine Kleemola, Irene Poetranto, John Scott-Railton, et al. 2014. "Communities at Risk - Executive Summary." Citizen Lab. https://targetedthreats.net/media/1-ExecutiveSummary.pdf.

Crete-Nishihata, Masashi, Jakub Dalek, Etienne Maynier, and John Scott-Railton. 2018. "Spying on a Budget: Inside a Phishing Operation with Targets in the Tibetan Community." The Citizen Lab. January 30, 2018. https://citizenlab.ca/2018/01/spying-on-a-budget-inside-a-phishing-operation-with-targets-in-the-tibetan-community/.

CyberX. 2017. "Operation BugDrop: CyberX Discovers Large-Scale Cyber-Reconnaissance Operation Targeting Ukrainian Organizations." *CyberX* (blog). February 15, 2017. https://cyberx-labs.com/en/blog/operation-bugdrop-cyberx-discovers-large-scale-cyber-reconnaissance-operation/.

Deibert, Ronald. 2013. *Black Code : Inside the Battle for Cyberspace*. Toronto: Signal.

———. 2015a. "The Geopolitics of Cyberspace After Snowden." *Global Trends*.

———. 2015b. "Cyberspace Under Siege." *Journal of Democracy* 26 (3): 64–78. https://doi.org/10.1353/jod.2015.0051.

Deibert, Ronald, John Palfrey, Rafal Rohozinski, and Jonathan Zittrain, eds. 2008. *Access Denied: The Practice and Policy of Global Internet Filtering*. The Information Revolution and Global Politics. Cambridge, Mass: MIT Press.

Deibert, Ronald, John Palfrey, Rafal Rohozinski, Jonathan Zittrain, and OpenNet Initiative, eds. 2010. *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace*. Information Revolution and Global Politics. Cambridge, Mass: MIT Press.

Deibert, Ronald, John Palfrey, Jonathan Zittrain, and Rafal Rohozinski, eds. 2012. *Access Contested: Security, Identity, and Resistance in Asian Cyberspace Information Revolution and Global Politics*. Information Revolution and Global Politics. Cambridge, MA: MIT Press.

Deibert, Ronald, and Rafal Rohozinski. 2009. "Tracking GhostNet: Investigating a Cyber Espionage Network." *The Citizen Lab* (blog). March 28, 2009. https://citizenlab.org/2009/03/tracking-ghostnet-investigating-a-cyber-espionage-network/.

ESET. 2016. "Operation Groundbait." https://www.welivesecurity.com/wp-content/uploads/2016/05/Operation-Groundbait.pdf.

FireEye. 2013. "SUPPLY CHAIN ANALYSIS: From Quartermaster to SunshopFireEye." https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-malware-supply-chain.pdf.

———. 2014. "Operation GreedyWonk." https://www.fireeye.com/blog/threat-research/2014/02/operation-greedywonk-multiple-economic-and-foreign-policy-sites-compromised-serving-up-flash-zero-day-exploit.html.

———. 2016a. “ONE UNITED DEFENSE AGAINST CYBER ATTACKS.” https://www.fireeye.com/content/dam/fireeye-www/global/en/company/pdfs/fireeye-advanced-threat-protection.pdf.

———. 2016b. “Redline Drawn.”

Freedom House. 2018. “Methodology: Freedom in the World 2018.” 2018. https://freedomhouse.org/report/methodology-freedom-world-2018.

Galeotti, Mark. 2016. “PUTIN’S HYDRA: INSIDE RUSSIA’S INTELLIGENCE SERVICES.” European Council on Foreign Relations. http://www.ecfr.eu/page/-/ECFR_169_-_PUTINS_HYDRA_INSIDE_THE_RUSSIAN_INTELLIGENCE_SERVICES_1513.pdf.

George, Alexander L., and Andrew. Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, Mass.: MIT Press.

Global Reporting Initiative. 2016. “SHINING A LIGHT ON HUMAN RIGHTS.” https://www.globalreporting.org/resourcelibrary/Shining%20a%20Light%20on%20Human%20Rights%202016.pdf.

Goodin, Dan. 2017. “Massive Equifax Hack Reportedly Started 4 Months before It Was Detected.” Ars Technica. September 21, 2017. https://arstechnica.com/information-technology/2017/09/massive-equifax-hack-reportedly-started-4-months-before-it-was-detected/.

Guerrero-Saade, Julian Andres. 2015. “THE ETHICS AND PERILS OF APT RESEARCH: AN UNEXPECTED TRANSITION INTO INTELLIGENCE BROKERAGE.” https://media.kaspersky.com/pdf/Guerrero-Saade-VB2015.pdf.

Haight, Laura. 2015. “NonProfit Hacks: Too Small to Be Hacked? Not!” *Portfolio* (blog). 2015. https://www.portfoliosc.com/blog/2015/7/2/nonprofit-hacks-most-at-risk-least-prepared.

HM Government. 2016. “Good Business Implementing the UN Guiding Principles on Business and Human Rights.”

HRW. n.d. “Oil, Mining, and Natural Resources.” Human Rights Watch. Accessed March 30, 2018. https://www.hrw.org/topic/business/oil-mining-and-natural-resources.

Hulcoop, Adam, Matt Brooks, Etienne Maynier, John Scott-Railton, and Masashi Crete-Nishihata. 2016. “It’s Parliamentary: KeyBoy and the Targeting of the Tibetan Community.” https://citizenlab.ca/2016/11/parliament-keyboy/.

Jay Jay. 2017. “38% of Small Businesses Spend next to Nothing on Cyber Security.” TEISS. October 5, 2017. https://teiss.co.uk/information-security/small-businesses-cyber-security/.

Kello, Lucas. 2013. “The Meaning of the Cyber Revolution: Perils to Theory and Statecraft.” *International Security* 38 (2): 7–40.

Kramer, Franklin D. 2009. “Cyberpower and National Security: Policy Recommendations for a Strategic Framework.” In *Cyberpower and National Security*, 3–25. Washington, D.C.: Potomac Books.

Kurtz, George. 2013. “CrowdStrike Falcon Unveiled: The Power of The Platform ».” February 25, 2013. https://www.crowdstrike.com/blog/crowdstrike-falcon-unveiled-power-platform/.

Kuzio, Taras. 2010. “The FSB Returns to Ukraine.” Jamestown. 2010. https://jamestown.org/program/the-fsb-returns-to-ukraine/.

Le Blond, Stevens, Adina Uritesc, and Cedric Gilbert. 2014. “A Look at Targeted Attacks Through the Lense of an NGO.” In *Proceedings of the 23rd USENIX Security Symposium*. San Diego.

Libicki, Martin C. 2009. *Cyberdeterrence and Cyberwar*. Santa Monica, CA: RAND.

Lynn III, William J. 2010. “Defending a New Domain.” *Foreign Affairs*, October 2010. http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain.

Marczak, Bill, Geoffrey Alexander, Sarah McKune, John Scott-Railton, and Ron Deibert. 2017. “Champing at the Cyberbit: Ethiopian Dissidents Targeted with New Commercial Spyware.” https://citizenlab.ca/2017/12/champing-cyberbit-ethiopian-dissidents-targeted-commercial-spyware/.

Marczak, BIll, Claudio Guarneri, John Scott-Railton, and Morgan Marquis-Bore. 2014. “Mapping Hacking Team’s ‘Untraceable’ Spyware.” *The Citizen Lab* (blog). February 17, 2014. https://citizenlab.org/2014/02/mapping-hacking-teams-untraceable-spyware/.

Marczak, Bill, and John Scott-Railton. 2016. “The Million Dollar Dissident: NSO Group’s IPhone Zero-Days Used against a UAE Human Rights Defender.” *The Citizen Lab* (blog). August 24, 2016. https://citizenlab.org/2016/08/million-dollar-dissident-iphone-zero-day-nso-group-uae/.

Marczak, Bill, John Scott-Railton, and Morgan Marquis-Bore. 2014. “When Governments Hack Opponents: A Look at Actors and Technology.” In . San Diego.

Marsh, Sarah. 2018. “US Joins UK in Blaming Russia for NotPetya Cyber-Attack.” The Guardian. February 15, 2018. http://www.theguardian.com/technology/2018/feb/15/uk-blames-russia-notpetya-cyber-attack-ukraine.

McLean, Asha. 2017. “Australian SMEs Consider Antivirus Software Sufficient Defence: MYOB.” *ZDNet*, September 6, 2017. http://www.zdnet.com/article/australian-smes-consider-antivirus-software-sufficient-defence-myob/.

Micek, Peter, and Deniz Duru Aydin. 2017. “Non-Financial Disclosures in the Tech Sector: Furthering the Trend.” In *The Responsibilities of Online Service Providers*, 241–61. Law, Governance and Technology Series. Springer, Cham. https://doi.org/10.1007/978-3-319-47852-4_13.

Miller, Matthew, Jon Brickey, and Gregory Conti. 2012. “Why Your Intuition About Cyber Warfare Is Probably Wrong | Small Wars Journal.” http://smallwarsjournal.com/jrnl/art/why-your-intuition-about-cyber-warfare-is-probably-wrong.

NATO. 2016. “Press Conference by NATO Secretary General Jens Stoltenberg Following the North Atlantic Council Meeting at the Level of NATO Defence Ministers.” NATO. June 17, 2016. http://www.nato.int/cps/en/natohq/opinions_132349.htm.

Nye, Joseph S. 2010. “Cyber Power.” DTIC Document.

Nye, Joseph S. 2011. “Nuclear Lessons for Cyber Security?” *Strategic Studies Quarterly* 5 (4): 18–38.

Perritt, Henry H. Jr. 1997. “Internet as a Threat to Sovereignty - Thoughts on the Internet’s Role in Strengthening National and Global Governance, The.” *Indiana Journal of Global Legal Studies* 5: 423.

ProofPoint. 2017. “Human Factor Report.” https://www.proofpoint.com/sites/default/files/pfpt-en-us-human-factor-report-2017.pdf.

Rattray, Gregroy J. 2009. “An Environmental Approach to Understanding Cyberpower.” In *Cyberpower and National Security*, 253–74. Washington, D.C.: Potomac Books.

RSA. 2012. “THE VOHO CAMPAIGN : AN IN DEPTH ANALYSIS.” http://blogsdev.rsa.com/wp-content/uploads/VOHO_WP_FINAL_READY-FOR-Publication-09242012_AC.pdf.

Scott-Railton, John, Bill Marczak, Abdul Razzak Bahr, Masashi Crete-Nishihata, and Ron Deibert. 2017. “Reckless Exploit: Mexican Journalists, Lawyers, and a Child Targeted with NSO Spyware.” https://citizenlab.ca/2017/06/reckless-exploit-mexico-nso/.

Scott-Railton, John, Bill Marczak, Claudio Guarneri, and Masashi Crete-Nishihata. 2017. “Bitter Sweet: Supporters of Mexico’s Soda Tax Targeted With NSO Exploit Links.” https://citizenlab.ca/2017/02/bittersweet-nso-mexico-spyware/.

Scott-Railton, John, Morgan Marquis-Bore, Claudio Guarneri, and Marion Marschalek. 2015. “Packrat: Seven Years of a South American Threat Actor.” *The Citizen Lab* (blog). December 8, 2015. https://citizenlab.org/2015/12/packrat-report/.

SecureWorks. 2013. “Secrets of the Comfoo Masters.” https://www.secureworks.com/research/secrets-of-the-comfoo-masters.

———. 2016a. “Hillary Clinton Email Targeted by Threat Group-4127.” https://www.secureworks.com/research/threat-group-4127-targets-hillary-clinton-presidential-campaign.

———. 2016b. “Threat Group-4127 Targets Google Accounts.”

Siegel, Tara. 2017. “Equifax Says Cyberattack May Have Affected 143 Million in the U.S. - The New York Times,” 2017. https://www.nytimes.com/2017/09/07/business/equifax-cyberattack.html?hp&action=click&pgtype=Homepage&clickSource=story-heading&module=first-column-region&region=top-news&WT.nav=top-news&mtrref=undefined&gwh=71B7526A17E38FDB4C41F47AB08B3254&gwt=pay.

Symantec. 2011. “Trojan.Taidoor.” https://www.symantec.com/content/dam/symantec/docs/security-center/white-papers/trojan-taidoor-12-en.pdf.

———. 2013. “Internet Security Threat Report 2013.” http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v18_2012_21291018.en-us.pdf.

———. 2015. “Attackers Target Both Large and Small Businesses.” https://www.symantec.com/content/dam/symantec/docs/infographics/istr-attackers-strike-large-business-en.pdf.

———. 2016. “Patchwork Cyberespionage Group Expands Targets from Governments to Wide Range of Industries.” http://www.symantec.com/connect/blogs/patchwork-cyberespionage-group-expands-targets-governments-wide-range-industries.

“The Dropping Elephant Actor.” n.d. Securelist - Information about Viruses, Hackers and Spam. Accessed January 22, 2018. https://securelist.com/the-dropping-elephant-actor/75328/.

Threat Intelligence Researcher. 2018Phone.

ThreatConnect. 2016. “Russia Hacks Bellingcat ME17 Investigation.” https://www.threatconnect.com/blog/russia-hacks-bellingcat-mE17-investigation/.

TrendMicro. 2012. “The Taidoor Campaign.” https://www.trendmicro.de/cloud-content/us/pdfs/security-intelligence/white-papers/wp_the_taidoor_campaign.pdf.

———. 2013. “Watering Hole 101 - Threat Encyclopedia - Trend Micro USA.” 2013. https://www.trendmicro.com/vinfo/us/threat-encyclopedia/web-attack/137/watering-hole-101.

———. 2014. “Operation Pawn Storm.”

Tsagourias, Nicholas. 2012. “Cyber Attacks, Self-Defence and the Problem of Attribution.” *Journal of Conflict and Security Law*, krs019.

Volexity. 2016. “PowerDuke: Widespread Post-Election Spear Phishing Campaigns Targeting Think Tanks and NGOs | Volexity.” https://www.volexity.com/blog/2016/11/09/powerduke-post-election-spear-phishing-campaigns-targeting-think-tanks-and-ngos/.

White, Gordon. 1994. “Civil Society, Democratization and Development (I): Clearing the Analytical Ground.” *Democratization* 1 (2): 375–90. https://doi.org/10.1080/13510349408403399.

Yang, Zhilin, Xuehua Wang, and Chenting Su. 2006. "A Review of Research Methodologies in International Business." *International Business Review* 15 (6): 601–17. https://doi.org/10.1016/j.ibusrev.2006.08.003.

Zscaler. 2015. "Chinese Cyber Espionage APT Group Leveraging Recently Leaked Hacking Team Exploits To Target A Financial Services Firm." Cloud Security Solutions | Zscaler. 2015. https://www.zscaler.com/blogs/research/chinese-cyber-espionage-apt-group-leveraging-recently-leaked-hacking-team-exploits-target-financial-services-firm.

Data Flows & National Security: A conceptual framework to assess restrictions on data flows under GATS security exception[18]

Martina Ferracane

Abstract

This paper explores the national security implications of a potential for a WTO dispute on data flow restrictions. It proposes a basic conceptual framework to assess data flows' restrictions under GATS security exception. The paper represents a contribution to the literature because it is the first paper to address systematically the issue of data flows and national security in the context of a GATS dispute and because it provides a unique perspective that looks both at legal and technical arguments.

If a case where to be brought before the WTO dispute settlement, there are certain national security concerns which can be considered essential and imminent, and that therefore the defender might bring up to support its case for invoking the security exception. These are: protection from cyber espionage, protection from cyber attacks on critical infrastructure, and access data in order to prevent terrorist threats. The paper presents both a legal and technical analysis of these three cases in order to assess the relevance of restrictions on data flows under GATS security exception. This

[18] Comments or questions can be sent to martina.ferracane@gmail.com. I would like to thank Columbia University SIPA and the Carnegie Corporation for their generous research support. I am also grateful to Francesco Casotto (CISCO), Lee Tuthill (WTO) and Prof. Anupam Chander (UC Davis) for their valuable comments and suggestions to improve the quality of the paper. In addition, I would like to thank the participants to the GIG-ARTS Conference 2018 in Cardiff, ECPR General Conference 2018 in Hamburg and GigaNet Symposium 2018 in Paris for their comments.

analysis can, more generally, inform the debate on the protection of national security in the digital era.

The paper finds that in the three cases, restrictions on data considered critical for national security might raise the cost of certain attacks. However, the risks would remain pervasive and national security would not be significantly enhanced both legally and technically. As a matter of fact, several studies claim that local processing requirements can have rather a detrimental impact on security. The implementation of good security standards and encryption techniques appears to be a more effective way to ensure a better response to cyber threats. All in all, it will be important to investigate on a case by case basis whether the scope of the measure (sectors and data covered) is considered proportionate and whether the measure in question in practice reduces the exposure of the country to cyber espionage, cyber attacks and terrorist threats.

Introduction

Mr Trump's tariffs on steel and aluminium have brought renewed interest on the security exception in the World Trade Organization (WTO). Under the WTO system, countries are not allowed to introduce new tariffs or to impose other trade restrictions on bound goods and services. However, a country can deviate from its free trade obligations in order to achieve important non-economic objectives such as data privacy, public morals and national security.

The security exception is the widest among the exceptions listed in the WTO texts and has only rarely been invoked by WTO members. This has been partly due to the fact that members did not wish for the exception to be employed as a cover for new protectionist measures. But things could change now as more countries consider how to justify trade restrictions under this exception. Recently, this exception has been invoked by the Russian Federation in a dispute involving transit restrictions that Moscow imposed on Ukraine in January 2016 (Palmer, 2018).[19] More generally, however, several countries are referring to the security exception when imposing new regulatory measures, especially those related to the digital economy. The most recent case is Vietnam's Cybersecurity Law passed in June 2018. The lawmakers of the country have made clear that the new measure is justified under the security exception in WTO texts and other free trade agreements (Nguyen, 2018).

Vietnam's law is an example of a new wave of restrictions affecting digital trade (Ferracane *et al.*, 2018). In their public pronouncements or related laws and regulations, many countries have cited national security as a rationale to restrict digital trade, although it is not yet clear whether, in practice, the restrictions contribute to improved national security (Peng, 2015; Sargsyan, 2016). Examples of measures that create thick digital borders include bans to the use of certain digital products in the public sector,[20] security screening on investment (European Commission, 2017a) and measures requiring data to be kept locally (Ferracane, 2017).

[19] See *Russia — Measures Concerning Traffic in Transit* (DS512) [hereafter *Russia-in Transit].*

[20] For example, in 2012, Australia blocked China's Huawei Technologies from tendering for contracts in the country's $38 billion National Broadband Network due to cyber security concerns (Lu-YueYang, 2012). See also a report by the National

The uncertainty surrounding the digital economy increases the policy space of countries to impose protectionist measures, as it is not always clear whether in practice these measures are needed to protect important non-economic interests or whether there are less trade-restrictive alternatives. When a WTO member considers that a certain measure is actually designed to restrict trade rather than to achieve its stated non-economic objective, including national security, it can challenge the measure under the WTO dispute system. Such a dispute is the topic of this paper. In particular, this paper focuses on disputes related to restrictions on data flows. While some scholars have looked into potential disputes related to cybersecurity threats connected to certain digital goods (Peng, 2015), no papers have yet addressed potential disputes related to restrictions on data flows and national security.

These restrictions can have a serious impact on the capacity of businesses to operate and provide services to their customers, and it is therefore not unlikely that a WTO member might challenge data flows' restrictions as a violation of a member's WTO commitments on services. Already in 2015, the European Commission's Report on Trade and Investment Barriers and Protectionist Trends criticised China for restricting data flows 'on the ostensible grounds of "national security"' going 'beyond essential national security concerns', and generally risking 'imposing unnecessary restrictions on commercial activities' (European Commission, 2016).

This paper therefore explores the national security implications of a potential for a WTO dispute on data flow restrictions. It proposes a basic conceptual framework to assess data flows' restrictions under WTO security exception and, in particular, under the General Agreement on Trade in Services (GATS), which enshrines obligations and disciplines on commercial services that apply to all WTO Members.[21] In doing so, the paper intends to fill a gap in the literature by clarifying how these restrictions could be assessed under the existing WTO language.

Ultimately, the decision of a country to start a dispute on data flows restrictions remains a political one and its impact will go well beyond trade. The deadlock on negotiation of digital trade commitments reflects the uncertainty on whether the current structure of the WTO is well suited to judge on issues of privacy, security and, ultimately, internet governance. The WTO members might want to refrain from bringing claims on digital issues which are not explicitly covered by current WTO language until this uncertainty is settled. The likelihood for such a claim will depend on whether one or more members want to take a political stance on internet governance in a forum that is meant to address trade issues.

This paper remains nevertheless relevant even if a WTO dispute on data flows never arises. On the one hand, an analysis on how data flows restrictions may influence, in practice, the capacity of a country to protect its national security can more generally inform discussions of trade agreements

Development and Reform Commission of China and the Ministry of Finance banning the purchase of certain foreign IT products for selected government procurement including by Apple, Microsoft, Dell and Hewlett-Packard Co (China Digital Times, 2014). See also US Directive requiring agencies to identify Kaspersky-branded products on Federal information systems and provide plans to discontinue use of Kaspersky-branded products (US Homeland Security Department, 2017). See also recent Bill proposed in the US that aims to ban US government agencies from using phones and equipment from Huawei and ZTE (US Congress, 2018).

[21] Members' obligations under GATS apply only to those services sectors for which members have voluntarily assumed obligations by inscribing commitments in their Schedules of Specific Commitments (GATS, 1994).

which are contemplating a language on data flows. On the other, this analysis can be useful for policy-makers that are considering imposing new restrictions on data flows driven by national security concerns. The paper can provide arguments on whether and how these measures can effectively improve the capacity of a country to protect its security interests.

This paper refrains instead from exploring GATS general exceptions (e.g. data privacy or public morals) or discussing which types of restrictions on data flows constitute a restriction on the cross-border provision of services and which can be considered necessary for legitimate policy objectives in a WTO context.[22] Many of the restrictions on data flows in force today are likely to fall under such general exceptions, especially the exception on data privacy. These measures are the subject of a separate analysis which is ongoing. An analysis on the GATS security exception remains nevertheless relevant given that the most sweeping measures restricting data flows today, such as local processing requirements for data in critical infrastructure and public procurement, have a clear national security rationale and could hardly be justified under the general exceptions.

Given that restrictions on data flows are not explicitly prohibited under the WTO, this paper starts by introducing the debate on whether data flows restrictions could be considered a trade barrier in the first place (Section I) and whether they could be challenged under GATS (Section II). Assuming that at least some of these measures can be considered a trade restriction, then it remains to be seen how these measures would be assessed in a potential WTO dispute.

The defendant would likely seek to justify the restriction under one of GATS exceptions, which are presented in Section III. The paper then turns to a detailed analysis of how restrictions on data flows can be assessed under the security exception (Section IV). In particular, Section IV looks at three main reasons why a country might most reasonably impose data flows restrictions under the national security rationale. These are: protection from cyber espionage, protection from a cyber attack on critical infrastructure and access to data to prevent terrorist threats. Finally, Section V concludes and provides food for thought for future research in this area.

I. Restrictions on data flows as a trade barrier

Under the World Trade Organization (WTO), countries commit not to restrict trade on specified goods and services included in national schedules with overarching principles, such as most-favoured-national treatment, that apply to all such trade. Certain exceptions are available when a measure is considered necessary to achieve important non-economic objectives, including data privacy and national security.

This system regulates today over 98 percent of global trade in goods and services, providing also a dispute resolution mechanism aimed at enforcing participants' adherence to their commitments. When a country imposes a measure that one or more WTO members perceive to be violating WTO commitments and obligations, they can use this disputes settlement mechanism.

[22] A first assessment on whether data flows restrictions can be considered necessary for legitimate policy objectives in a WTO context can be found in Crosby (2016).

As of today, there have been no disputes at the WTO specifically related to restrictions on data flows.[23] However, several governments have complained about the costs raised by restrictions on data flows or potential national treatment implications, and DG Trade Commissioner Malmström notably stated that 'restrictions on cross-border data flows inhibit trade of all kinds: digital and non-digital, products and services. We cannot just pretend that this doesn't exist, or that data has nothing to do with global trade' (Malmström, 2016).

Given the increasing importance of data flows for trade and the recent surge in the number of data flows restrictions being implement worldwide (Ferracane, 2017), any WTO member could challenge a measure affecting data flows, arguing that it is unnecessarily restricting trade. Data flows constitute today the lifeblood of trade in services, which in turn support manufacturing and trade in goods (Meltzer, 2013). It is therefore unsurprising that restrictions on data flows have risen to the top of the international trade policy agenda of some important trading partners, especially in countries whose businesses rely heavily on the internet for the provision of goods and services.[24]

Restrictions on cross-border data flows are often referred to as data 'localisation' or 'residency' requirements. These measures raise the cost of conducting business across borders by either *mandating companies to keep data within a certain border* or by *imposing additional requirements for data to be transferred abroad* (Ferracane, 2017).

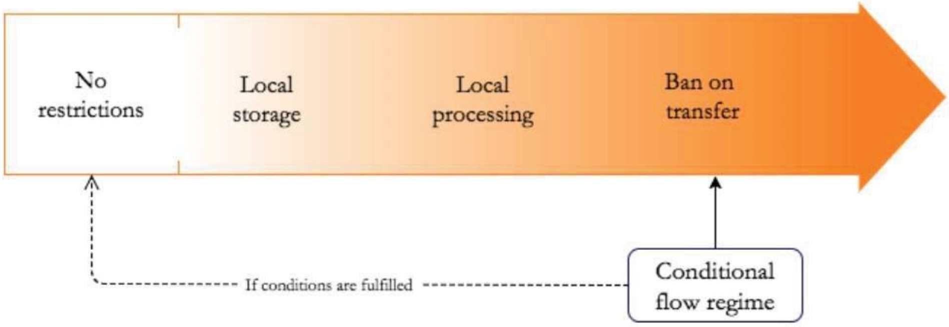
Although these measures share common traits, they can be quite diverse. Four main categories can be identified:

- Ban on the transfer of data abroad (data can never leave the country);
- Local processing requirement (data can leave the country but the main processing has to be done locally);
- Local storage requirement (a copy of the data has to be stored locally); and
- Conditional flow regime (data can travel abroad only under certain conditions, such as consent of the data subject).

Figure 1: Types of restrictions on data flows

[23] There have been three trade disputes that indirectly address data flows. These are *United States-Measures Affecting the Cross-Border supply of Gambling and Betting Services* (DS285) [hereafter *US-Gambling]*, *China-Measures Affecting Trading Rights and Distributions Services for Certain Publications and Audiovisual Entertainment Products* (DS363) [hereafter *China-Publications and Audiovisual Products]* and *China- Certain Measures Affecting Electronic Payment Services* (DS413) [hereafter *China-Electronic Payment Services].* These cases are presented in Section II.

[24] Among others, the U.S. International Trade Commission (USITC) has recently announced new investigations on digital trade, including a report for USTR that 'will assess the rate of adoption of digital technologies in the United States as well as in foreign markets with further study of the importance of both domestic and cross-border data-flows' (USITC, 2017). Two earlier USITC reports had already pointed out the preoccupation for restrictions on data flows (USITC, 2013, 2014). The European Commission has proposed a Regulation on free flow of non-personal data in September 2017 (European Commission, 2017b).



Source: Ferracane (2017).

Conditional flow regimes tend to be imposed under a data privacy rationale and include all those privacy regulations requiring, for example, the consent of the data subject before data leaves the country. Local storage requirements, on the other hand, are often imposed with the objective to facilitate access to certain data for law enforcement (for example in the case of accounting data or metadata). In this case, as long as a copy of the data is kept locally, data can flow freely outside the country.

Stricter measures requiring data to be processed locally or banning any transfer of data altogether are instead more often justified under a national security rationale. These are only a limited share of the restrictions imposed today (Ferracane, 2017), but are likely to create the highest costs for businesses (Ferracane *et al.*, 2018). The analysis in this paper applies mainly to these stricter measures, while a discussion on the GATS general exceptions would more likely focus on conditional flow regimes and local storage requirements.

There are also other measures that affect data when it flows *into the country,* which include blocking and filtering of online content. These measures usually apply to specific websites, online services or political content and have often the objective to censor certain information and maintain public order, or in other cases are meant to protect local companies. Given that these measures target a limited set of actors, they usually do not impact how the overall internet architecture is designed by forcing data to be processed locally, but are rather implemented with targeted actions such as IP blocking. While content blocking and filtering can create serious costs for certain businesses, these measures are different in quality from the restrictions covered in this paper, which have to do with the transfer of data when it flows *outside the country*.

There is little doubt that restrictions on cross-border data flows can constitute a trade restriction, yet they have been explicitly addressed only in a handful of trade agreements (Ferracane, 2016; Burri, 2017). In the absence of an express prohibition to restrict data flows, the current WTO language defining limitations and restrictions remains the most relevant to investigate whether data flows restrictions could be challenged as a trade barrier. In particular, given that restrictions on data flows more directly impact trade in services, the most relevant agreement to analyse in this context is the GATS.

II. GATS and restrictions on data flows

No WTO member has yet instituted proceedings against another for violating the GATS based on restrictions on data flows, and the more general debate over when measures regulating data flows should be considered a violation of the GATS is still in its infancy. The WTO texts say nothing about the internet, censorship, e-commerce or data flows.[25] Despite attempts to include binding language on data flows under the Work Program on Electronic Commerce (WTO, 1998), as of today there are no rules on data flows being negotiated in the WTO. This raises the question of whether, under the existing GATS language, restrictions on data flows constitute a barrier that could be challenged in a trade dispute (Crosby, 2016; Aaronson, 2017; Burri, 2017; Ferracane, 2017).

Despite the lack of legal adaptation of WTO texts, it seems that there is an agreement on the fact that digital trade can be subsumed under the provisions of the GATS (Crosby, 2016; Drake, 2016; Tuthill, 2016; Burri, 2017). This is mainly based on the interpretation of certain WTO rulings and GATS Council documents issued as part of the e-commerce work program (WTO, 1999). Three cases are particularly relevant when looking at restrictions on trade of online services: *US-Gambling*, *China-Publications and Audiovisual Products and China-Electronic Payment Services*.

In *US-Gambling*, the report concluded that cross-border supply of services under the GATS (so-called mode 1)[26] encompasses all possible means of supplying services from the territory of one WTO Member into the territory of another WTO Member. Therefore, a full market access commitment for mode 1 implies the right for other Members' suppliers to supply a service through all means of delivery, including online delivery.[27]

In *China-Publications and Audiovisual Products,* the Panel found that the scope of China's scheduled commitment on 'sound recording distribution services' extends to recordings distributed in non-physical form through technologies such as the internet.[28] These conclusions are in line with the principle of technological neutrality, which seems to be largely shared among WTO Members and was already mentioned in the 1999 Progress Report on the Work Program on Electronic Commerce (WTO, 1999).[29]

Similarly, in *China-Electronic Payment Services,* the Panel reached the conclusion that China's commitments on 'payment and money transmission services' include electronic payments

[25] The GATS Annex on Telecommunications indirectly touches upon the issue of data flows as it specifies that services suppliers are entitled to use public telecommunications for the movement of information within and across borders as well as for cross-border access to information stored in databases.

[26] WTO members have so far not agreed upon a clear determination of whether the electronic cross-border delivery of a service is a service supplied through GATS mode 1 (cross-border) or mode 2 (consumption abroad).

[27] This would be the case unless the member has specified otherwise in its schedule of commitments. *See* Panel Report, *US-Gambling,* paragraph 6287*.*

[28] Panel Report, *China-Publications and Audiovisual Products*, paragraph 71209.

[29] Paragraph 4 of the Progress Report on the e-Commerce working program states that: 'It was also the general view that the GATS is technologically neutral in the sense that it does not contain any provisions that distinguish between the different technological means through which a services may be supplied'.

services.[30] The Panel suggested that electronic payment services are an integral part to certain payment services and therefore, as long as data transfers are an integral part of a service in a committed sector, they are also covered by the commitments.

Measures restricting cross-border data flows could be assessed as a restriction on cross-border supply of both 'traditional' services and computer and related services. For example, a measure requiring to process financial data locally might be seen both as a market access restriction on its operations, as well as a de facto national treatment restriction in the provision of financial services because of the higher costs that the company would incur to process data locally and the fact that it would be disadvantageous to foreign companies that would normally process or store their data abroad.[31] As a restriction on computer and related services, the impact is more direct. Where a government has a full commitment on computer and related services, the data restrictions would prevent an IT service supplier abroad from securing cross border clients in the country. A great number of WTO members have made far-reaching commitments on computer and related services (Renee and Reisman, 2012).

While it has been argued that an analysis of existing WTO texts leads to the conclusion that 'data localization measures violate existing GATS rules and commitments to allow unrestricted cross-border trade in digital services and cross-border data flows' (Crosby, 2016), some research is still needed to confirm that the GATS language as it stands today possesses the basic legal strength to address specifically the concerns on restrictions on data flows. Among others, the E15 group and the G20 have already recommended that the WTO clarify the application of GATS commitments to digital trade and data flows (Meltzer, 2016; IMF, 2016).

This paper assumes that at least some of the restrictions imposed on data flows can be considered a trade restriction. With this assumption, it turns to exploring how these measures could be assessed under GATS security exception.

III. GATS security exception

As noted by Tim Wu, when the GATS entered into force, no one realised that 'almost by accident, the WTO has put itself in an oversight position for most of the national laws and practices that regulate the Internet' (Wu, 2006). Wu also made the point that members of the WTO would have to decide 'how much control is legitimate domestic regulation, and how much is a barrier to trade'. Despite the fact that the agreement has not been updated to take into account new internet issues, it is inevitable that the WTO would find itself in the position to adjudicate when certain internet measures would be justified, or not, under the current exceptions.

As mentioned above, GATS provides a number of specific grounds for maintaining and adopting restrictions on data flows based on legitimate, non-economic policy objectives. These are

[30] Panel Report, *China-Electronic Payment Services*, paragraphs 7180 an 7182.

[31] A ban on transfer of financial data cross-border might also be interpreted as a prohibition to provide financial services under mode 1, as the company would need to have a commercial presence in the country in order to provide its services (so-called mode 3) and could not provide them cross-border.

enumerated in Articles XIV (General Exceptions) and XIV *bis* (Security Exception). In case of a dispute related to restrictions on data flows, the respondent might argue that the measure falls under one of the listed policy objective relevant to taking exceptions. This paper focuses on the security exception, while another analysis will look at the general exceptions.

GATS Art. XIV *bis* is the most relevant article when looking at data flows and services. The provision (b)(iii) of GATS security exception states that nothing in the agreement should be construed to 'prevent any Member from taking any action which it considers necessary for the protection of its security interests (...) taken in time of war or other emergency in international relations'. A defendant invoking this exception could justify restrictions on movement of data by claiming that it is taking actions 'it considers necessary' to protect its 'essential security interests' in the context of an 'emergency in international relations' caused, for example, by threats of cyber espionage or a cyber attack that could destabilise the country.

While the WTO jurisprudence has provided a certain degree of clarity when it comes to the interpretation of GATS general exceptions,[32] the same is not true in regards to the security exception.[33] The article has been invoked only rarely in trade disputes and the WTO judiciary has consistently avoided issuing findings on these merits. Therefore, the degree of uncertainty on the interpretation of this clause remains significant. This situation might change with the case *Russia-in Transit* in which the Panel is requested for the first time to rule on a defence based on the security exception (in this case the exception in Article XXI of GATT) (European Union, 2017).

No reliable statistical data is available as to the unilateral application of the provision given that Members were not obliged to notify the invocation of national security measures under GATT. There are four cases with reference to Art. XXI GATT that reached the panel level.[34] These cases have recently been mentioned in the context of consultations in the dispute DS526 *United Arab Emirates — Measures Relating to Trade in Goods and Services, and Trade-Related Aspects of Intellectual Property Rights* (WTO, 2017). The small body of relevant jurisprudence on the security exception reveals that the practice of the WTO is still inconclusive on the issue of national security (Peng, 2015). In addition, while certain scholars have looked into interpretation of restrictions on data flows under certain GATS general exceptions (Kobrin, 2004; Future of Privacy Forum, 2013; Burri and Schär, 2016; Aaronson, 2017; Burri, 2017; Kuner, 2017), there is no analysis available today on data flows restrictions being taken under GATS security exception.

[32] The Appellate Body in *US — Gambling* elaborated on the similarities between Article XX of the GATT 1994 and Article XIV and stated that the article sets out general exceptions under the GATS much in the same way as Article XX of the GATT 1994 does under the GATT. The Appellate Body also found previous decisions under Article XX of the GATT 1994 relevant for the analysis under Article XIV.

[33] Yet Delimatsis & Cottier (2008) argue that, given the semantic similarity between the GATT and GATS articles, interpretations and case law under Art. XXI GATT are relevant and useful when interpreting Art. XIV *bis*.

[34] *See* GATT Panel Report, *United States – Restrictions on Exports to Czechoslovakia*, GATT Doc CP.3/SR22 (8 June 1949) ('US – Export Restrictions (Czechoslovakia)'); GATT Panel Report, *United States – Imports of Sugar from Nicaragua*, GATT Doc L/5607 (2 March 1984); GATT Panel Report, *United States – Trade Measures Affecting Nicaragua*, GATT Doc L/6053 (13 October 1986, unadopted) ('US – Trade Measures Affecting Nicaragua'); *Trade Measures Taken by the European Community against the Socialist Federal Republic of Yugoslavia*, GATT Doc L/6948 (2 December 1991) (Communication from the European Communities).

The security exception differs from the general exceptions in two important ways. First, the measures do not explicitly prohibit arbitrary or unjustifiable discrimination. Second, in order to invoke the exception, a Member only needs to 'consider' that its security interests are endangered, so it could appear like a Member can self-interpret its own security interests (Carr, 2001; Ayres and Mitchell, 2012). Thus, the role of the Panels and the Appellate Body is more circumscribed, and it has been argued that 'the security exceptions might not be judiciable at all' (Ayres and Mitchell, 2012; Schill and Briese, 2009; USTR, 2018).

However, recent trends in thought in regards to the interpretation of the security exceptions support the idea that these provisions are subject to review by the WTO dispute settlement mechanism. Among others, scholars have argued that the doctrine of 'unfettered discretion deserves a critical assessment' (Delimatsis & Cottier, 2008) and that there is nothing in the wording of the security exception that suggests that 'the performance of the conditions in the subparagraphs is totally self-judging' (Peng, 2015; *see also* European Union, 2018).

A similar point has been raised by the GATT panel in *United States – Trade Measures Affecting Nicaragua,* although the final report has not been adopted. While stating that its mandate precluded it from examining the motivations of the US in raising the security exception, the panel also posed the question in regards to GATT security exception in Article XXI. It states that 'If it were accepted that the interpretation of Article XXI was reserved entirely to the contracting party invoking it, how could the CONTRACTING PARTIES ensure that this general exception to all obligations under the General Agreement is not invoked excessively or for purposes other than those set out in this provision?'.[35]

Therefore, it is likely that there is certain role for the WTO dispute mechanism in determining the validity of the measure, although the margin of appreciation accorded to the State remains wider than under GATS general exceptions in Art. XIV (Delimatsis & Cottier, 2008). The limits are to be found in the doctrine of abuse of rights as the principle of good faith under Article 31 of the Vienna Convention on the Law of Treaties is expected to inform any provision of the WTO agreements.[36] Thus far, a body of academic literature has strongly argued that this principle is an appropriate standard applicable to security exceptions as well (Hahn, 1991; Carr, 2001; Bonnan, 2010; Peng, 2015; Yoo, 2016).

Therefore, while the necessity test can only be applied in regards to the general exceptions of Art. XIV, the WTO judiciary has the jurisdiction to assess whether the right of invocation has been abused and can review whether the enhancement of security is 'manifestly absent' based on the

[35] The panel also adds: 'If the CONTRACTING PARTIES give a panel the task of examining a case involving an Article XXI invocation without authorizing it to examine the justification of that invocation, do they limit the adversely affected contracting party's right to have its complaint investigated in accordance with Article XXIII:2? Are the powers of the CONTRACTING PARTIES under Article XXIII:2 sufficient to provide redress to contracting parties subjected to a two-way embargo?'. Notice that the report has not been adopted.

[36] Article 31 of the Vienna Convention on the Law of Treaties states that '[a] treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose'. This is also in line with one early report on this matter by the GATT panel in *United States – Restrictions on Exports to Czechoslovakia* in 1949. In the report the panel states that 'every country must be the judge in the last resort on questions relating to its own security. On the other hand, every Contracting Party should be cautious not to take any step which might have the effect of undermining the General Agreement'.

definition of security as intended by the member (Delimatsis & Cottier, 2008). Some authors have claimed that focusing on the jurisprudence on the test of least trade-restrictiveness, instead of the objectives of each different measure, could be an appropriate guideline (Yoo, 2016). The WTO judiciary would therefore need to assess whether the respondent has added sufficient evidence showing that the measure remains within the broad bounds of necessity for the protection of its essential security interests and that the specific measure does not intentionally serve protectionist purposes (Delimatsis & Cottier, 2008).

The word 'essential' indicates that general security is not sufficient and therefore the 'security interests should meet a higher standard that can be distinguished from other normal security interests' (Peng, 2015). The instances pertaining to the security exception should be different in quality than those related to 'public order' that fall instead under the general exceptions. Important is also that the threat to the essential security interest is credible and especially imminent (Delimatsis & Cottier, 2008). This is clear from the reference to a 'war or other emergency in international relations'. Measures that are implemented on a long-term basis, for example the establishment of a national-only cloud, could hardly be justified under a temporary condition of emergency.

In this context, there is therefore a certain margin of discretion for the WTO judiciary as well, although it is likely that the definition of security interest would be left to the discretion of the country, while the panel would be left to judge the relevance of the measures imposed to achieve the government's objective.

IV. Data flows and national security

It is not unusual to hear political statements like 'cybersecurity risks pose some of the most serious economic and national security challenges of the XXI century' (Obama, 2015) and 'without cybersecurity there is no national security' (Bandurski, 2017).[37] However, despite the frequent reference to security concerns as a justification for the imposition of restrictions on data flows, governments have not provided clear arguments on how in practice certain restrictions would respond to these security concerns.

Therefore, an analysis of the security implications of these measures requires to start with an assumption on the possible reasons why a government would consider that keeping data within its borders is a matter of national security. The question of what is a serious concern of national security eventually boils down to the threat scenario of the country and what each country perceives as a national security issue. Certain countries, for example, might consider the threat of economic interests to be a national security issue, while others might not.

The fact that, in the digital era, access to the Internet is necessary for the ordinary functioning of a country cannot be *per se* a justification for a requirement to process all data locally. Forcing companies to process data locally does not ensure that the network would still be fully functional

[37] This words appear in a banner across the top of the official website of the Cyberspace Administration of China together with the image of President Xi Jinping.

in case of an attack nor does it make the communication system more resilient in case of an attack. A weak security system remains weak no matter where the data is stored. To ensure the more general stability and resilience of the network, the country could rather focus on a detailed analysis of the risk scenario specific to the country in order to identify idiosyncratic vulnerabilities and to secure the internet infrastructure. Important points are, among others, the necessity to build internet exchange points and a recovery system. Such a general claim therefore would hardly be brought into a dispute.

There are however other concerns that a country could bring up in a dispute to defend data restrictions under the security exception. These measures should address essential security concerns and respond to a situation of emergency. There are in particular three instances that are likely to concern the majority of countries today in relation to the movement of data and that could likely be brought up in a dispute. These are:

1. Cyber espionage: restricting the flow of data from leaving the country makes it harder for other countries to surveil certain communications considered of national security interest;
2. Cyber attacks on critical infrastructure: by restricting flow of data, the critical infrastructure is better protected from and more resilient to cyber attacks;
3. Terrorist threats: keeping data locally improves the capacity of a government to conduct surveillance at home with the objective to identify possible threats and prevent terrorist attacks.

In this section of the paper, these three cases are taken as a basis for both a legal and technical analysis on the impact of data flows restrictions on the ability of a country to protect its national security. The analysis will focus on two types of restrictions: local processing requirements and bans to transfer data. This is because, as mentioned above, the other two categories of data flows restrictions (local storage requirement – that is keeping only a local copy of the data - and conditional flow regimes) would hardly be implemented under the national security rationale. Examples of local processing requirements recently implemented include China's Cybersecurity Law in China and Russia's latest amendments to its Data Protection Law.[38]

The question to address is how in practice data flows' restrictions can contribute to protect the country's national security and whether it can be expected that they, in fact, enhance the protection of the country's essential security interests. Given that from a technical perspective a certain level of risk is an intrinsic feature of the digital era (O'Harrow, 2014), it is not expected that these measures should eliminate the risk of cyber espionage or cyber attacks altogether in order to be considered consistent with the GATS framework. However, it would be crucial for the country to prove that the measure is somewhat contributing to reducing cyber risks connected to essential security interests and that it does not intentionally serve protectionist purposes.

IV. a. *Cyber espionage*

Cyber espionage occurs in secret. This implies that there is a lack of public knowledge on the issue. Nevertheless, the Snowden revelations have sparked interest among the general public and have

[38] Standing Committee of the National People's Congress, Cybersecurity Law, 7 November 2016; Federal Law no. 152-FZ "On Personal Data" (OPD-Law) as amended in July 2014 by Federal Law No. 242-FZ "On Amendments to Certain Legislative Acts of the Russian Federation for Clarification of Personal Data Processing in Information and Telecommunications Networks".

enhanced the public's knowledge on government surveillance and the techniques used to conduct it.

When it comes to a potential WTO dispute, a country might argue that certain restrictions on data flows, and in particular a ban to transfer certain information abroad or the requirement to process data locally, are legitimate under GATS security exception as they prevent espionage and therefore protect national security. This section looks into this hypothesis presenting legal and technical arguments.

First of all, we need to clarify the main means through which cyber espionage takes place. Cyber espionage exploits vulnerabilities in each of the three layers of cyberspace (physical layer, logical layer and social layer).[39]

In the physical layer, a certain mechanism (so-called 'back-door')[40] can be built in the hardware, including during the manufacturing process, that would subsequently permit remote access. Hardware products with back-doors could be computers but also some parts of the internet infrastructure such as routers and switches (Zetter, 2013). Alternatively, data can be captured when it is transmitted over communications cables. In this case, the wiretap can happen in the country where data is stored or in one of the countries through which data is transiting.

Data can also be redirected through specific countries for tapping purposes (so-called 'BGP hijacking'). In this case, the attack would happen at the logical layer as it affects an internet protocol. BGP is a protocol used to determine the most efficient way to route data between independently operated networks.[41] This protocol can be hijacked (intentionally or unintentionally) and result in a manipulation of Internet routing paths. When intentional, BGP hijacking can allow data to be intercepted (or in some cases manipulated). While to date there is no known evidence of BGP hijackers successfully decrypting rerouted traffic (Goodin, 2017), such attacks enable bulk collection of data that can be stored for future decryption. These attacks remain however quite complicated and they usually require the support of an ISP or any organization controlling an autonomous system. In addition, there are some technological solutions available that might make these attacks harder in the future. An example is the recent progress in BGPSec Protocol Specification (Timberg, 2015; Siddiqui, 2017).

[39] This is only one of the possible classifications of internet layers and it is based on Schmitt (2017). The Tallinn Manual 2.0 defines these three layers as follows. The physical layer is defined as the 'physical network components (i.e., hardware and other infrastructure, such as cables, routers, servers, and computers)'. The logical layer 'consists of the connections that exist between network devices' and 'it includes applications, data, and protocols that allow the exchange of data across the physical layer'. Finally, the social layer 'encompasses individuals and groups engaged in cyber activities'.

[40] Back-doors are also common at the logical layer. In the latter case, a certain vulnerability is intentionally created in order to enable access to communications.

[41] For a detailed explanation of BGP and BGP hijacking, *see* Julian (2017).

Vulnerabilities[42] in the logical layer are more commonly exploited by malware ('malicious + software') designed to monitor communications, but also by human actors that conduct targeted attacks in order to intrude a system. In these cases, attacks are often conducted remotely, although there are cases in which software modifications can also be placed by physically accessing a device. An example of malware is the Remote Access Trojan (RAT). RATs are very common and designed to provide the attacker with complete control over the victim's system. They can be used to steal sensitive information, to spy on victims, and remotely control infected computers (Shamir, 2015). These attacks can target systems in computers, but also in components of the internet infrastructure such as routers (Zetter, 2013).

Finally, social engineering techniques such as phishing and spear-phishing[43] can be employed in the social layer with the objective to gather access credentials to facilitate seemingly authorised access to information that possesses intelligence value. These attacks consist mainly in sending emails that appear to come from a reputable source with the objective to access sensitive information such as passwords.

The question is whether, by restricting data flows and keeping data locally, the country would be less exposed to cyber espionage. The answer has two faces: a legal one and a technical one.

From a customary international law perspective, espionage is not prohibited *per se* (Schmitt, 2017). The Tallinn Manual 2.0 provides certain arguments for which local processing of data might provide certain safeguards to the implementing country. In fact, a tapping operation in the territorial or archipelagic waters of another State is considered by the experts of the Tallinn Manual as a violation of that State's sovereignty, while the operation is considered acceptable in the State waters ('without prejudice to the application of other international legal norms', Rule 54). Therefore, in the extent to which a restriction on data flows makes it less likely for data to pass through communications cables outside the country, data is legally better safeguarded when stored locally.

In addition, certain experts have argued that close access cyber espionage operations, such as the insertion of a USB flash drive into a computer located on one State's territory, might constitute a violation of sovereignty (Rule 32). Therefore, it seems that certain means of espionage would be unlawful if data remains stored and processed locally. However, all other means that enable cyber espionage remotely, such as phishing and malwares, remain legally and technologically available and therefore the risk for cyber espionage is likely to remain substantial.

While certain states have enacted domestic legislation that criminalises cyber espionage carried out against them, a number of States have by domestic law authorised their security services to engage in espionage, including cyber espionage.[44] According to certain national rules on cyber

[42] Vulnerabilities are weaknesses resulting from 'chinks in the armor of [a] code, where the system does not behave precisely as designed'. These can be exploited to access communications or manipulate the software (Singer and Friedman, 2014, at 42).

[43] Phishing is a generally exploratory attack that targets a broader audience, while spear phishing is a targeted version of phishing.

[44] Schmitt (2017) cites as examples *Lag om signalspaning i försvarsunderrättelseverksamhet* (2008:717), Secs. 1–2 (Sweden); *BND-Gesetz* (20 December 1990), Sec. 2(1)(40) (Germany); *Wet op de inlichtingen – en veiligheidsdiensten* (WIV) (7 February

espionage, the legal hurdle of government agencies for intercepting data is lower when data is kept outside their territory. In fact, these laws provide fewer legal protections to data that is stored abroad than when it is stored or processed in the country. This is the case of the United States where the interception of communications overseas has looser restrictions and less oversight than when data is stored domestically.

The case of the United States is especially relevant given that much of the world's electronic communications passes through the country. It is therefore worth clarifying the country's legal regime. While large-scale collection of Internet content would be illegal in the United States, when the operations take place overseas, the NSA is allowed to presume that anyone using a foreign data link is a foreigner and therefore statutory restrictions on surveillance seldom apply. Therefore, if a country is imposing a restriction on data flows with the objective to avoid cyber espionage by the US, the measure could paradoxically have the opposite effect of (legally) facilitating cyber espionage activities by the US authorities.

Surveillance from American soil is regulated by the Foreign Intelligence Surveillance Act (FISA), which prohibits interception within the United States if there is any possibility the communications might include United States 'persons'[45].[46] When it is clear that the data is of non-US citizens, the Fourth Amendment protections do not apply and the government does not have to apply 'even the minimal requirement of reasonableness' (Daskal, 2015, at 341). However, given the practical difficulty to know in advance if all parties of a communication over the internet are foreigners outside the United States, in practice collection of data or communications from a wire or switch inside the United States requires a court order by the Foreign Intelligence Surveillance Court (FISC or FISA court) (Edgar, 2017, at 35).

By contrast, the NSA's surveillance outside the American soil (including satellite surveillance) is governed by Executive Order 12,333 of 1981 and it does not require a court review.[47] The Order only requires the NSA and other intelligence agencies to have rules that protect the privacy of US persons. Therefore, data would usually be harder to access by US agencies when it is stored in the US than when it is kept outside the US soil.

2002), Arts. 6.2.d, 7.2.a.1°, 7.2.e, 27(1) (Netherlands); *Regulation of Investigatory Powers Act* (RIPA) (2000), Sec. 8(4) (United Kingdom); *Bundesgesetz über die Zuständigkeiten im Bereich des zivilen Nachrichtendienstes* (3 October 2008), Art. 1.a (Switzerland); *Bundesgesetz über Aufgaben und Befugnisse im Rahmen der militärischen Landesverteidigung (Militärbefugnisgesetz – MBG)* (27 April 1999), Sec. 20(a) (Austria).

[45] A 'United States person' is defined in the NSA's website as: (1) a citizen of the United States; (2) an alien lawfully admitted for permanent residence; (3) an unincorporated association with a substantial number of members who are citizens of the U.S. or are aliens lawfully admitted for permanent residence; or (4). a corporation that is incorporated in the U.S. See NSA website FAQs at the link https://www.nsa.gov/about/faqs/sigint-faqs.shtml#

sigint4

[46] Foreign Intelligence Surveillance Act (FISA) of 1978, 50 U.S. Code § 1881b - Certain acquisitions inside the United States targeting United States persons outside the United States.

[47] Executive Order 12333—United States Intelligence Activities, US Federal Register, December 4, 1981.

Although legal restrictions on collection and access to data (also metadata)[48] are heightened within the United States, concerns on surveillance of data which passes through the United States remain significant and are connected to the potential abuse of Section 702 of FISA Amendment Act of 2008 (FAA).[49] While the section was not a secret, the Snowden leaks revealed the existence of two programs authorized by this section: Prism (also known as 'downstream collection') and 'upstream collection'.

Under Prism, the NSA collected from domestic companies the phone calls, emails, texts and other electronic messages of foreigners abroad without a warrant.[50] Upstream collection, on the other hand, gave the NSA access to data in transit across the internet backbone facilities of American telecommunication companies. These programs leveraged on the fact that data was physically passing through the United States and, although it appears that after the Snowden revelations these practices have been limited (Edgar, 2017; The White House, 2014), they created an incentive for governments to control how data moves and whether it passes through the United States.

From a technological perspective, the arguments are different. It might be technically easier for a government agency to access data when it is passing through their national network as they have easier physical access to the cables, routers, data centers or other components of the network infrastructure. Also, in these cases, it would be harder to find proofs of the snooping as evidence lays in the hardware that is physically hacked. Strict restrictions on data flows with border controls might therefore limit this type of attacks and might also make it harder to redirect internet traffic to a certain network through BGP hijacking.

Besides these cases in which restrictions on data flows might increase the cost of cyber espionage, there are several other options to conduct surveillance activities that would remain open even when data is processed locally. The most common techniques such as malware, phishing activities or hardware hacking would still be available.[51] In some cases, even unplugging from the internet would not work. For example, documents leaked by Edward Snowden have allegedly revealed that the NSA spied on foreign intelligence targets, including the Chinese and Russian militaries, by inserting tiny circuit boards or USB cards into 100,000 computers and using radio waves to transmit data therefrom without the computers having to be connected to the Internet (BBC, 2014).

[48] Domestic bulk collection of metadata was permitted under Section 215 of the FISA (as amended by the 2001 Patriot Act). Yet, this practice was declared unlawful in 2015 by a US federal appeals court. *See American Civil Liberties Union v. Clapper*, 785 F. 3d 787 (2d Cir. 2015); *see also* Gellman and Soltani (2013).

[49] The Protect America Act of 2007, Pub.L. 110–55, 121 Stat. 552, enacted by S. 1927, was a legislative forerunner to Section 702 of FISA and it was a temporary measure that was set to expire 180 days after its enactment. It was reauthorized with additional safeguards and the introduction of Section 702 of Foreign Intelligence Surveillance Act (FISA) with the FISA Amendments Act of 2008, H.R. 6304, enacted October 7, 2008. Under this section, the FISA court can authorize surveillance of 'persons reasonably believed to be outside the United States' without individual review of targets. This section was reauthorized in 2012 and again in January 2018 with a six-year extension.

[50] In certain cases, the NSA received also support to get direct access to encrypted messages (Greenwald *et al.,* 2013).

[51] It is worth mentioning that hardware hacking might be more complicated when data is processed locally because there would be a better way to inspect the product against tampering. However, back-doors and other vulnerabilities might still be exploited when not detected in the inspection.

In summary, restrictions on data considered critical for national security might raise the cost of certain attempts of cyber espionage, especially in cases in which companies are legally compelled to hand over to the government any data passing through the country. However, the risk of cyber espionage remains pervasive even when data is processed locally as already today cyber espionage is generally conducted remotely (Schmitt, 2017, at 247). It will be important to investigate on a case by case basis whether the scope of the measure (sectors and data covered) is considered proportionate and whether the measure in question in practice reduces the exposure of the country to cyber espionage.

In this context it is relevant to mention that only a complete ban to transfer data abroad (or very strict restrictions on data export) could have a certain impact in raising the cost of cyber espionage while any local processing requirement in practice would hardly make any difference for the capacity of the country to protect itself from cyber espionage. In this latter case, in fact, a copy of the data can still leave the country and therefore being subject to cyber espionage. As an example, the recent amendments of data privacy law in Russia require data operators to ensure that the recording, systematisation, accumulation, storage, update, amendment and retrieval of personal data of the citizens of Russia is made using databases located in Russia.[52] The Russian government has framed these measures as a necessary security measure in the light of the Snowden revelations (Sargsyan, 2016), but it could hardly sustain this argument in a dispute given that a copy of the data can still freely leave the country.

Finally, in the context of a dispute, it will be relevant to investigate whether there are alternative technological solutions available that can guarantee the same level of protection from cyber espionage while having less restrictive effects on trade. Certain solutions in this regard have recently been developed and rely on keeping encryption keys with the customer (among others, Brodkin, 2015).

IV. b. *Cyber attack on critical infrastructure*

While the issue of cyber espionage has mainly to do with confidentiality of data, the main concerns connected with cyber attacks on critical infrastructure are integrity and availability of information systems.[53] In these cases, the main goal of certain restrictions on data flows would be to prevent data from being improperly altered or changed without authorization as well as to make sure that the system remains available and can be used as usually expected.

There is no single definition of critical infrastructure as countries have different views on what is truly critical for their functioning. Systems which can be considered critical from a national security perspective include, among others, electricity, energy, telecommunication, water supply,

[52] Federal Law no. 152-FZ “On Personal Data” (OPD-Law) as amended in July 2014 by Federal Law No. 242-FZ “On Amendments to Certain Legislative Acts of the Russian Federation for Clarification of Personal Data Processing in Information and Telecommunications Networks”.

[53] Confidentiality, integrity and availability are sometimes referred to as the ‘CIA triad’ of cyber security and are considered to be the main goals of security in an information environment.

financial services, and security services (police, military).[54] These sectors rely heavily (or will soon rely heavily) on the internet, and are only likely to become more exposed to cyber threats in the future with advancements connected to the Internet of Things (IoT) and 5G technology (Lee-Makiyama, 2018). A severe cyber attack to these services is likely to have not only a significant economic impact but also national security implications.

The new Cybersecurity Law in China, which entered into force in June 2017, can be considered an example of regulation applied to critical infrastructure which restricts flows of the data on the basis of the national security. The law includes requirements for ‘important data’ collected by ‘key information infrastructure operators’ (KIIOs) to be kept within the borders of China. If there are business needs for the KIIOs to transfer this data outside of China, security assessments must be conducted (Ferracane and Lee-Makiyama, 2017).

In this case, the arguments to investigate are merely technical. In fact, the legal discussion of what constitutes a cyber attack on critical infrastructure that would endanger national security is not affected by the location in which the data is processed. Tampering with a state’s network and critical infrastructure is in fact always likely to be considered a violation of national sovereignty (Schmitt, 2017, at 170; 312; 339-356).[55] Therefore, the question to ask in this case is whether an obligation to process data locally makes critical systems or sectors more resilient in case of a cyber attack.

To answer this question, it is relevant to investigate how cyber attacks on critical infrastructure can occur. The tools are somewhat similar to the case of cyber espionage, although in this case the objective is to gain access to the system in order to sabotage it or control it, rather than simply gathering certain communications.

First of all, the attackers might prey on systems that have ignored basic precautions, such as products that have default login names and passwords that the user has not changed (Singer and Friedman, 2012, at 39-45). In other cases, attackers could exploit software vulnerabilities either known or unknown (the latter are usually referred to as ‘zero days’). An attacker could also use malware to take advantage of a certain vulnerability to infiltrate the system. Malware can be transferred through infected USB sticks but they can also spread through the internet and be downloaded into a computer simply by visiting a website, clicking on a link or opening an attachment.

A recent example of malware that impacted critical infrastructure is NotPetya. This malware, which leveraged on the modified version of two NSA's stolen and leaked exploits, took down many

[54] See wikipedia entry on ‘critical infrastructure’ for an overview of different programs in place with the objective to protect critical infrastructure, https://en.wikipedia.org/wiki/Critical_

infrastructure, accessed on 20 December 2017.

[55] This is the view of the experts of the Tallinn Manual 2.0. Rule 66 states that ‘A State may not intervene, including by cyber means, in the internal or external affairs of another State’. Accidental loss of functionality of infrastructure as a result of cyber espionage is also considered a violation of sovereignty. The discussion on what constitutes use of force presented in the Tallinn Manual 2.0 might be relevant in this context.

Ukrainian government agencies, the Ukraine National Bank, its transportation services and largest power companies in June 2017 (Thomson, 2017). Only a month earlier, Wannacry ransomware[56] had affected operations of hundreds organisations, including delaying surgeries and disrupting operations of the UK National Health System (Graham, 2017).

A malware can also allow a single actor to control a network of private computers as a group without the owners' knowledge. These are referred to as 'botnets' and can be used for activities such as sending spam messages. Botnets can have an impact on the availability of critical infrastructure when they are used to conduct 'distributed denial of service' (DDoS) attacks. These consist in the flooding of a server, website or other network resource with data sent simultaneously from many individual computers so that the system is eventually forced to slow down or even crash and shut down. An example of a DDoS attack on critical infrastructure happened in Estonia in 2007. The attack swamped websites of several organizations, including the Estonian parliament, banks, ministries, newspapers and broadcasters (among others, Haeley, 2013).

There are several protections a company or a public authority can put in place in order to protect their activity from a cyber attack. Defensive measures include technological solutions (such as automated updates, hardware-based security, encrypted data, and multifactor authentication), innovations in operations (user education and awareness through training and certificates, creation of Computer Emergency Response Teams (CERTs), and cyber kill chain) and public policies (data breach notification laws, international coordination and conventions, education curricula, and liability rules).[57]

The National Institute of Standards and Technology (NIST), which is the American agency that promotes and maintains standards in several areas including cyber security, released in early 2017 an updated *Framework for Improving Critical Infrastructure Cybersecurity* (NIST, 2017). The framework, which was recently promoted to a compulsory requirement for risk management of critical infrastructure in the US,[58] suggests cyber security practices and protocols including encryption techniques for data security both for data-at-rest and data-in-transit.[59]

The extreme level of defence to avoid attacks would be to have a physical separation between the network and critical systems (so-called 'air-gapped' system) (Singer and Friedman, 2014, at 63). This provides companies with an extra-level of security as they are significantly less exposed to

[56] A ransomware is a specific type of malware that "prevents or limits users from accessing their system, either by locking the system's screen or by locking the users' files unless a ransom is paid" (MicroTrends, 2017).

[57] The list is based on a broader analysis presented in New York Cyber Task Force (2017).

[58] Executive Order 13800, *Presidential Executive Order on Strengthening the Cybersecurity of Federal Networks and Critical Infrastructure*, 11 May 2017.

[59] PwC (2016, at 4) found that organisations adopting the NIST Cybersecurity Framework greatly improve their ability to identify risks as well as detect and mitigate threats.

threats coming from the internet. However, this comes at a substantial cost in terms of efficiency and effectiveness, and it does not always work in practice (*Ibid.)*.[60]

Restrictions on data flows such as requirements to process data locally or bans to transfer data abroad would hardly have any impact on the security of the company or public authority as long as their network is connected to the internet. Implementing restrictions on data flows might increase the cost of the attack in certain cases, but the risk remains substantial. As a matter of fact, several studies point at the opposite direction and claim that local processing requirements can have rather a detrimental impact on security (Chander and Lê, 2015; Mauer *et al*., 2014).

This is because the security advantage brought by the distributed nature of cloud solutions would be lost.[61] Keeping data locally might actually reduce the resilience of the system if data is not distributed over several data centers and therefore is lost in case of a targeted attack. In addition, companies would be forced to use local providers of data processing services rather than international providers that might have more resources to implement high security standards. Lower security would also facilitate cyber attacks coming from within the country, with similarly detrimental consequences on national security.

The implementation of good security standards and encryption techniques appears to be a more effective way to lead to improvements of the resilience of critical infrastructure and ensure a better response to cyber threats.

IV. c. *Terrorist threats*

Another issue that has climbed to the top of the national security agenda is governments' access to data to identify terrorist threats and prevent (offline) terrorist attacks as well as similar incidences which can destabilize the country. In this case, the rationale for imposing restrictions on data flows would be to enable the government to conduct better surveillance within its country in order to identify terrorist activities. This topic is connected to the broader theme of access to data by government authorities for law enforcement.[62] However, only instances connected to emergencies and security threats that could destabilize the country are likely to qualify as a genuine national security issue that could be justified under GATS security exception. Other issues on law

[60] The authors mention that there have been several vulnerability assessments by the National Security and Communications Integration Center that have confirmed that such attempts of complete separation from the firm's other computer enterprise network fail most of the times.

[61] Amoroso (2014) writes that 'when deployed properly, the cloud provides several critical security advantages over perimeter-based models including greater automation, self-tailoring, and self-healing characteristics of virtualized security' and also adds that in these way security teams can 'decouple security software from hardware and deliver on-demand protection rapidly and flexibly via APIs'.

[62] Relevant in this context is the *United States v. Microsoft Corp.* case, also known as the "Microsoft Ireland" case, that is a pending data privacy case currently being heard by the Supreme Court of the United States. The case involves the extraterritoriality of law enforcement seeking electronic data under the 1986 Stored Communications Act, Title II of the Electronic Communications Privacy Act of 1986 (ECPA), in light of modern computing and Internet technologies such as data centers and cloud storage. *See* 'In the Matter of a Warrant to Search a Certain E-Mail Account Controlled and Maintained by Microsoft Corporation', Case 14-2985, United States Court of Appeals FOR THE SECOND CIRCUIT.

enforcement could fall more generally under the public order rationale and therefore the analysis would relate to the GATS general exception.

The question is: would restrictions on data flows contribute to better investigate terrorist threats and prevent attacks?

From a legal perspective, governments have repeatedly pointed out the inadequacy of the current regime for requesting access to data stored abroad with the objective to investigate terrorist threats. The current mechanism to obtain evidence - including communications data - across borders is a web of Mutual Legal Assistance Treaties (MLATs). These regimes are used more broadly for law enforcement purposes, including terrorists investigations. Their wide reach has created serious inefficiencies with the system practically failing to provide timely access to data stored abroad and requests taking up to several months or even years to be handled (Kent, 2015).

As noted in a recent report by the President's Review Group on Intelligence and Communications Technologies, foreign countries seeking information from the United States through an MLAT request 'face a frustrating delay in conducting legitimate investigations. These delays provide a rationale for new laws that require e-mail and other records to be held in the other country, thus contributing to the harmful trend of localization laws' (Clarke *et al.*, 2013). This is why there are several initiatives, including by the Council of Europe, the International Working Group on Data Protection in Telecommunications, and the European Commission, which are currently developing frameworks for transborder access to electronic evidence.

The fact that there are these initiatives ongoing that could offer a solution to the inefficiencies of the current system and the strong concerns on the possible implications for human rights connected to government's access to communications of its citizens offer room to assess that local processing requirements would not be considered a proportionate means to achieve the objective of preventing terrorist attacks. First of all, in order to access metadata and content of communications, it would be enough to request that only a copy of the data is kept locally (local storage requirement). Instead, requesting the local processing of data or imposing a ban to the transfer of data abroad would increase considerably trade costs while not making it any easier for the authorities to gain access to communications data.

Second, the fact that a copy of the data is stored locally would only support investigation of terrorist threats which originate within the country. All potentially relevant communications taking place outside the country would still remain outside the jurisdiction of the country.

Third, the fact that the data is stored locally would not translate in unconstrained access to such data, but the intelligence agencies and other authorities would still find several additional legal constraints in those countries that have imposed legal requirements to balance data privacy with national security. There is already a heated debate in several countries on where it stands the line between legitimate access to data and violation of citizens' privacy. Certain governments have openly denounced the internet and internet companies as responsible for providing terrorism 'the safe space it needs to breed',[63] while at the same time concerns have been raised over how

[63] Theresa May in a speech given in June 2017 (Parker, 2017).

governments could abuse their access to data stored locally, for example to identify political dissidents (Plaum, 2014; Sargsyan, 2016).

The invalidation of the European Data Retention Directive in 2014 is relevant in this context. The directive, more formally "Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC" required member states had to store citizens' telecommunications data for a minimum of 6 months and at most 24 months. The European Court of Justice (ECJ) ruled the directive invalid and it considered that the directive 'entails a wide-ranging and particularly serious interference with the fundamental rights to the respect for private life and to the protection of personal data, without that interference being limited to what is strictly necessary'.[64]

Two recent regulatory texts offer an example of possible government response to the inefficiency of the MLAT process. One is China's Counter Terrorism Law.[65] The second is US Cloud Act.[66] Both these laws have been designed to facilitate access to data by public authorities with the objective to prevent terrorist threats, among other objectives. Both laws have been criticized for different reasons, but only the Chinese law requires local storage as a necessary requirement to investigate terrorist threats.

In case of a dispute, the defendant would need to prove how this measure is contributing to the country's national security objective despite the arguments presented above and despite the fact that there is little evidence that any particular terrorist plot that would have been carried out if government surveillance had not taken place (Savage, 2013).

V. Conclusions

No WTO member has yet instituted proceedings against another for imposing restrictions on data flows and the debate over whether these restrictions could or should be considered a violation of the GATS commitments is still in its infancy. Although it has been argued that restrictions on data flows, in practice, may constitute a violation of GATS commitments, further research is still needed to confirm whether the existing language in GATS possesses the basic juridical strength to address restrictions on data flows.

If a certain restriction of data flows were challenged under the WTO framework as a violation of the country's trade commitments, the defendant might argue that the measure is imposed for national security reasons and is therefore justified under GATS security exception. If that were to

[64] JUDGMENT OF THE COURT (Grand Chamber), 8 April 2014, in Joined Cases C-293/12 and C-594/12, http://curia.europa.eu/juris/liste.jsf?num=C-293/12

[65] Standing Committee of the National People's Congress, *Counterterrorism Law of the People's Republic of China*, Order No. 36 of the President, issued in December 2015.

[66] US Congress, *Clarifying Lawful Overseas Use of Data Act or CLOUD Act*, H.R. 4943, 115th Congress (2017-2018) of the United States, 6 February 2018.

happen, the WTO judiciary would need to assess the self-judging nature of the security exception and whether the right of invocation of the exception was abused under the overarching good-faith principle which underpins all WTO agreements.

Ultimately, it is likely to remain under the discretion of the member to identify what is considered to be a national security interest. Yet, the panel has still a role to play to assess the relevance of the measure to achieve the government's objective. A good-faith review would need to consider whether security interests are manifestly absent and whether the measure is not intentionally serving protectionist interests.

Given the technicality of the issue, the panel should assess whether the measure actually makes a contribution to reducing cyber risks connected to national security (as defined by the member), which might entail finding at least a minimum degree of proportionality between the protection of national security interests and the overall impact on trade resulting from the measure.

The decision to challenge restrictions on data flows under the WTO remains ultimately a political one. In a moment of deadlock of the trade negotiations on digital issues, the WTO members have to face the question of whether the current structure of the multilateral system is equipped and entitled to deal with sensitive issues such as privacy and security online, and eventually have a say on the direction in which the internet will develop.

This paper presents both legal and technical arguments that can be relevant to assess restrictions on data flows under GATS security exception and, more generally, can inform the debate on how to protect national security in the digital age. If a case where to be brought before the dispute settlement, there are certain national security concerns which can be considered essential and imminent, and that therefore the defender might bring up to support its case for invoking the security exception. As noted above, these are combatting: cyber espionage, cyber attacks on critical infrastructure, and terrorism.

In relation to the first case of cyber espionage, the legal hurdle of government agencies for intercepting and accessing sensitive data of another country might increase when data is kept within the country. This is because several national laws provide fewer legal protections to data that is stored abroad than when it is in the country. At the same time, according to soft law, certain means of espionage would be unlawful only if data remains stored and processed locally. In addition, interception and access to data might be technologically simpler when the data leaves the home country and especially when it passes through the country that aims at intercepting the communications. Detecting espionage might also be harder when it is conducted in another country's telecommunication infrastructure.

Yet, several means for espionage remain technologically available to the country that wants to conduct cyber espionage and the risk of cyber espionage remains substantial despite local processing of data. It will therefore be important to investigate on a case by case basis whether the scope of the measure (sectors and data covered) is considered proportionate and whether the measure in question in practice reduces the exposure of the country to cyber espionage. As an example, the Chinese data processing requirement under which user data needs to be processed locally in order to obtain a license for online taxis companies could hardly be defended in a WTO context under the security exception.[67]

In relation to cyber attacks that might destabilise national security, this paper finds that restrictions on data flows such as requirements to process data locally or bans to transfer data abroad would hardly have any impact on the security of the company or public authority as long as their systems remain connected to the internet. While the cost of certain attacks might increase in certain cases, the risks remain substantial. As a matter of fact, local processing requirements can have instead a detrimental impact on security and could also facilitate cyber attacks coming from within the country. The implementation of good security practices and encryption techniques appears to be a more effective way to improve the resilience of the critical infrastructure and to ensure a better response to cyber threats.

The third case analysed in this paper concerns government's access to data with the objective to identify terrorist threats and prevent attacks. Although the current MLATs mechanism for accessing data located abroad is in need of reform, there is little evidence that restrictions on data flows would have any impact on the capacity of a country to prevent terrorist attacks. Measures that keep data within the country's border would only support investigation of terrorist threats which originate within the country, and the intelligence agencies would still need to face several legal hurdles which are put in place to balance privacy with security.

Overall, the assessment of restrictions on data flows under GATS security exceptions is complicated by several factors. First of all, the border between what in an 'essential' security interest and what is not is not clear. Second, it is not always clear to what extent a particular restriction would strengthen the capacity of the country to respond to a cyber threat. If the risk of cyber espionage or cyber attack is diminished to a very limited extent and yet the measure creates major disruption of trade flows, a WTO panel might not consider the measure appropriate or effective in relation to the defendant's national security interests, even if the standard of proportionality is not, in itself, the determining argument.

In addition, the complaining party could argue that there are other solutions, such as security standards and screenings of the internet infrastructure, that can achieve the same results in terms of supporting national security while being less trade-restrictive. While most countries would agree that dramatic measures might be needed in time of war or cyberattack, the fact that only a limited number of countries seem to consider strict data or internet restrictions implemented in a pervasive and routine manner as a means of ensuring national security might be relevant in this context. Indeed, most of world economies are not currently considering strict restrictions on data flows as an approach to protecting national security interests.

The weighing and balancing of different factors should also take this into account the fact that no measure is expected to lead to a situation of zero-risk, which is practically impossible in a cyber security context. Moreover, restrictions on data flows can impact countries in different ways depending on their legal environment, other cyber security policies in place, and their telecom

[67] *Interim Regulations for the Management of Network Appoint Taxi Services Operations*, Article 27, 1 November 2016. *See* Livingston and Greenleaf (2016).

infrastructure, among other issues. Ultimately, each measure would need to be assessed on a case by case basis according to each country's threat scenario.

References

Aaronson, S., "What are We Talking About When We Discuss Digital Protectionism?", Working Paper for the Economic Research Institute of Asia (ERIA), July 2017, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3032108.

Amoroso, E., *"The New Security Architecture, Dark Reading",* 20 November 2014, https://www.darkreading.com/compliance/the-new-security-architecture-/d/d-id/899845.

Ayres, G. and A. D. Mitchell, "General and Security Exceptions under the GATT and the GATS", in Indira Carr, Jahid Bhuiyan & Shawkat Alam (eds), International Trade Law and WTO, Federation Press, 2012.

Bandurski, D., "Xi Jinping's Web of Laws", China Media Project, 3 May 2017, http://chinamediaproject.org/2017/05/03/xi-jinpings-web-laws/, accessed on 30 December 2017.

BBC, "NSA could 'spy on offline computers', says latest leak", 15 January 2014.

Bonnan, R. "The GATT Security Exception in a Dispute Resolution Context: Necessity or Incompatibility?", XIX(1) Currents International Trade Law Journal 3, 2010.

Brenner, J., "Glass Houses: Privacy, Secrecy and Cyber Insecurity in a Transparent World", Pinguin Books, 2013.

Brodkin*, J., "Box hands cloud encryption keys over to its customers"*, ArsTechnica, 10 February 2015, https://arstechnica.com/information-technology/2015/02/box-hands-cloud-encryption-keys-over-to-its-customers/, accessed on 18 December 2017.

Burri, M., "The Regulation of Data Flows Through Trade Agreements", Georgetown Journal of International Law, Vol. 48, No. 1, 2017, University of Lucerne, 28 August 2017.

Burri, M. and R. Schär*, "*The Reform of the EU Data Protection Framework: Outlining Key Changes and Assessing Their Fitness for a Data-Driven Economy", 6 J. INFO. POL'Y 479, 2016.

Carr W.A. Jr, "Creating Standards and Accountability for the Use of the WTO Security Exception: Reducing the Role of Power-Based Relations and Establishing a new Balance between Sovereignty and Multilateralism", 26, Yale Journal of International Law 413, 2001.

Chander, A. and U. Lê, "Data Nationalism", Emory Law, Volume 64, Issue 3, 5 March 2015.

China Digital Times, "Apple Excluded from Official Procurement", 7 August 2014, https://chinadigitaltimes.net/2014/08/apple-excluded-chinese-government-procurement/, (accessed on 6 April 2018).

Clarke, R.A., M.J. Morell, G.R. Stone, C. R. Sunstein and P. Swire, “Liberty and Security in a Changing World, Report and Recommendation of the President’s Review Group on Intelligence and Communications Technologies”, at 227, 12 December 2013.

Crosby, D. “Analysis of Data Localization Measures Under WTO Services Trade Rules and Commitments”, E15 Initiative Policy Brief, E15Initiative, International Centre for Trade and Sustainable Development and The World Economic Forum, March 2016.

Daskal, J. “The Un-Territoriality of Data”, 125 Yale Law Journal 326, March 2015, p.341.

Drake, W. J., “BACKGROUND PAPER for the workshop on Data Localization and Barriers to Transborder Data Flows”, The World Economic Forum, Geneva, September 2016.

Edgar, T.H., “Beyond Snowden: Privacy, Mass Surveillance, and the Struggle to Reform the NSA”, Brookings Institution Press, Washington, DC, 2017.

European Commission, “REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT on Trade and Investment Barriers and Protectionist Trends”, COM(2016) 406 final, p. 11, Brussels, 20 June 2016, http://trade.ec.europa.eu/doclib/docs/2016/june/tradoc_154665.pdf.

European Commission, “Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Welcoming Foreign Direct Investment while Protecting Essential Interests”, COM(2017) 494 final, Brussels, 13 September 2017.

European Commission, “Proposal for a Regulation of the European Parliament and of the Council on a framework for the free flow of non-personal data in the European Union”, COM(2017)495, Brussels, 19 September 2017.

European Union, “RUSSIA — MEASURES CONCERNING TRAFFIC IN TRANSIT (DS512)”, Third Party Oral Statement by the European Union, Ref. Ares(2017)5434182, Geneva, 8 November 2017, http://trade.ec.europa.eu/doclib/docs/2018/february/tradoc_156602.pdf

Ferracane, M.F., “After TPP: the making up of trade rules for data flows”, Borderlex, PRO Monthly Trade Briefing, April 2016, http://borderlex.eu/wp-content/uploads/2016/07/2016-04-BORDERLEX-PRO-MONTHLY.pdf.

Ferracane, M.F., “Restrictions on Cross-Border Data Flows: a Taxonomy”, ECIPE Working Paper No. 1/2017, European Center for International Political Economy (ECIPE), Brussels, November 2017.

Ferracane, M.F., and H. Lee-Makiyama, “China’s technology protectionism and its non-negotiable rationales”, ECIPE Trade Working Paper, European Center for International Political Economy (ECIPE), Brussels, June 2017.

Ferracane, M.F., H. Lee-Makiyama and Van der Marel, E., “Digital Trade Restrictiveness Index Report”, European Center for International Political Economy (ECIPE), Brussels, April 2018, http://ecipe.org/app/uploads/2018/04/DTRI-final1.pdf

Future of Privacy Forum, “Overextended: Jurisdiction and Applicable Law Under the EU General Data Protection Regulation”, January 2013, https://fpf.org/wp-content/uploads/FINAL-Future-of-Privacy-Forum-White-Paper-on-Jurisdiction-and-Applicable-Law-January-20134.pdf .

Gellman, B. and A. Soltani, “NSA infiltrates links to Yahoo, Google data centers worldwide, Snowden documents say”, The Washington Post, October 13, 2013, https://www.washingtonpost.com/world/national-security/nsa-infiltrates-links-to-yahoo-google-data-centers-worldwide-snowden-documents-say/2013/10/30/e51d661e-4166-11e3-8b74-d89d714ca4dd_story.html?utm_term=.4d49f8f4049d.

General Agreement on Trade in Services (GATS), Marrakesh Agreement Establishing the World Trade Organization, Annex 1B, 1869 U.N.T.S. 183, 33 I.L.M. 1167 (1994), 15 April 1994.

Goodin, D. *““Suspicious” event routes traffic for big-name sites through Russia”*, ArsTechnica, 13 December 2017, https://arstechnica.com/information-technology/2017/12/suspicious-event-routes-traffic-for-big-name-sites-through-russia/, (accessed on 27 December 2017).

Graham, C., “NHS cyber attack: Everything you need to know about 'biggest ransomware' offensive in history”, The Telegraph, 20 May 2017, http://www.telegraph.co.uk/news/2017/05/13/nhs-cyber-attack-everything-need-know-biggest-ransomware-offensive/.

Greenwald, G., E. MacAskill, L. Poitras, S. Ackerman and D. Rushe, “Revealed: How Microsoft handed the NSA access to encrypted messages”, The Guardian, July 11, 2013.

Hahn, M.J., “Vital Interests and the Law of GATT: An Analysis of GATT’s Security Exception”, 12 Michigan Journal of International Law 558, 1991.

Healey, J., “A Fierce Domain: Conflict in Cyberspace, 1986 to 2012”, Cyber Conflict Studies Association (CCSA), June 2013.

IMF, “Reinvigorating Trade to Support Growth: A Path Forward”, Note for Ministers and Governors for the July G-20 Ministerial Prepared by IMF Staff, September, p. 5, 2016, http://www.g20.org/English/Documents/Current/201608/P020160815370397652241.pdf.

Kent, G., “The Mutual Legal Assistance Problem Explained”, Blog Post at the Center for Internet and Society website, Stanford Law School, 23 February 2015, http://cyberlaw.stanford.edu/blog/2015/02/mutual-legal-assistance-problem-explained.

Kuner, C., “Reality and Illusion in EU Data Transfer Regulation Post Schrems”, 18 GERMAN L.J. 881, 2017.

Kobrin, S.J., “Safe Harbours Are Hard to Find: The Trans-Atlantic Data Privacy Dispute, Territorial Jurisdiction and Global Governance”, 30 REV. INT’L STUD. 111, 2004.

Julian, Z., "An Overview of BGP Hijacking", Blog post on Bishop Fox Website, 17 August 2015, https://www.bishopfox.com/blog/2015/08/an-overview-of-bgp-hijacking/, (accessed on 12 December 2017).

Lee-Makiyama, H., "Stealing Thunder: Cloud, IoT and 5G will Change the Strategic Paradigm for Protecting European Commercial Interests. Will Cyber Espionage be Allowed to Hold Europe Back in the Global Race for Industrial Competitiveness?", ECIPE Occasional Paper No. 2/2018, European Center for International Political Economy (ECIPE), Brussels, February 2018, http://ecipe.org/publications/stealing-thunder/

Livingston, S. and G. Greenleaf, "Data Localisation In China and other APEC Jurisdictions", (2016) 143 Privacy Laws & Business International Report, 22-26, October 2016.

Lu-YueYang, M., "Australia blocks China's Huawei from broadband tender", Reuters, 26 March 2012, https://www.reuters.com/article/us-australia-huawei-nbn/australia-blocks-chinas-huawei-from-broadband-tender-idUSBRE82P0GA20120326, accessed 6 April 2018.

Malmström, C., "Trade in a Digital World", Speech at the Conference on Digital Trade, European Parliament, 17 November 2016, Brussels, http://trade.ec.europa.eu/doclib/docs/2016/november/tradoc_155094.pdf.

Mauer, T., R. Morgus and I. Skierka, "The Anti-Surveillance Strategies That Could Ruin the Internet", TIME, 10 December 2014, http://time.com/3628212/strategies-ruin-internet/ (accessed on 2 September 2018).

Meltzer, J.P., "The Internet, Cross-Border Data Flows and International Trade", Issues in Technology Innovation, No. 22, Center for Technology Innovation at Brookings, Washington DC, February 2013.

Meltzer, J.P., "Maximizing the Opportunities of the Internet for International Trade", on behalf of the E15 Expert Group on the Digital Economy, p. 20, 2016.

MicroTrends, " Definition: Ransomware", https://www.trendmicro.com/vinfo/us/security/definition/ransomware, accessed on 15 December 2017.

New York Cyber Task Force, "Building a Defensible Cyberspace", School of International Public Affairs (SIPA), Columbia University, 18 September 2017, https://sipa.columbia.edu/sites/default/files/3668_SIPA%20Defensible%20Cyberspace-WEB.PDF.

Nguyen, M., "Vietnam lawmakers approve cyber law clamping down on tech firms, dissent", Reuters, 12 June 2018, https://www.reuters.com/article/us-vietnam-socialmedia/vietnam-lawmakers-approve-cyber-law-clamping-down-on-tech-firms-dissent-idUSKBN1J80AE.

National Institute of Standards and Technology (NIST), "Framework for Improving Critical Infrastructure Cybersecurity, Draft Version 1.1", 10 January 2017, https://www.nist.gov/sites/default/files/documents////draft-cybersecurity-framework-v1.11.pdf.

Obama, B., "Fact Sheet: Enhancing and Strengthening the Federal Government's Cybersecurity", online by Gerhard Peters and John T. Woolley, The American Presidency Project, 17June 2015,. https://obamawhitehouse.archives.gov/blog/2015/06/17/fact-sheet-enhancing-and-strengthening-federal-government-s-cybersecurity

O'Harrow, R. Jr., “Zero Day: the Threat in Cyberspace”. New York: Washington Post e-book, 2014.

Palmer, D., “US sides with Russia in WTO national security case against Ukraine”, Politico, July 30, 2018, https://www.politico.eu/article/us-sides-with-russia-in-wto-national-security-case-against-ukraine/

Panos D. and T. Cottier, “Article XIV *bis* GATS: Security Exceptions”, 2008, https://www.researchgate.net/publication/228121387_Article_XIV_bis_GATS_Security_Exceptions.

Parker, G., “Theresa May warns tech companies: ‘no safe space’ for extremists”, Financial Times, 4 June, 2017.

Peng, S., “Cybersecurity Threats and the WTO National Security Exceptions”, Journal of International Economic Law, Volume 18, Issue 2, 1 June 2015, Pages 449–478, https://doi.org/10.1093/jiel/jgv025.

Plaum, A., “The impact of forced data localisation on fundamental rights”, Access Now, 4 June 2014, https://www.accessnow.org/the-impact-of-forced-data-localisation-on-fundamental-rights/.

PwC, *“*Turnaround and Transformation in Cybersecurity: Key Findings from the Global State of Information Security Survey 2016”, 8 February 2016.

Renee, B. and M. Reisman, “Policy Challenges of Cross-Border Cloud Computing”, Journal of International Commerce and Economics, US Trade Commission, Web version: May 2012, https://usitc.gov/journals/policy_challenges_of_cross-border_cloud_computing.pdf

Sargsyan, T., “Data Localization and the Role of Infrastructure for Surveillance, Privacy, and Security”, International Journal of Communication 10 (2016), 2221-2237, 2016.

Savage, C., “N.S.A. Said to Search Content of Messages to and From U.S.”, New York Times, August 8, 2013.

Schill, S. and R. Briese., “‘If the State Considers’: The Self-Judging Clauses in International Dispute Settlement”, in A. von Bogdandy et al. (eds), Max Planck Yearbook of United Nationals Law (Martinus Nijhoff Publishers, 2009), Vol.13, 61-140.

Schmitt, M., “Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations”, Cambridge: Cambridge University Press, 2017.

Siddiqui, A., “BGPSec – A reality now”, 16 October 2017, The Internet Society, https://www.internetsociety.org/blog/2017/10/bgpsec-reality-now/.

Singer, P.W. and A. Friedman, "Cybersecurity and Cyberwar: What everyone needs to know", Oxford University Press, 2014.

The White House, "Presidential Policy Directive -- Signals Intelligence Activities" (PPD-28), 17 January 2014, https://obamawhitehouse.archives.gov/the-press-office/2014/01/17/presidential-policy-directive-signals-intelligence-activities.

Thomson, I., "Everything you need to know about the Petya, er, NotPetya nasty trashing PCs worldwide", The Register, 28 Jun 2017, https://www.theregister.co.uk/2017/06/28/petya_notpetya_ransomware/, accessed 15 December 2017.

Timberg, C., "Quick fix for an early Internet problem lives on a quarter-century later*", The* Washington Post, 31 May 2015, accessed on 26 December 2017, http://www.washingtonpost.com/sf/business/2015/05/31/net-of-insecurity-part2/?utm_term=.965aaa36a9b8.

Tuthill, L., "Cross-border data flows: What role for trade rules?", Research Handbook on Trade in Services, 357–382, ISBN: 9781783478057, 30 September 2016.

US Congress, "H. R. 4747, Defending U.S. Government Communications Act", 9 January 2018, 115th Congress, https://www.congress.gov/bill/115th-congress/house-bill/4747/all-info, (accessed 6 April 2018).

US Homeland Security Department, "Binding Operational Directive 17-01, 82 FR 43782", 13 September 2017, https://www.federalregister.gov/documents/2017/09/19/2017-19838/national-protection-and-programs-directorate-notification-of-issuance-of-bindingoperational, (accessed 6 April 2018).

USITC, "Digital Trade in the U.S. and Global Economies, Part I", July 2013.

USITC, "Digital Trade in the U.S. and Global Economies, Part 2", August 2014.

USITC, "Digital Trade and the Impact of Barriers to Digital Trade on the Competitiveness of U.S. Firms in International Markets Will Be Focus of Three New USITC Investigations," News Release 17-024, February 2017.

WTO, "Work Program on Electronic Commerce", adopted by the General Council on 25 September 1998, WT/L/274.

United State Trade Representative (USTR), "Russia – Measures Concerning Traffic in Transit (DS512)", Responses of the United States of America to Questions From the Panel and Russia to Third Parties, 20 February 2018, https://ustr.gov/sites/default/files/enforcement/DS/US.3d.Pty.As.Pnl.and.Rus.Qs.fin.(public).pdf

WTO, "Work Program on Electronic Commerce: Progress Report to the General Council", S/L/74, 27 July 1999.

WTO, “Qatar seeks WTO panel review of UAE measures on goods, services, IP rights”, WTO website, 23 October 2017, https://www.wto.org/english/news_e/news17_e/dsb_23oct17_e.htm.

Wu, T., “The World Trade Law of Censorship and Internet Filtering”, Chicago Journal of International Law, Volume 7, No. 1, Article 12, June 1, 2006.

Yoo, J.Y. and D. Ahn, “Security Exceptions in the WTO System: Bridge or Bottle-Neck for Trade and Security?”, Journal of International Economic Law, Volume 19, Issue 2, 1 June 2016, Pages 417–444, https://doi.org/10.1093/jiel/jgw049

Zetter, Z. “NSA laughs at PCs, prefers hacking routers and switches, Security”, 4 September 2013, www.wired.com/2013/09/nsa-router-hacking/ , (accessed 2 April 2018).

Toward a ‘Digital Silk Road’ Strategy?
Chinese State-Firm Coordination and Technology Policy Developments in Southeast Asia

Shazeda Ahmed

Five years after Chinese President Xi Jinping unveiled his government’s massive “One Belt, One Road” (OBOR, 一路一带 *yilu yidai*) international infrastructure plan, critics now warn that Chinese construction of overseas ports, bridges, highways, and other traditional forms of infrastructure may lead alleged benefactor states to incur substantial debt. Sri Lanka has become the go-to example of an OBOR deal gone wrong, left with no choice but to turn a controlling share of a Chinese-built port over to the partially state-owned China Merchants Port Holdings (Stacey, 2017). While some observers now argue that international enchantment with OBOR has started to fade, the same cannot be said of the parallel development of the project’s powerful digital economic counterpart, the “Digital Silk Road” (DSR, 数字丝绸之路 *shuzi sichou zhi lu*).[68]

Although an official definition of the Digital Silk Road project’s scope has not been issued, government representatives and tech companies alike now regularly affix the phrase to initiatives involving overseas Chinese-built Internet infrastructure, adoption of Chinese technical standards, and the popularization of Chinese online service and content platforms in countries with which China aims to conduct digitally-mediated trade. At present the plans for this information and communications technology (ICT) network comprise most of the Association of Southeast Asian Nations (ASEAN) member states, and a selection of Central Asian and Middle Eastern countries. Within China, experts have even begun to discuss how to extend the DSR to include Latin America (Lou and Yang, 2018) and sub-Saharan Africa (Xin and Yao, 2018).

Discussions about OBOR often highlight projects in countries at a far geographic remove from China, at the expense of capturing the transformative developments between China and its

[68] The Indian Ocean-traversing companion of OBOR, called the “Maritime Silk Road” (海上丝绸之路, *haishang sichou zhi lu*) has been further spun off into the interchangeably used terms “Online Maritime Silk Road” (网上海上丝绸之路, *wangshang haishang sichou zhi lu*), “Information Silk Road” (信息丝绸之路, *xinxi sichou zhi lu*), and “Digital Silk Road” (DSR) or “Digital OBOR”.

neighbors. On-the-ground engagements between Chinese companies, the Chinese government, and their counterparts in ASEAN countries are laying the groundwork for a regional digital economy with China at its center. Whether explicitly branded as part of the DSR or not, these developments will have long-term consequences for digital trade and ICT policies in Southeast Asia. Similar partnerships with China's tech firms and government may serve as a model other developing states seek to pursue.

At this early a stage in the DSR's development, a question with near-term significance is whether or not this strategy represents a high level of coordination between the Chinese government and the country's biggest tech firms. Drawing upon an extensive review of original Mandarin policy documents, tech industry periodicals, and scholarly literature, this paper argues that the Digital Silk Road represents Beijing's efforts to more closely monitor and shape the trajectory of Chinese tech firms' ongoing internationalization, rather than a collaborative initiative involving concerted state-firm cooperation.

Longer-term considerations stem from the nature of the relationship between the state and tech firms in shaping the DSR. How might the DSR influence the development of ICT infrastructure, markets, and policy in the ASEAN region? Do these exchanges nudge certain ASEAN states to enact technology policies mirrored after China's own restrictive model of domestic Internet governance? Finally, what are the options for counterbalancing this disproportionate influence? The significance of this last question has amplified since the United States decided not to sign the Trans-Pacific Partnership (TPP), an international trade agreement in which stipulations regarding e-commerce, data protection, and other digital economic concerns could have counteracted growing Chinese geopolitical and economic influence in the Asia-Pacific.

While it is all too easy to label the DSR merely another manifestation of the "authoritarian turn" alternately cited as either cause or effect of China's growing international influence, this paper argues that more economically-motivated dynamics are fueling the DSR. In Southeast Asia alone, several states' censorship practices, arrests of online dissidents, and nascent ICT regulations designed to legalize these efforts suggest a long-standing affinity for the top-down control of the internet in China. It is not coincidental that most DSR target states are notorious for their repressive domestic internet governance. The most troubling outcome of the DSR's development would be if Chinese firms develop an uncontested monopoly over both ICT markets and user data in the DSR's target countries. Concerning long-term outcomes include stymied domestic innovation in these countries due to over-reliance on Chinese firms to maintain and secure technological infrastructure, and the potential adoption of restrictive local technology policies that may deter foreign investment and the entry or expansion of certain multinational firms in these markets.

Through a literature review of Chinese-language secondary sources this study identifies how China's government and top technology companies are cooperating to 1) construct and maintain internet and telecommunications infrastructure 2) establish and internationalize the technical standards on which this infrastructure is dependent, and 3) provide the digital content and services that sit atop this infrastructure (e.g., mobile payment systems, digital messaging applications, and ecommerce platforms) to emerging markets. Promises of "leapfrogging" technological growth in the least technologically developed countries, the presentation of Chinese state policy goals as matching specific domestic technology development strategies within certain ASEAN states, and technological training of workers in small and medium-sized enterprises (SMEs) seeking access to China’s ecommerce and mobile payment platforms, and joint ventures with local companies are among some of the more common strategies China’s top technology firms are pursuing with varying degrees of Chinese government support.

While this list is not exhaustive and differs based on the technological objectives, infrastructure, and economic development of each ASEAN member state, the general promise of expediting the globalization process within these countries links these approaches together. This paper highlights on-the-ground developments that reflect a number of the fifteen high-level goals laid out in the “Proposal for International Cooperation on the 'One Belt, One Road' Digital Economy”, an agreement several countries signed at the 2017 World Internet Conference in Wuzhen, China. Given that this agreement’s signatories comprised mostly non-ASEAN countries, it remains to be seen whether the strategies employed in Southeast Asia may be effective elsewhere.

Wuzhen and the Appeal of a “Chinese Model” of Internet Governance

Since 2014, Chinese government and tech industry representatives have hosted an annual World Internet Conference (WIC) in Chinese village of Wuzhen. High-level representatives from foreign governments, tech companies, and research centers have attended a showcase of homegrown Chinese tech giants’ products and services alongside forums discussing Chinese objectives for domestic and international technological growth and regulation. The WIC’s tacit premise is that these foreign states and companies—regardless of their vastly different levels of technological advancement—all stand to benefit from partnering with their increasingly powerful Chinese counterparts, yet must respect China’s model of Internet governance in order to access this benefit. The originally overt political messaging of the Wuzhen Summit has been refined into a subtler form of telegraphing that the recent “Proposal for International Cooperation on the 'One Belt, One Road' Digital Economy” (hereafter the Wuzhen Digital OBOR document) reflects.

At a ‘Digital Silk Road’ International Cooperation Forum held at the 2017 WIC, representatives from China, Laos, Thailand, Turkey, Saudi Arabia, the United Arab Emirates, and Serbia signed an agreement on a variety of digital infrastructure and trade initiatives. The “voluntary, non-binding” (*New Media*, 2017) Wuzhen Digital OBOR document flagged fifteen high-level objectives for its signatories to collectively pursue (People’s Web, 2017), notably including:

1. Expand broadband access [and] raising broadband quality.

2. Stimulate the digital transformation… of agricultural production… manufacturing sector… cultural education, healthcare and medicine, environmental protection, urban planning… service sectors such as smart logistics, online tourism, mobile payment, digital creativity and the shared economy.

3. Stimulate e-commerce cooperation.

4. Support internet startups and innovation.

5. Stimulate development of micro, small, and medium-sized enterprises (MSMEs).

6. Strengthen digitized skills training.

9. Increasing digital inclusivity… to close the digital divide, including the digital divide between and within countries.

10. Encouraging transparent digital economy policymaking.

Other important points in the document included calls for cooperation on international cybersecurity and digital standards-settings initiatives, which are well under way in areas such as Chinese construction of 5G networks (Kania, 2017). Helping developing countries rapidly digitize through a mix of investment, education, infrastructure-building, and joint ventures are all common themes in the DSR overtures Chinese government officials and, increasingly, tech companies, are making in Southeast Asia.

These appeals come at a time when observers are documenting the growing attractiveness of a loosely defined "Chinese model" of domestic Internet governance and digital economic growth. Chinese digital economy expert Samm Sacks has argued that certain developing countries who admire the rapid expansion of China's tech sector as well as the heavy hand with which the Chinese Communist Party (CCP) has managed to control domestic Internet access and content are beginning to emulate practices associated with the latter while courting investment assumed to bring about the former. Sacks cites Sino-Tanzanian government consultations on censorship tools and data localization policies, along with the recent development of the controversial Vietnamese cybersecurity law (discussed below) modeled after the Chinese national cybersecurity law, as critical examples of this shift (Sacks, 2017). While US tech giants have increasingly strained relationships with law enforcement and other government officials in some of the same countries that are drawn to the Chinese model, Beijing, by contrast, might be laying the groundwork to transition from exporting hardware to exporting policy frameworks.

Methods

At least one other report (Lewis, 2018) documents a host of major bilateral and multilateral means through which Chinese firms and the state have supported ICT development across ASEAN. Furthermore, a thorough review of the economic motivations driving the Digital Silk Road has identified five objectives tech companies are pursuing with varying degrees of state support: "mitigate industrial overcapacity, facilitate corporate China's global expansion, support the internationalization of the renminbi, construct a China-centered transnational network infrastructure, and promote an Internet-enabled 'inclusive globalization'" (Shen, 2018 PP 2683). While both studies seek to explain the origins and potential future of the DSR, the question of the extent to which the state and the tech sector are coordinating behavior remains unanswered.

In this analysis, each ASEAN country is separately assessed using examples of its state and commercial exchanges with China, and, where relevant, recent domestic developments in technology policy (cybersecurity, data privacy, and related standards and laws) that echo elements of similar policies in China. The accounts of bilateral engagements provided in each case study originates from secondary source material written in Mandarin and published in Chinese academic, policy, and tech industry journals and newspapers.[69]

China-ASEAN Technological Ties

Although the ASEAN countries are but a sample of Digital Silk Road focal sites, the political and economic cohesiveness of the region may be significant if DSR buy-in from a plurality or majority of member states creates a winner-takes-all outcome for China. Historically, ASEAN has tended to act as a bloc on economic and security issues, using relatively informal consensus mechanisms to reach resolutions. There has been speculation about ASEAN attempting to build a digital single market modeled after the EU's own efforts, despite the vastly different levels of technological development of its member states. (Tey, 2017).

By one estimate, the Southeast Asian ecommerce market is projected to reach USD $88 billion by 2025 (Setboonsarng and Zhu, 2017). A co-authored article by a researcher from China's influential Ministry of Industry and Information Technology (MIIT)-housed government think tank China Academy of Information and Communications Technology (CAICT) and a representative from a Chinese ecommerce firm notes that China and ASEAN combined form the world's biggest internet market, projecting that ASEAN's digital economy will grow by 6.5 times its current size by 2025 (Chen and Liu, 2017, p. 59). The growth of this network brings hundreds of millions of users in contact with ASEAN markets while solving end-to-end logistics and mobile payment infrastructural problems in Southeast Asia. Part of the strategy for entering these markets is to claim alignment of domestic Chinese industrial policy objectives with those such as ASEAN's ICT 2020 strategy (Chen and Liu, 2017, p. 61), a move echoed on a country-by-country level in Thailand and the Philippines. Chinese policymakers and industry figures also advocate for creating "smart grid" (智能电网, *zhineng dianwang*) systems connecting China and ASEAN. One DSR-related initiative within China's own borders, the China-ASEAN Information Harbor, is expediting this type of development.

Information Harbor

At the 2014 China-ASEAN Cyberspace Forum, representatives from China and all ten ASEAN members' governments formally agreed to establish a "China-ASEAN Information Harbor" (中国-东盟信息湾, *zhongguo dongmeng xinxi wan*)— a China Unicom-led project with bases in

[69] The author conducted all translations of Chinese source material, except for cases in which English translations of a source's title or a journal's name were already presented alongside the original Chinese. In these instances the author used the provided translations even when they inaccurately or inelegantly represented the Chinese text, as this may simplify the process of other scholars locating these resources. Manual searches of two major Chinese databases, China National Knowledge Infrastructure (CNKI) and the Chaoxing (Superstar) Database for terms including the names of each ASEAN member state, ASEAN (东盟, *dongmeng*), ecommerce (电子商务, *dianzi shangwu*), the names of some of China's biggest tech firms, China's national "Internet Plus" (互联网+, *hulianwang*+), and truncated variations of several of these terms (e.g., both 阿里巴巴 *Alibaba* and the nickname 阿里 *Ali* for the ecommerce giant) were several among a longer list of phrases used.

Guangxi province’s Nanning and Qinzhou (Ji and Huang, 2017, pp. 58) and has seen involvement from ecommerce firms including JD.com and retail giant Suning, as well as top Chinese express delivery companies. Two years later, China’s State Council granted formal approval for the construction of the Information Harbor, including it in the Thirteenth Five-Year Plan (Wang Gongqing, 2017, p. 57) Although the Information Harbor has received relatively little foreign publicity in the newly emerging discussions of the DSR, it was enough of a priority in 2016 to attract representatives from the National Development and Reform Commission (NDRC) and MIIT to attend a China-ASEAN Information Harbor Forum, where NDRC deputy director Lin Nianxiu delivered a speech entitled “Joint Construction of a China-ASEAN Information Harbor Enables the Establishment of a Maritime Silk Road for the Benefit of All States’ Citizens” (*Henan Technology*, 2016).

Media and industry portrayals of the Information Harbor suggest that it will become a hybrid of an industrial park, an expo center, and cloud computing center, with some of its main tech sub-sectors—fintech, telecommunications infrastructure, ecommerce— as varied as those promoted under the DSR label itself. Given Guangxi’s proximity to the Vietnamese border as an entry point to the rest of Southeast Asia, the Information Harbor has also been touted as a potential source of information sharing and resource mobilization in response to emergencies, natural disasters, and epidemics (Ji and Huang, 2017, pp. 61). As of May 2017, over 64 so-called “major projects” have been associated with the Information Harbor, with investments totaling 45.9 billion RMB (*Times Finance*, 2017). As the most explicit example of a multilateral, DSR-branded project, the Information Harbor lacks the demonstrable and deliverable results of many of the firm-led expansions detailed below.

Malaysia

Both state- and firm-level Chinese engagement with Malaysia’s government and tech companies reveal largely private sector-led long-term strategic thinking about how to access the young, burgeoning population of Malaysian technology users and ecommerce vendors. In 2017 Alibaba, China’s biggest ecommerce firm and one of the country’s monopolistic tech giants, established the first “digital hub” of its Electronic World Trade Platform (eWTP, 电子世界贸易平台, *dianzi shijie maoyi pingtai*) in Malaysia. Not only does the eWTP provide opportunities to internationalize Alibaba’s ecommerce platforms, it also enables Alibaba’s logistics arm, Cainiao, and ten additional Chinese logistics partners to operate outside of China’s borders (*China Storage and Transport*, 2017). This development is explicitly framed as contributing to the OBOR strategy, with one author noting that Malaysia has been among the first and most enthusiastic ASEAN supporters of the Belt and Road (Liu 110).

Liu Tingting comments that as China’s largest trading partner in ASEAN for the last eight years, Malaysia has been an ideal site for a series of Alibaba parternships: Alipay, the Chinese market-dominating mobile payment platform of Ant Financial (an Alibaba spin-off fintech company), has paired with Malaysia’s digital wallet company Touch n’ Go as well as major Malaysian banks including Public Bank, Maybank, and CIMB, to provide a “Malaysian version of Alipay.” Alibaba Cloud, the company’s cloud computing branch, plans to establish an enterprise cloud computing system in Malaysia. Moreover, Alibaba has promised to create a local ecommerce worker training program to teach a minimum of 1,000 professionals in small and medium-sized enterprises (SMEs) per year how to become vendors on Alibaba's ecommerce platforms. Worker training initiatives in Malaysia are uniquely positioned to take advantage of the multi-generational community of ethnic Chinese in the country who have established business networks spanning a variety of sectors (Hamilton-Hart, 2004, p. 176).

Alipay's main Chinese competitor, Tencent's WeChat Pay, has applied for a license to operate in Malaysia (S. Jiang, 2018). The app will allow local users to tie their Malaysian bank accounts to the payments platform, marking WeChat Pay's inaugural overseas market entry. This expansion also signals the opening of a new arena of the Alibaba-Tencent rivalry, at a time when both companies may be preemptively wary of market saturation in China. Public discussions of the Chinese tech juggernauts' expansion into Malaysia have yet to address concerns about cross-border data protection, or the ill-defined path path to fulfilling nebulous claims of technological "leapfrogging." Alipay and WeChat Pay have crowded out competitors in China; how might potential Malaysian mobile payment competitors fare? The same question applies to SME worker training initiatives, which carry the advantage of connecting local Malaysian businesses to markets abroad, while also placing Alibaba in command of a trove of user data on these transactions that could enable the highly valuable spin-off of new services.

On the hardware side, mobile devices such as Huawei and Xiaomi phones are commonly used in Malaysia. In November 2015, Huawei established a new location of its international chain of research centers, OpenLabs, in Kuala Lumpur (Huawei 2017). The Malaysian OpenLab, which Huawei calls a "base for Asian enterprise cloud data," is meant to promote collaboration between Huawei and local Malaysian enterprises and is located in the Iskandar region, referred to as the "next Shenzhen," a comparison to one of China's major technological manufacturing and design cities (Chen, 33). On a bilateral level, the Chinese and Malaysian governments have also established a "two countries, two parks" (两国双园, *liangguo shuangyuan*) plan to build the Malaysia-China Kuantan Industrial Park in the port city of Kuantan, Malaysia along with the Malaysia Qinzhou Industrial Park in Guangxi, China (Hong Kong Trade Development Council, 2017).

Malaysia was one of the main ASEAN champions of OBOR under former Prime Minister Najib Razik, who was arrested for corruption in the spring of 2018 and lost the same year's election to Mahathir Mohamad. Mahathir has taken a more sober approach to questioning the value OBOR brings to Malaysia, halting over USD $20 billion worth of Chinese railway and pipeline construction projects pending further review. On this matter, Mahathir has commented that "We will be friendly with China, but we do not want to be indebted to China," while in Tokyo (Wen and Latiff, 2018), it remains to be seen whether his regime will apply the same degree of caution to the DSR.

The Malaysian case may come to demonstrate how even as political attitudes towards Beijing shift, overseas dependence on Chinese tech firms such as Alibaba may nonetheless continue to expand. In particular, the Chinese literature on Alibaba's globalization strategy describes Malaysia as a test site for the eWTP model of building and applying the infrastructure underlying Alibaba's ecommerce ecosystem in countries lacking the capital and experience to do so themselves (Sun and Huang, 2018).

Indonesia

Comparable to its presence in Malaysia, Alibaba has likewise begun to establish high-level connections in Indonesia. The extent to which the founders and senior leadership of China's biggest tech companies consults with government officials has already drawn concern from critics abroad, many of whom have failed to recognize that Jack Ma has agreed to serve as an ecommerce advisor to the Indonesian government under President Joko Widodo. In addition, Ant Financial (which Ma has spoken on behalf of despite the company's independence from Alibaba) has rolled out its popular Alipay mobile wallet app in Indonesia through a joint partnership with the latter's Elang Mahkota Teknologi (Lee, 2018).

A published interview with Indonesian Ambassador to China Soegeng Rohardjo from 2017 presents Indonesia as yet another major proponent of OBOR. The article quotes Rohardjo as saying that "the Belt and Road Initiative will inevitably strengthen connections between China, other Asian states, Africa, and Europe" (Wang Shiyu, 22). The article likewise quotes Rohardjo's commentary on how Indonesia will reevaluate its list of rejected investments, relax restrictions on foreign investors, and further ease investment restraints on majority-shareholder foreign investors.

Indonesia may also be a welcoming market for services that have encountered regulatory pressures in China, such as peer-to-peer (P2P) micro-lending. Aside from Alipay parent company Ant Financial's efforts to expand fintech services including mobile payments and microlending in the country, Chinese lending platforms such as WeShare have applied for licenses from Indonesia's central bank and aim to use their expertise in artificial intelligence to gather and analyze troves of data on potential borrowers (Lee, 2018). This type of expansion raises the question of whether home-grown Indonesian tech startups will be able to compete with their Chinese counterparts who have already developed sophisticated modeling and analysis techniques along with user data on Indonesians.

The Indonesian case presents a trend whereby Chinese tech firms establish strong ties with senior government leaders in DSR target countries, often prior to and without the assistance of representatives of China's own government. Once again, this demonstrates more strategic thinking from the private sector than the state, and it remains to be seen if and how the latter will "catch up" to the former in support of the DSR.

Singapore

As the most technologically advanced of the ASEAN countries, Singapore is a unique forerunner that China seeks to emulate. Chinese experts have praised smart city initiatives in Singapore and Malaysia as "enlightening" and worth studying (Chen and Liu, 2017, p. 60). Describing the first China-Singapore Internet Forum in 2016, one Chinese tech industry magazine articulated what each country offers and gains by partnering on digital economic initiatives: Singapore possesses the market access to and familiarity with differing levels of development of Southeast Asian countries that could help Chinese companies striving to expand overseas as part of OBOR, whereas China is a massive market for Singaporean goods and services (*Innovation Time*, 2016).

Alibaba acquired a 10.35% stake in Singapore Post, the country's national postal service provider and an inroad into Singaporean ecommerce logistics (Bai, 85). In 2017, Ant Financial acquired the Singaporean mobile payment company helloPay Group, which is housed within the ecommerce company Lazada. Originally founded in Singapore through the German startup accelerator Rocket Internet, Lazada was designed to mimic Amazon's business model in Southeast Asia. Not only does Alibaba now have an 83% stake in Lazada, but as of Alibaba's acquisition of helloPay Group, helloPay's mobile wallet application has been renamed Alipay Singapore, Alipay Malaysia, Alipay Indonesia, and Alipay Philippines in each of these respective ASEAN states (Xiao 2017).

In 2012, Chinese search engine giant Baidu established a joint research laboratory with Singapore's Agency for Science, Technology and Research (A*STAR), an institute managed under Singapore's Ministry of Trade and Industry. The scope of joint research projects at the lab include natural language processing (NLP) of Thai and Vietnamese, along with voice recognition and information retrieval (*China Internet*, 2012). Aside from the Baidu lab, Huawei also built an OpenLab in Singapore (Huawei, 2017).

Singapore's uniqueness among the DSR countries is further bolstered by its proactive approach to developing domestic technology policies and laws earlier than its fellow ASEAN member states. Thus Singapore's tech policies may be the most impervious to influence among the ASEAN countries, while the country may serve as a base from which Chinese tech firms can gain legitimacy and expand within Southeast Asia.

Thailand

Thailand provides numerous examples of parallels between Chinese state-firm strategic alignment in pursuing bilateral ICT market connectivity. A 2017 paper in the Chinese journal *Global Markets* notes similarities between the "Thailand 4.0" strategy and the digital economic goals of OBOR, in that the former's strategic objectives include advancing domestic development of automation, intelligent systems, the Internet of Things, and "smart" medicine and agriculture, with the overarching goal of "avoiding the middle-income trap" (Huang 2017, 8). The article argues that the biggest challenges to the Thai economy include rising household debt levels, protectionist US trade policies, Brexit, and slower economic growth in Europe. In contrast, China's OBOR plan and "Made in China 2025" strategy are posed as buttressing "Thailand 4.0" with the claim that "The launch of the 'Thailand 4.0' strategy, especially the Eastern Economic Corridor's ten major industries—including smart electronics, the automobile industry, robotic manufacturing, aviation, and digitization—are all in line with the rising and mature new industries in China, which is conducive to China's 'going out' (走出去, *zouchuqu*) and to actualization of the 'Thailand 4.0' strategy" (Huang 2017, 8).

The large diaspora of Thai citizens descended from multiple generations of Chinese merchants mitigates some of the tensions experienced in other Southeast Asian countries, with the same e-commerce representative remarking that his company employs many ethnically Chinese Thais. Yet even the hiring of local labor may not yet compensate for the absence of local competition in e-commerce. The domestic Chinese rivalry between Alibaba and JD.com is slowly crossing over to the Thai market. Although JD has taken up a USD $500 million joint venture with Thai retail conglomerate Central group with the hopes of not only breaking into Thailand but also of using it as a launching point to later serve Malaysian and Vietnamese markets (Setboonsarng and Zhu, 2017), Alibaba is at an advantage for its majority ownership of Lazada, and neither of the two e-commerce companies appears to be challenged by home-grown Thai competition.

Chinese tech firms may rely on the rhetoric of DSR and its alleged similarities to "Thailand 4.0" to enter the Thai market, yet it remains to be seen whether they stand by these promises after gaining access. Much like in the example of Indonesia, it is not yet clear whether these DSR benefactor countries are weighing the benefits of technological investment and infrastructure building against potential long-term reliance on Chinese tech firms that could stymie the development of local startups.

Vietnam

Vietnam has conducted state-level policy planning for ecommerce since at least 2006 (Fanshi Meijuan et al, 2014, p. 26), although as one Chinese CEO of a business-to-business (B2B) retail platform operating in Vietnam noted, "Vietnam is similar to China when it was opening up in the 1990s" in terms of reliance on foreign imports, limited purchasing power, and a manufacturing sector that faces the challenge in which "and the same goods at a lower price from China are [buyers'] first choice" (Chen, 2013). This comparison may provide two advantages to China's tech giants: familiarity with the hurdles of developing infrastructure and services in developing countries, along with these same countries' governments admiration of China's own transition as an advertisement for how Chinese firms can transform Southeast Asian markets.

There are limits to this appeal in Vietnam, however. Hostilities stemming from ongoing maritime disputes in the East South China Sea are among the more recent sore spots in already fraught Vietnam-China relations. Despite these tensions, a seemingly less controversial vector of engagement between China and Vietnam has been in combating cross-border cybercrime. This partnership between law enforcement in both countries may be considered less remarkable, however, given that Vietnam has actively engaged in international information security industry trainings and other initiatives with Microsoft as well as with the South Korean and UK governments (Zhou, 2015).

How can the gap between China as purveyor of a desirable model of Internet governance and China as a historically untrustworthy neighbor be reconciled in analyzing the Vietnamese case? Prior to the June 2018 passage of Vietnam's cybersecurity law, local protestors attempted to block its adoption, and a range of NGOs, trade organizations, as well as domestic and foreign firms[70] spoke out against the law as a setback for economic growth, free speech, and innovation (Nikkei Asian Review, 2018). Both the US Embassy and the Asia Internet Coalition (drawing membership from Facebook and Google, among other major foreign tech firms) have cited the cybersecurity law's data localization stipulations and the power it grants Vietnamese law enforcement to make

[70] This list includes Human Rights Watch, Amnesty International, the American Chamber of Commerce, Lazada, and Toshiba. (2018, June 12). Vietnam's cybersecurity law sparks concerns from businesses. *Nikkei Asian Review*. Retrieved from: https://asia.nikkei.com/Politics/Vietnam-s-cybersecurity-law-sparks-concerns-from-businesses.

decisions about censorship as some of the law's chief impediments to Vietnam's digital economy (Uyen and Boudreau, 2018). The law has drawn criticism for resembling the rigidity of China's own national cybersecurity law (Al Jazeera, 2018), which multinational firms spanning many sectors beyond tech alone have long criticized on similar grounds. Domestic protests against the cybersecurity law were somewhat eclipsed by much larger concurrent anti-China protests against the legalization of special economic zones with 99-year land leases for foreign firms, which many Vietnamese suspect will become a source of Chinese control over Vietnam (Elmer, 2018).

More so than in other ASEAN countries, Vietnam's contentious history with China may inhibit the entrance of Chinese tech firms into local markets. The higher level of public scrutiny applied to domestic tech policy developments further complicates both state and firm-level Chinese engagement with Vietnam.

Cambodia and the Philippines

Although there has been limited discussion of the DSR's extension to Cambodia, the potential for this development tends to be discussed as an eventual follow-up to the perceived successes of traditional BRI enterprises that have created thousands of jobs in places such as the Sihanoukville Special Economic Zone (Sun and Wei, 2017). In a distinctively public-private arrangement that may not appear to be connected to China at first glance, one of the projects initiated through the China-ASEAN Information Harbor was the construction of Cambodia's National Data Center (Cheng, 2016, p. 24) through Southeast Asia Telecom (SEATEL), a company whose chairman identifies it as a "purely Chinese-invested private enterprise" (Zuo, 2017).

As in the case of Cambodia, the Philippines—specifically since the 2016 inauguration of Rodrigo Duterte—is slowly favoring Chinese infrastructural investment over admittedly limited Western alternatives. China Mobile has been awarded a license to build mobile networks in the Philippines and a localized version of Alipay has been rolled out in the country as well.

In March 2017, the International Ecommerce Trade Promotion Association of China (IETPA) made an official visit to the Philippines to promote cross-border e-commerce and other international trade initiatives under the auspices of the "Maritime Silk Road" (Xiao F., 2017). Xiao Fangchen's report of the event for *China Informatization* drew parallels between the controversial national industrial upgrading strategy "Made in China 2025" and the Philippines' "AmBisyon 2040," a nationwide plan to alleviate poverty and establish a uniform standard of living for all Filipinos. Citing the Duterte presidency as a "honeymoon period" for Sino-Philippine relations, Xiao's article is one of the notable few to comment on the opportunities that OBOR will provide for overseas Chinese communities who have developed strong commercial networks across Southeast Asia over multiple generations.

Myanmar

As an example of an ASEAN member state that is not as technologically advanced as some of the countries China has deep and multifaceted engagements with, Myanmar is nonetheless an overseas market flooded with Chinese hardware and software. The popularity of Huawei mobile phones and rising use of Tencent's WeChat messaging application are two examples provided for the ubiquity of Chinese products and services in both first-tier Burmese cities and more remote regions of the country (Tang, 111).[71] A representative from Wave Money, a Myanmar-based fintech platform, described the current business environment in the country as a "land grab" in which mostly foreign companies are vying to establish critical ICT infrastructure. Huawei and ZTE, for instance, have partnered with Qatari telecommunications provider Ooredoo and the multinational Telenor group in establishing telecommunications networks across Myanmar.

While Myanmar is in the midst of an overhaul from paper-based banking to digitized systems, a mix of Japanese firms as well as U.S. companies such as Microsoft and Oracle have contributed to building out end-user and enterprise versions, respectively. Notably, Chinese companies have not stepped into the banking infrastructure arena in Myanmar, yet they are making tentative moves toward the data-rich fintech market that may one day sit atop this national banking system. Ascend Group's True Money product, in which Ant Financial has a 30% stake, is one of the few mobile payment competitors in Myanmar. As in the case of Vietnam, it appears Chinese companies are cautiously entering the Burmese market through joint ventures given the state of political ties between China and Myanmar. One fintech startup representative described the Myanmar government as "on the fence" about welcoming Chinese firms into the country, speculating that negative international perceptions of Myanmar for the Rohingya refugee crisis may leave the country with fewer supporters of domestic development projects.

Efforts to connect China and Myanmar through OBOR and other infrastructure projects persist despite other challenges. In the Chinese literature on Myanmar's technological propsects, Myanmar is conceptualized in terms of its contiguous border with China's Yunnan province, an agriculture and tourism hub that Xi Jinping has described as having the potential to serve as a connective center between China and Southeast Asia (Jiang, 2017, p. 82). Geography is one of the many obstacles for ICT companies expanding in a country as large as Myanmar, with the Wave Money respondent adding regulatory changes as the country seeks to modernize and low digital literacy in spite of high smartphone adoption rates to this list. Yet some Chinese commentators see long-term potential in creating broadband infrastructure "arteries" linking Guangxi to Singapore, Myanmar, and Vietnam.

Laos and Brunei

Much like in the case of Myanmar, in the Chinese literature Laos is primarily spoken of in terms of its shared border with China's Yunnan province. This proximity is presented as an opportunity to expand cross-border ecommerce between China and Laos, explicitly framed in at least one article as contributing to OBOR (Li and Wu, 30). Unlike Malaysia, where Alibaba is conducting local trainings for SMEs to open online shops via Taobao, Laos is offered as a potential site for building warehouses to manage ecommerce logistics. As far back as 2006, China won the Laotian

[71] One Myanmar-based interviewee pointed out that WeChat is not the predominant messaging platform in the country, as it competes with the more popular alternatives of Facebook and Japan's Viber.

government's right to construct major telecommunications and internet infrastructure in the country (China Engineering Construction Newsletter, 2006).

At the 2017 WIC, Laos was one of the signatories of the OBOR digital economy agreement, with Laos' Vice Minister of Posts and Telecommunications Bounsaleumsay Khennavong praising the initiative's potential to close the digital divide (Xinhua 2017). Laos and Thailand are also notable for adoption of the Beidou navigational system, which could popularize China's alternative to the Global Positioning System (GPS) used in much of the rest of the world (Chen and Liu, 2017, p. 60). Notably, Brunei is the one ASEAN country for which no substantive data or DSR-related planning could be found.

Discussion

The Trans-Pacific Partnership

The United States' withdrawal from the Trans-Pacific Partnership (TPP) in 2017 drew criticism for signaling the US government's retreat from serving as a trade mediator within the Asia-Pacific region, ending the Obama administration's "pivot to Asia". Yet given that the digital economy was but a small part of this mega free trade agreement's (MFTA) agenda, questions about how the TPP's stipulations on issues such as data protection and localization, and mandatory partnerships with host country firms have received minimal attention.

As Jane Kelsey notes in a highly critical assessment of the TPP as establishing the "wrong model of e-commerce" for ASEAN, the TPP is only one example of numerous FTAs containing a "common core of text" with "a chameleon-like presence: the same provisions may appear in chapters on cross-border services, investment, e-commerce, intellectual property, and transparency" (Kelsey 21), which, in aggregate, produce a pattern in which developing countries are rule-takers subjected to the preferences of US and European rule makers. Aside from the argument that the TPP would impede economic growth of developing states' digital economies while further increasing their dependence on the core set of mostly US tech firms that dominate the sector, civil society groups lambasted the TPP as diminishing these countries' rights to autonomy in determining their domestic technology policies. What did the US lose on the digital economic front from leaving the TPP, and what does this mean for China?

Singapore, Malaysia, Brunei, and Vietnam are the four ASEAN members to have ultimately signed the TPP, leaving open the question of how changes in these countries' technology policies that contradict the TPP may be contested, and by whom. Most policy analyses of the TPP's digital trade components were written before the United States abandoned the agreement, and of these the unfavorable accounts zeroed in on the self-interested behavior of US firms and business organizations. Kelsey, for example, pointed out that the US Information Technology Industry Council (ITIC) disapproved of proposed data localization, source code disclosure, encryption key sharing, and other technology policies that Vietnam, along with non-TPP member states Indonesia and the Philippines, proposed (Kelsey, 2017, p. 9-10). Some of these policies have been passed under Vietnam's recent cybersecurity law. Although the American Chamber of Commerce in

Vietnam has complained about the law's contradiction of both WTO and TPP commitments (Nikkei Asian Review, 2018), the United States' absence from this MFTA has cut off one major avenue of collective pushback against both the Vietnamese cybersecurity law and future iterations of similar regulations across the TPP signatory states.

Although China was deliberately excluded from the TPP and has tried to counter this with its own Regional Comprehensive Economic Partnership (RCEP), a broader current shaping these FTAs is one that a small but growing number of scholars and industry figures have identified as a dynamic in which the United States, the European Union, and other Western economies will capture the digital markets of developed countries, whereas China is focused on doing the same for developing countries across the world. Venture capitalist Kai-Fu Lee has predicted this division will play out in the so-called "AI race" between the United States and China, noting the significant advantages China would hold in capturing user data from a wide variety of developing economies. Kelsey's list of requirements for ASEAN to attain autonomy in the development of individual member states' and the regional organization's digital economies reads as almost an exact replica of what many Chinese government officials and technology companies are promising through invocations of the Digital Silk Road: "international, regional, and national rules that facilitate digital industrialization, close the digital divide, and correct the development asymmetries that currently favor developed countries and their corporations" (Kelsey 1).

There is a sense that a "China model" has risen in the absence of an alternative effort to meet the needs of ASEAN countries and other developing states seeking technological investment and skills training. Whether or not China can facilitate autonomy in ASEAN's digital economy remains to be seen, given that this objective appears at odds with Kelsey's scenario that the "parallel risk for ASEAN [is] that Alibaba and its affiliates will control Asia's regional infrastructure, platforms and data, and become the gatekeeper for ASEAN countries wanting to harness new technologies and value chains for development" (Kelsey 8). Ultimately, US forfeiture of TPP membership may have cleared the path for further development of the "China model," removing one among many possible options for counterbalancing this influence.

International Response?

Interview respondents largely lamented the absence of local and foreign competition to match Chinese e-commerce and fintech outfits expanding in Southeast Asia, prompting a broader examination of how foreign firms are responding to the DSR. While the OBOR project has come under increasing scrutiny overseas as some benefactor countries have failed to repay exorbitant loans, early understandings of the DSR's ambitions are only beginning to arise in international media. In response to OBOR, some potential competitors are instead choosing to preemptively cooperate; Germany's Siemens, for example, has established a Belt and Road Initiative office in Beijing, and the company's chief executive has notably claimed that the BRI "going to be the new World Trade Organisation — like it or not" (Suokas 2018). Notably, some US-based companies have participated in China-led efforts to broaden engagement with Southeast Asian markets, such as Google and Microsoft sending attendees to the China-ASEAN E-Commerce Summit in 2016 (Cheng, 2016, p. 24).

Conclusion

Although the ASEAN region provides ample opportunities for outlining a Chinese model of technology-driven globalization, there are limits to this emerging model's replicability. Some of the factors that make Southeast Asia an obvious first choice for Chinese tech firms' internationalization processes—geographical proximity and cultural similarities to China, diasporic trade networks, and historical interdependency and cooperation on regional security and economic issues—are notably absent in other markets to which the promise of the Digital OBOR is directed, including the Middle East, sub-Saharan Africa, and Latin America. The strengths China possesses in Southeast Asia may be areas the Chinese government and the state's unofficial tech ambassadors may seek to compensate for through other means.

Features of the region that make it a fertile testing ground for reproducing a Chinese model of dominating ICT infrastructure, hardware, and software include a relative dearth of local companies that can compete with Chinese counterparts in terms of technological skill, quality of service, and product pricing; developing countries' hopes of technological leapfrogging in spite of major infrastructural barriers to growth (a challenge China is viewed as having rapidly overcome); limited efforts from the United States, the European Union, Japan, South Korea, and other developed democracies to counterbalance Chinese ICT expansion in the developing world; and finally still-developing ICT regulatory regimes that may draw influences from China, the US, and the European Union.

References

(2017). Ant Financial Acquires Singaporean Payment Service in SE Asia [蚂蚁金服收购新加坡支付服务 SE Asia]. *Informatization of China Construction* [中国建设信息化] (8), 6.

(2006). China Conducts National Telecommunications Reconstruction Project in Laos [中国承建老挝国家电讯改造工程]. *China Engineering Construction Newsletter* [中国工程建设通讯](9), 14.

(2017). China-ASEAN Information Harbor Construction Picks Up Speed [中国—东盟信息港建设全面提速]. *Times Finance* [时代金融](13), 47.

(2001). China-ASEAN Symposium on Information and Communications Technology Convened [中国-东盟信息通信技术研讨会召开]. *Telecom World* [通讯世界](5), 65.

(2017). Chinese Ecommerce Partakes in Adjustment of Global Logistics Chains [中国电商参与全球物流链大调整]. *China Storage and Transport* [中国储运](6), 90.

(2017). Electronic World Trade Platform (eWTP) "Experimental Zone" Established in Malaysia [世界电子贸易平台（eWTP）"试验区"落地马来西亚]. *Construction Machinery Digest* [工程机械文摘](2), 5.

(2012). Formal Launch of Baidu-Singapore Joint Laboratory [百度新加坡联合实验室正式启动]. *China Internet* [互联网天地](8), 64.

(2018, 15 June). The Great 'Firewall' of China: Digital economy thrives despite robust internet regulations. *Verdict*. Retrieved from: https://www.verdict.co.uk/great-firewall-china-digital-economy-robust-internet-regulations/.

(2010). Huawei Wins Malaysian Maxis Full Business NGBB Network Construction and Management Service Contract [华为赢得马来西亚 Maxis 全业务 NGBB 网络建设和管理服务合同]. *Network Telecom* [网络电信], 12(8), 71-72.

(2016). Inaugural China-Singapore Internet Forum Focuses on Digital Economic Cooperation. [首届中新互联网论坛聚焦数字经济合作]. *Innovation Time* [创新时代] (2), 98.

(2017, December 4). Initiative on Belt and Road digital economy cooperation launched. *Xinhua*. Retrieved from: http://news.xinhuanet.com/english/2017-12/04/c_136797807.htm.

(2017). Jack Ma Serves as E-Commerce Advisor to the Indonesian Government [马云担任印尼政府电子商务顾问]. *Times Finance* [时代金融](25), 44-45.

(2016). Lin Nianxiu: Establishment of China-ASEAN Information Harbor is a Major Strategic Measure That Enables the Establishment of a Maritime Silk Road [林念修：建立中国-东盟信息港是推动海上丝绸之路建设重大战略举措. 河南科技](19), 5.

(2017). ‘One Belt, One Road’ Digital Economy International Agreement Initiative Published

[《“一带一路”数字经济国际合作倡议》发布]. *New Media* [网络传播](12), 82-83.

(2018, June 13). Vietnam cybersecurity law a devastating blow to freedom: Amnesty. *Al Jazeera*. Retrieved from: https://www.aljazeera.com/news/2018/06/vietnam-cybersecurity-law-devastating-blow-freedom-amnesty-180613074931697.html.

(2018, June 12). Vietnam's cybersecurity law sparks concerns from businesses. *Nikkei Asian Review*. Retrieved from: https://asia.nikkei.com/Politics/Vietnam-s-cybersecurity-law-sparks-concerns-from-businesses.

Bai Zongyi [白宗义]. (2014). Alibaba’s Investment Logic [阿里巴巴的投资逻辑]. *The Internet Economy* [互联网经济](2)

Chen Cai and Liu Xiaoqing [陈才、刘晓晴]. (2017). China-ASEAN Cooperation on Informatization Drives Development of 21st Century Maritime Silk Road [以中国-东盟信息化合作推动 21 世纪海上丝绸之路发展]. *Telecommunications World* [世界电信] (2), 59-62.

Chen Fen [陈芬]. (2016). Malaysia: Learning From China’s “Internet Plus” [马来西亚：向中国学习互联网+]. *China Economic Information* [中国经济信息](9)

Chen Xiaolin [沉晓琳]. (2013). Creating Vietnam’s “Alibaba” [打造越南的“阿里巴巴”]. *Zhejiang Business* [浙商] (6), 100-101.

Cheng Zi [程子]. (2016). Co-Constructing a 21st Century Maritime Silk Road, Building a Closer China-ASEAN Destiny and Community [共建 21 世纪海上丝绸之路

共筑更紧密的中国—东盟命运共同体], *Food and Beverage Industry* [中国食品工业], (10) 22-25.

Elmer, K. (2018, June 13). Anti-China protests in Vietnam set to aggravate tensions with Beijing. *South China Morning Post*. Retrieved from: http://www.scmp.com/news/china/diplomacy-defence/article/2150653/anti-china-protests-vietnam-set-aggravate-tensions.

Fanshi Meijuan, Deng Xinting, Shi Feng [范氏梅娟, 邓新婷, 史烽]. (2014). A Comparative Study on E-Commerce Development in China and Vietnam [中越电子商务发展对比研究]. *Enterprise Science, Technology, and Development (Monthly Edition)* [企业科技与发展(上半月)](5), 11-12.

Hamilton-Hart, Natasha, “The Regionalization of Southeast Asian Business” in Pempel (ed.) *Remapping East Asia* (Ithaca: Cornell University Press, 2004):170-191.

Hong Kong Trade Development Council. (2017, May 16). Prospects for the Malaysia-China Kuantan Industrial Park and Kuantan Port. Retrieved from http://economists-pickresearch.hktdc.com/business-news/article/Research-Articles/Prospects-for-the-Malaysia-China-Kuantan-Industrial-Park-and-Kuantan-Port/rp/en/1/1X000000/1X0AA0CO.htm.

Huang Weirong [黄伟荣]. (2017). Thailand’s Economy and Business Opportunities Under the “Thailand 4.0” Strategy [“泰国 4.0”战略背景下泰国的经济与商机.] *Global Markets* [全球市场](30), 8 and 21.

Huawei. (2017, November 9). Huawei Announces New OpenLab in Malaysia to Drive Digital Transformation in APAC. Retrieved from: http://www.huawei.com/en/press-events/news/2017/11/Huawei-New-OpenLab-Malaysia-APAC.

Ji Chunyang and Huang Linlin [计春阳、黄琳琳]. (2017). Research on the Construction of Path Model for China-ASEAN Information Port Construction [中国-东盟信息湾建设路径模式研究]. *Around Southeast Asia* [东南亚纵横] (3), 58-62.

Jiang, S. (2017, July 11). Tencent's WeChat Pay seeks license for local payment services in Malaysia. *Reuters*. Retrieved from: http://www.reuters.com/article/us-tencent-holdings-malaysia/tencents-wechat-pay-seeks-license-for-local-payment-services-in-malaysia-idUSKBN19W0RN.

Jiang Wenrong [蒋文荣]. (2017). Research on the Development of “Internet Plus” in Dehongzhou [德宏州“互联网+”发展研究.] *Rural Science and Technology* [乡村科技](11), 82-85.

Kania, E. (2017, June 27). China’s play for global 5G dominance—standards and the ‘Digital Silk Road’. *ASPI Strategist*. Retrieved from: https://www.aspistrategist.org.au/chinas-play-for-global-5g-dominance-standards-and-the-digital-silk-road/.

Kelsey, J. (2017) The Risks for ASEAN of New Mega-Agreements that Promote the Wrong Model of e-Commerce. *ERIA Discussion Paper Series*, 1-51. Retrieved from: http://www.eria.org/ERIA-DP-2017-10.pdf.

Lee, G. (2018, June 9). China’s fintech companies are exporting AI and big data to Asia’s ‘laggard’ banking markets. *South China Morning Post*. Retrieved from: http://www.scmp.com/business/banking-finance/article/2149989/chinas-fintech-companies-are-exporting-ai-and-big-data.

Lewis, D. (2017). ICS Occasional Paper # 14, China’s Global Internet Ambitions: Finding Roots in ASEAN, 2-28. Retrieved from: http://www.icsin.org/uploads/2017/07/24/95ef81ecfcb1118bcfbf94b369d0ef1e.pdf.

Li Zheng, Wu Xiaosong [李峥, 吴晓松]. (2017). A New Type of Warehouse Model Applied in Cross-Border E-Commerce—A Case Study of Yunnan and Laos [一种新型仓储模型在跨境电子商务中的应用——以云南与老挝为例]. *Market Weekly* [市场周刊(理论研究)](6), 30-33, 35.

Liu Tingting [刘婷婷]. (2017). Analysis of China's and Malaysia's Trade Cooperation Developments and Prospects Under "One Belt, One Road" ["一带一路"背景下中国与马来西亚经贸合作发展与前景分析]. *Trade* [经贸实践](22), 110-111.

Liu Yifang, Fu Yunwei, Xu Lei, An Xiaoyin, Ren Jun, Bao Xuelin, Xu Haijing, Yang Zhou, Le Yanna, Tao Jun, Deng Qian, Zhang Xiaojun [刘轶芳, 傅云威, 许雷, 安晓荫, 任军, 包雪琳, 徐海静, 杨舟, 乐艳娜, 陶军, 邓茜, 张小军]. (2016). China Electricity Providers Connect All over the World [中国电商全球连线]. 中国名牌(23), 70-72.

Lou Xiangfei and Yang Jian [楼项飞, 杨剑]. (2018). Ending Latin America's Digital Divide and Joing Chinese-Latin American Construction of the "Digital Silk Road"[拉美数字鸿沟消弭与中拉共建"数字丝绸之路"]. *International Outlook* [国际展望] (5), 159-160.

Sacks, S. (2017, June 18). Beijing Wants to Rewrite the Rules of the Internet. *The Atlantic*. Retrieved from: https://www.theatlantic.com/international/archive/2018/06/zte-huawei-china-trump-trade-cyber/563033/.

Setboonsarng, C. and Zhu, J. (2017, August 23). JD.com, Thai retailer Central in talks for $500 million e-commerce JV: sources. *Reuters*. Retrieved from: https://www.reuters.com/article/us-jd-com-centralgroup-idUSKCN1B3113.

Sun Yang, Wei Xinning [孙阳, 魏欣宁]. (2017). Fast-Track Development of "One Belt, One Road" [搭乘"一带一路"快车谋发展]. *Culture Corporate* [企业文化](31), 32-34.

Suokas, Janne. (2018, March 23). Germany's Siemens sets up Belt and Road office in Beijing. *GB Times*. Retrieved from: https://gbtimes.com/germanys-siemens-sets-up-belt-and-road-office-in-beijing.

Stacey, K. (2018, December 11). China signs 99-year lease on Sri Lanka's Hambantota port. *Financial Times*. Retrieved from: https://www.ft.com/content/e150ef0c-de37-11e7-a8a4-0a1e63a52f9c.

Sun Qian and Huang Lina [孙倩, 黄莉娜]. (2018). 阿里之略 [Alibaba's Strategy]. 中国邮政 [*China Post*](8):8-10.

Tang Xiaoqian [唐小茜]. (2017). China and Myanmar's Tracks in the Field [中缅田野足迹]. *China Foreign Exchange* [中外交流](23).

Tey, S. (2017, July 6). Is ASEAN ready for the fourth industrial revolution? *Asian Development Bank Blog*. Retrieved from: https://blogs.adb.org/blog/asean-ready-fourth-industrial-revolution.

Uyen, N. and Boudreau, J. (2018, June 11). Vietnam Parliament Passes Cyber Law Denounced in Street Protests. *Bloomberg*. Retrieved from: https://www.bloomberg.com/news/articles/2018-06-12/vietnam-parliament-passes-cyber-law-denounced-in-street-protests.

Wang Gongqing [王功清]. (2017). China-ASEAN Expo Witnesses Growth of China-ASEAN Information Harbor [东博会见证中国—东盟信息港成长]. *China-ASEAN Panorama* [中国-东盟博览](3), 56-57.

Wang Junwei [王军伟]. (2016). Ecommerce Becomes a New Growth Point—Trends in China-ASEAN Trade Relations [电子商务成为新增长点—中国-东盟经贸关系趋势(下)]. *China Comment* [半月谈], (18), 17-19.

Wang Shiyu [王世钰]. (2017). "One Belt, One Road" Brings More Opportunities to Sino-Indonesian Cooperation—An Interview with Indonesian Ambassador to China Soegeng Rahardjo ["一带一路"为中印合作带来更多机遇——访问印度尼西亚驻华大使苏庚·拉哈尔佐]. *China's Foreign Trade* [中国对外贸易](5), 22-23.

Wei Liuting, Zhang Jing, Cao Lingdan, Zhou Mengzhen, Shijianbin [韦柳婷、 张晶、 曹令丹、 周锰珍、石建斌]. (2015). Research Survey on Southeast Asian Nations' Cross-Border E-Commerce [东南亚国家跨境电子商务调查研究]. *Cooperative Economy and Science* [合作经济与科技](14), 124-126.

Wen, P. and Latiff, R. (2018, July 5). Malaysia's Mahathir to visit China after putting $20 billion of projects on ice: sources. *Reuters*. Retrieved from: https://www.reuters.com/article/us-malaysia-politics-china/malaysias-mahathir-to-visit-china-after-putting-20-billion-of-projects-on-ice-sources-idUSKBN1JV1JP.

Xiao, E. (2017, April 18) Jack Ma's Ant Financial merges with Lazada's HelloPay Group. *Tech in Asia*. Retrieved from: https://www.techinasia.com/ant-financial-merge-hellopay-group.

Xiao Fangchen [肖方晨]. (2017). A Tour of the Maritime Silk Road—International Ecommerce Trade Promotion Association of China Goes to the Philippines for Exchanges, Integration, and Discussions [海洋丝绸之路之旅—中国民营科技国际电商贸易促进会赴菲律宾交流对接洽谈]. *China Informatization* [中国信息化] (4), 49-51.

Xin Xing and Yao Jingzhou [辛星, 姚景周]. (2018). On the 21st Century Digital Silk Road, Private Enterprises Have Bright Prospects [21 世纪数字丝绸之路，民营企业大有可为]. *Information and Communications Technology and Policy* [信息通信技术与政策] (9):5-8.

Zhang Fang, Duan Yuanping [张芳, 段元萍]. (2016). International Marketing Strategy for China's E-Commerce Industry in Southeast Asia [我国电商企业在东南亚地区的国际营销策略]. *E-Commerce* [电子商务](8), 44-45, 73.

Zhang Ying, Lu Xianting [张莺, 卢羡婷]. (2014). Golden Age Comes for Transnational E-commerce between China and ASEAN [中国东盟跨境电商迎来黄金年代]. *China Top Brands* [中国名牌](20), 26-27.

Zhong Mingzhang [钟鸣长]. (2016). Lessons From the Construction of Singapore’s Fintech Ecosystem [新加坡 FinTech 生态系统建设及其启示]. *Journal of University of Electronic Science and Technology of China (Social Sciences Edition)* [电子科技大学学报(社科版)], 18(6), 30-38.

Zhou Jili [周季礼]. (2015) Overview of Vietnam’s Cybersecurity Developments in 2014 [2014 年越南网络空间安全发展综述]. *China Information Security* [中国信息安全] (4), 94-99.

Zuo Huazheng [左华政]. (2017, March 17). Development Opportunities for Private Telecommunications Enterprises in the “Belt and Road” Initiative [“一带一路”倡议中 民营电信企业的发展机遇]. *China Pictorial* [中国画报]. Retrieved from: http://www.rmhb.com.cn/zt/ydyl/201703/t20170317_800091645.html.

Upstream Bundling and Leverage of Market Power*

Alexandre de Cornière[†] and Greg Taylor[‡]

May 18, 2018

Abstract

Motivated by the recent Google-Android antitrust case, we present a novel rationale for bundling by a multiproduct upstream firm. Consider a market where downstream firms procure components from upstream suppliers. U_1 is the only supplier of component A, but faces competition for component B. Suppose that component A increases demand for the downstream product and that contractual frictions induce positive wholesale markups. By bundling A and B, U_1 reduces its B-rivals' willingness to offer slotting fees to the downstream firm, thereby allowing U_1 to capture more of the industry profit. Bundling harms the downstream firm and the B rivals, and can be anticompetitive.

Keywords: bundling, exclusion, vertical relations.
JEL Classification: L1, L4.

1 Introduction

Competition authorities in Europe and in the US have recently been investigating potentially anti-competitive practices by Google on the mobile applications market. Google, which develops the open-source mobile operating system Android and many mobile applications, has in particular been accused by the European Commission of abusing its dominant position by imposing restrictions on Android device manufacturers.[1]

[*] We are grateful for useful discussions with Jay Pil Choi, Natalia Fabra, Sjaak Hurkens, Doh-Shin Jeon, Bruno Jullien, Markus Reisinger, and Patrick Rey. We also thank participants at numerous seminars and conferences for their constructive comments. Taylor acknowledges financial support from the Carnegie Corporation of New York.

[†] Toulouse School of Economics, University of Toulouse Capitole, Toulouse; alexandre.de-corniere@tse-fr.eu; https://sites.google.com/site/adecorniere

[‡] Oxford Internet Institute, University of Oxford; greg.taylor@oii.ox.ac.uk; http://www.greg-taylor.co.uk

[1] The EC's statement of objections is summarized at `http://europa.eu/rapid/press-release_IP-16-1492_en.htm`. See Edelman and Geradin (2016) for an analysis criticizing Google's practices. A Google response is available at `https://blog.google/topics/google-europe/androids-model-of-open-innovation/`.

One such restriction is application bundling: manufacturers who want to install Google Play also have to pre-install other Google applications (notably Google Search and the Google Chrome browser). Because Google Play is by far the largest Android application store,[2] the Commission argues that it is commercially important for manufacturers to be able to offer it to their customers. On the other hand, the "tied" applications (Search, Chrome, and others) face stronger competition, and Google's practices prevent its competitors from being installed either exclusively or in a prominent position on most devices.

The main existing theories of anticompetitive tying (see the literature review below) rely on a "predatory" logic: tying is only profitable to the extent that it successfully induces the exit or prevents the entry of rivals (see Rey and Tirole, 2007 for a discussion). In the Android case, the predation story is unconvincing: Google's practices have been in place for several years, and there are still credible rivals on the browser or search engine markets. Motivated by features of the Android case, we present a new rationale for (potentially anticompetitive) bundling that does not rely on a predatory logic.

Suppose a final product (e.g., a smartphone), sold by a downstream firm D, is made of various components (e.g. applications) provided by upstream firms. There are two categories of components, A (e.g., an app store) and B (e.g., a browser). A is solely produced by upstream firm U_1, whereas two versions of B exist, one produced by U_1 and the other by U_2. Upstream firms offer contracts to the downstream firm, who chooses which component(s) to use and then sells to consumers. For our theory to apply, the following three conditions need to hold: (i) substitutability between the two versions of B leads the downstream firm to install at most one version; (ii) the demand for the final product is higher if component A is installed than if it is not (*retail-complementarity*); (iii) *contractual frictions* leave upstream firms with a positive mark-up. In other words, upstream firms cannot offer efficient two-part tariffs. An example of such frictions is if upstream firms can exert some non-contractible effort to increases final demand.[3]

In such an environment, because of contractual frictions, providers of the B component obtain a positive markup for each consumer served. Since D can only choose one B provider, each one is willing to offer a positive slotting fee. This slotting fee is increasing in the expected demand for D's product. By bundling A and B_1, U_1 can reduce the slotting fee offered by U_2: indeed, under bundling U_2 expects that a final product that has component B_2 will not have A, and will therefore be bought by fewer consumers. Facing a less aggressive rival, U_1 can reduce the slotting fee it offers to D and thereby increase its profit. Such a strategy is always profitable when B_1 is more efficient than B_2, but also when the reverse is true provided that the presence of A has a large enough effect on the final demand. In the latter case, bundling is anti-competitive.

[2] An application store allows consumers to search for and install applications that are not already on their device.

[3] Another example of friction is downstream risk aversion coupled with a stochastic demand.

After discussing related literature in Section 2, we present our mechanism in Section 3 by focusing on the simplest form of contractual friction, where upstream firms can only offer fixed fees. There we discuss how our mechanism differs from the standard rationales for bundling. In Section 4 we allow for more general contracts. There we show that some form of contractual friction is necessary for bundling to be profitable. We then discuss a model with upstream moral hazard and two-part tariffs which delivers results that are qualitatively similar to those of the model with fixed fees. One difference is that two-part tariffs enable U_1 to leverage its market power without actually bundling A and B_1. This suggests that a ban on bundling would not be sufficient to restore efficiency, even though the anticompetitive outcome would no longer be the unique equilibrium. Section 5 concludes with a discussion of some extensions. In particular, our model can naturally be reinterpreted as one of wholesale bundling in a standard retail supply chain.

2 Literature

Bundling and foreclosure First dealt a blow by the Chicago School's Single Monopoly Profit Theory (e.g., Director and Levi, 1956; Stigler, 1963), the leverage theory of bundling was reinvigorated by various scholars who showed bundling could be profitably used to deter entry (e.g., Whinston, 1990; Choi and Stefanadis, 2001; Carlton and Waldman, 2002; Nalebuff, 2004).[4] Our mechanism does not rely on entry deterrence and is thus quite different from these.

In Carbajo, De Meza, and Seidmann (1990) and Chen (1997), bundling softens competition by generating horizontal differentiation (one firm offers product A while the other offers A and B as a bundle).

An important feature of our model is the vertical dimension of the market: bundling occurs at the upstream level. Previous papers have looked at this practice from different angles (see, e.g., Burstein, 1960; Shaffer, 1991a; O'Brien and Shaffer, 2005; Ho, Ho, and Mortimer, 2012). Closest to us is Ide and Montero (2016), who show how bundling by an upstream multiproduct firm can be profitably used to exclude an upstream rival. The mechanisms are different though, as illustrated by the different implications: in Ide and Montero (2016) bundling is necessary to achieve leverage (unlike here, see Section 4) and, more importantly, downstream competition is necessary for bundling to be profitable.

In our model, contracting frictions introduce cross-group externalities between upstream firms and consumers: upstream firms benefit from greater downstream demand. The paper therefore also relates to the literature on bundling in two-sided markets: (Choi, 2010; Amelio and Jullien, 2012; Choi and Jeon, 2016). In particular, Choi and Jeon (2016) is also motivated in part by the Google Android case. The modelling setup is quite different

[4] Fumagalli, Motta, and Calcagno (2018) provides an up-to-date review of the various theories and their applications.

however, since they do not model the vertical chain, and rely on a different kind of friction (the impossibility of charging negative prices to consumers) to show the possibility of leverage through tying, whereas our theory relies on the possibility of negative payments, i.e. slotting fees.[5]

Slotting fees Earlier literature has emphasized the role of slotting allowances as signalling/screening mechanisms (Chu, 1992), as well as their potential anticompetitive effects (Shaffer, 1991b; Shaffer, 2005; Foros and Kind, 2008; Caprice and Schlippenbach, 2013). In our paper slotting fees result both from the positive wholesale markup induced by the contractual friction (a mechanism discussed by Farrell, 2001) and from the constraint preventing the downstream firm from procuring both B components (see, e.g., Marx and Shaffer, 2010, for a discussion of this point). The purpose of bundling is then to reduce U_2's willingness to offer high slotting fees, thereby softening the competition for access to final consumers.

Exclusive contracts Because of the constraint preventing the downstream firm from using two different B components, a bundled offer is a sort of exclusive contract whereby the downstream firm agrees to buy both components from the same supplier. The difference with the standard models of exclusive dealing (e.g., Aghion and Bolton, 1987; Rasmusen, Ramseyer, and Wiley Jr, 1991; Segal and Whinston, 2000) is that the upstream firm can commit not to deal with a firm who rejects the exclusivity clause. Within that literature, Calzolari, Denicolò, and Zanchettin (2016) recently emphasized the role of contractual frictions in making exclusive dealing profitable. While they also focus on frictions that lead upstream firms to charge unit prices above marginal costs, their mechanism is quite different from ours. In particular, they do no rely on the kind of strategic effect (making rivals softer competitors) that is at the core of our argument.

3 Baseline model

Basic institutional environment A downstream firm, D, sells a finished good to consumers at price p. The finished good is made of components, obtained from upstream suppliers. There are two categories of components, A and B. Upstream firm U_1 is the sole producer of the A component, but firms U_1 and U_2 each compete to sell their own version of B: B_1 and B_2 respectively. D can only install one version of component B.[6]

[5] See also Lee (2013) and Pouyet and Trégouët (2016) for papers on vertical integration in multi-sided markets, the latter with a particular focus on the smartphone industry.

[6] The debate around bundling of smartphone applications has mostly focused on the manufacturer's choice of a default application (or on which application makes it onto the phone's home screen). Capacity is constrained because there can be only one default for each task and space on the home screen is limited. Jeon and Menicucci (2012) also study bundling in a setup where the buyer has a limited capacity. The difference between their model and ours is that the capacity constraint is over the whole set of

Our main motivating example is the market for smartphones (where components are pre-installed applications). In keeping with this motivation, we assume that component B_i generates a direct revenue nr_i for U_i when it is used by n consumers. This revenue may come from advertising, sale of consumer data to third parties, or "in-app purchases".[7]

Demand for the final product is $Q(p, S)$, where p is the price and $S \in \{\{B_i\}, \{A, B_i\}\}$ is the set of components installed by D.[8] We assume that, for any S, D's revenue function $pQ(p, S)$ is quasi-concave in p and maximized at p_S. We also assume $Q(p, \{A, B_1\}) = Q(p, \{A, B_2\})$ and $Q(p, \{B_1\}) = Q(p, \{B_2\})$—the two B components are perfect substitutes from consumers' perspective (this assumption is not essential but makes the exposition cleaner).

We write $\Pi \equiv p_{\{A,B_i\}} Q(p_{\{A,B_i\}}, \{A, B_i\})$ and $\pi \equiv p_{\{B_i\}} Q(p_{\{B_i\}}, \{B_i\})$ respectively for the profit when A is and is not installed alongside B.

The two key ingredients of our theory are retail complementarity and a contractual friction.

Retail complementarity We assume demand is such that

$$Q \equiv Q(p_{\{A,B_i\}}, \{A, B_i\}) > Q(p_{\{B_i\}}, \{B_i\}) \equiv q \quad \text{and} \quad \Pi > \pi.$$

In words: when component A is installed, (i) more consumers buy the finished good (ii) downstream sales revenue is larger.

Contractual friction Our final ingredient is a contractual friction that leaves upstream firms with a positive per-unit income from each consumer. To make the mechanism clear, we begin with a very simple such friction: upstream firms can only offer lump-sum transfers (implying that U_i earns r_i per consumer served). We write F_X for the lump-sum that the upstream producer of component X demands from D ($F_X < 0$ corresponds to a payment to D, i.e. a slotting fee).

Payoffs Given D's optimal choice of price conditional on S, firms' payoffs are as follows. If the downstream firm installs A and B_i, its profit is $V_D = \Pi - F_A - F_{B_i}$. If it only installs B_i, $V_D = \pi - F_{B_i}$. Firm U_1's profit if both A and B_1 are installed is $V_1 = F_A + F_{B_1} + r_1 Q$. If only B_1 is installed, $V_1 = F_{B_1} + r_1 q$. Firm U_2's profits is $V_2 = F_{B_2} + r_2 Q$ if B_2 is installed alongside A, and $V_2 = F_{B_2} + r_2 q$ if B_2 is installed without A.

products, whereas we impose a constraint on the B-applications only. More specifically, we don't allow the manufacturer to install B_1 and B_2 only, i.e., A never competes against the B applications.

[7] For brevity, we normalize application A's revenue to zero. But our analysis easily extends to positive revenues for A.

[8] For brevity we assume that component B is essential.

Timing and equilibrium The game proceeds as follows: At $t = 0$, U_1 announces whether it bundles A and B_1. At $t = 1$, upstream firms make simultaneous offers to the downstream firm. At $t = 2$ the downstream firm decides which component(s) to install, and chooses a final price. Payoffs are realized at $t = 3$. We restrict attention to subgame-perfect equilibria in undominated strategies. We study the two subgames without bundling and with bundling in turn.

3.1 Separate marketing

Let us start with the case where components A and B_1 are sold separately.

Lemma 1. *Suppose that $r_i \geq r_j$. Under separate marketing:*

i The downstream firm chooses components A and B_i in equilibrium.[9]

ii B_j's (rejected) offer is $F_{B_j} = -(Qr_j - \epsilon)$.[10]

iii The accepted offers are $F_A = \Pi - \pi$ and $F_{B_i} = -Qr_j$.

iv If $r_1 \geq r_2$, firm U_1's profit is $V_1 = \Pi - \pi + Q(r_1 - r_2)$. If $r_1 < r_2$, it is $V_1 = \Pi - \pi$. Firm U_2's profit is then $V_2 = Q(r_2 - r_1)$. In both cases the downstream firm's profit is $V_D = \pi + \min\{r_1, r_2\}Q$.

Proof. (i) Suppose $S = \{A, B_j\}$. B_j cannot offer a slotting fee above Qr_j as this would generate negative profits. But then there exists an F'_{B_i} that B_i can offer to D representing a Pareto improvement for the pair (e.g., $F'_{B_i} = -Qr_j - \epsilon$). A similar reasoning holds for A. (ii) Given $A \in S$, each U_k is willing to offer up to Qr_k. The standard logic of asymmetric Bertrand competition implies that the least efficient firm makes the best offer it could afford, in this case $F_{B_j} = -r_jQ$. (iii) Given $F_{B_j} = -r_jQ$, the downstream firm prefers to install A and B_i rather than B_i alone (denoted $\{A, B_i\} \succsim \{B_i\}$) iff $\Pi - F_A - F_{B_i} \geq \pi - F_{B_i}$. Similarly, $\{A, B_i\} \succsim \{B_j\}$ implies $F_A + F_{B_i} \leq \Pi - \pi - r_jQ$. Lastly, $\{A, B_i\} \succsim \{A, B_j\}$ requires $F_{B_i} \leq F_{B_j}$. Together, these constraints imply $F_A = \Pi - \pi$ and $F_{B_i} = -r_jQ$. (iv) Component A generates profit F_A for U_1; B_i generates profit $Qr_i + F_{B_i}$ for U_i; $V_D = \Pi - F_A - F_{B_i}$. ■

Under separate marketing, competition on the B market forces firms to offer slotting fees $F_{B_i} < 0$, and therefore to transfer part of the rent to the downstream firm.

On the A market, firm U_1 can capture the *direct* value it brings to the downstream firm, $\Pi - \pi$. Component A also brings some *indirect* value to the downstream firm, through

[9] If $r_i = r_j$ then there is also the mirror allocation.

[10] Here we assume that ϵ, small, is the minimal size of a price change. In the remainder of the paper we simplify notations by removing the ϵ. Without the minimal size assumption the equilibrium in undominated strategies would be such that firm j mixes over $(-Qr_j, -Qr_j + \epsilon)$ for ϵ small enough, leading to equivalent outcomes. See Kartik (2011).

B firms' increased willingness to pay slotting fees (from qr_i to Qr_i). However, U_1 cannot capture this indirect value. This is a key difference with standard models of bundling with complements, where, if consumption of A increases the utility from B by Δ, the A firm can charge $v_A + \Delta$ and therefore capture all its marginal value. To see why such a logic does not work here, suppose that $r_i = r_j = r$, and that $F_A = \Pi - \pi + r(Q - q)$ so that firm 1 fully captures the marginal value of A. The downstream firm would never agree to pay such a fee, as it would be better-off only buying from the B firm making the most generous offer.

As we now show, bundling the two components allows firm 1 to capture more of A's marginal value.

3.2 Bundling

Now let firm 1 bundle A and B_1 with a single transfer offer $\hat{F}_1 = \hat{F}_A + \hat{F}_{B_1}$. Thus, D is constrained to choose $S \in \{\{B_2\}, \{A, B_1\}\}$. Firm 1 would only bundle if it expects to be chosen by D; we thus restrict attention to this case. We have:

Lemma 2. *Under bundling:*

i U_2 *offers* $\hat{F}_{B_2} = -qr_2$;

ii Firm 1 offers $\hat{F}_1 = \Pi - \pi - qr_2$;

iii Firm 1's profit is $\hat{V}_1 = \Pi - \pi + Qr_1 - qr_2$. *The downstream firm's profit is* $\hat{V}_D = \pi + qr_2$.

Proof. (i) $F_{B_2} < -r_2 q$ is dominated: if it were accepted U_2's profit would be $r_2 q + F_{B_2} < 0$. Suppose $\hat{F}_{B_2} > -qr_2$ and firms do not expect B_2 to be installed. D must be indifferent between installing B_2 and the bundle (otherwise, U_1 could increase $\hat{F}_1$ a little). But that means that U_2 could reduce $\hat{F}_{B_2}$ and be installed for positive profit. (ii) Given $\hat{F}_{B_2} = -r_2 q$, D chooses the bundle if $\Pi - \hat{F}_1 \geq \pi + r_2 q$, yielding $\hat{F}_1$. (iii) U_1's profit is $\hat{V}_1 = \hat{F}_1 + r_1 Q$. D's profit is $\hat{V}_D = \Pi - \hat{F}_1$. ■

Bundling allows firm U_1 to extract the whole joint marginal value of components A and B_1 by keeping the downstream firm at its outside option $\pi + qr_2$. The key to understand this is that bundling reduces firm U_2's willingness to pay a slotting fee. Indeed, U_2 anticipates that, should B_2 be chosen, component A would not be installed. It is therefore only willing to offer up to $r_2 q$ to be installed.

Proposition 1. *If* $r_1 < r_2$, *firm 1 is better-off under bundling (i.e.* $\hat{V}_1 > V_1$*) if* $r_1 Q > r_2 q$. *If* $r_1 \geq r_2$, *firm 1 is always better-off under bundling than under separate marketing.*

The proof follows immediately as a corollary of Lemmas 1 and 2. When $r_1 < r_2$, bundling creates an inefficiency. The gain for U_1 stems from the weaker competition from

U_2, who now only bids $r_2 q$ instead of $r_2 Q$. Bundling is more likely to be profitable if (i) the inefficiency $(r_2 - r_1)$ is small, and (ii) component A is important to attract consumers ($Q - q$ is large).

When $r_1 \geq r_2$, there is no inefficiency associated with bundling. Because firm 2 is still less aggressive than under separate pricing, firm 1 can demand a larger fixed fee, and bundling is always profitable.

3.3 Discussion

Novelty of the mechanism That joint marketing of complementary products can increase profit is certainly not a new result. However the mechanism we highlight is, to the best of our knowledge, novel. Let us briefly discuss how it differs from established theories of joint marketing and bundling.

First, the increase in profit does not come from solving the double-marginalization problem (Cournot, 1838). This point is made clearer by our focus on lump-sum transfers: there are no pricing externalities between the products and joint control cannot be used to raise overall demand for them.

Second, bundling can also be profitable when there are no externalities, by reducing the level of heterogeneity in the population (Adams and Yellen, 1976; Schmalensee, 1984). Again, this is not what is driving our result: we only have one buyer (the downstream firm), and therefore no heterogeneity. Buyers' homogeneity also makes mixed-bundling redundant.

Third, our theory differs from the one offered by Whinston (1990). We do not rely on firm U_2 incurring entry costs (or other economies of scale). Indeed, while Whinston (1990)'s theory is one of entry deterrence, ours can also be interpreted as exclusion of an established rival. In particular bundling is profitable in the short run even if the rival does not exit immediately.

Timing and commitment The simultaneity of the offers plays a role in making bundling profitable. To see this, suppose that $r_2 > r_1$. If negotiation over component A occurred first, bundling would no longer be optimal: U_1 would offer a payment $F_A = \Pi - \pi + r_1(Q - q)$. In the second stage, both firms would offer $F_{B_i} = r_1 Q$ if the first period offer had been accepted, $F_{B_i} = r_1 q$ otherwise. U_1's profit would be $\Pi - \pi + r_1(Q - q)$, greater than the profit under bundling $\hat{V}_1$.

U_1 would therefore have incentives to push the negotiations over A early. Two points are worth mentioning here. First, the downstream firm would have the opposite incentives, and would do its best to accelerate the negotiations over B. Second, a strong degree of commitment is required for such a strategy to work: U_1 must commit not to make a subsequent offer at the start of the second period of negotiations if D has rejected the first offer. Given that details of the negotiations are secretly held most of the time, it would be hard for outsiders to observe a deviation from the commitment not to make a second offer, and therefore reputation *vis-à-vis* third parties is unlikely to help sustain this commitment.

Of course our model also requires a certain degree of commitment power by U_1, as do all models where pure bundling occurs in equilibrium: U_1 must be able to commit not to offer A on a stand-alone basis if D accepts B_2's offer. Unlike the type of commitment discussed above, reputation *vis-à-vis* third parties is more likely to help here: it would be fairly easy to observe that D has installed B_2 alongside A, and therefore that U_1 has reneged its commitment.

Side payments Would bundling still be profitable if upstream firms could contract with one another? This question is particularly relevant when B_2 is more efficient than B_1. Suppose accordingly that $r_2 > r_1$.

A first possibility is a contract whereby firm U_1 agrees not to offer B_1 to the downstream firm. For U_1 to accept such a contract, U_2 must offer a payment at least equal to $Qr_1 - qr_2$—the extra profit generated by bundling. If firm U_1 accepts, firm U_2 no longer needs to offer any positive payment to the manufacturer, and its profit is at least Qr_2, which is larger than $Qr_1 - qr_2$. Even though such a contract dominates bundling, it would likely be deemed anti-competitive.

A second possibility would be for U_2 to pay U_1 not to bundle A and B_1, without requiring it not to offer B_1. As before, firm U_1 must receive a payment at least equal to $Qr_1 - qr_2$ to accept. This time, though, firm U_2 still faces competition on the B market, and its profit is $V_2 = Q(r_2 - r_1)$ (see Lemma 1). Therefore, when $2Qr_1 > (Q + q)r_2$, U_2 cannot profitably induce firm U_1 to unbundle A and B_1.

4 More general contracts

We now allow upstream firms to offer more general contracts, in the form of two-part tariffs. Under a tariff $T_i = (w_i, F_i)$, D pays $nw_i + F_i$ to the producer of component i if it chooses to install it and if the final demand is n.

4.1 Frictionless contracting

The timing is as follows: at $t = 0$, U_1 publicly announces whether it bundles A and B_1 or not. At $t = 1$, U_1 and U_2 offer two-part tariffs to D. A $t = 2$, D selects the set of components it installs, and chooses a final price p. At $t = 3$ payoffs are realized.

Unlike fixed fees, the level of the unit fees w affects the optimal price chosen by D. If D installs components A and B_i, the joint profit of the involved firms would be maximized by setting $w_A = 0$ and $w_{B_i} = -r_i$, so that D's price reflects the true marginal cost of the vertical structure.[11] We denote this maximal joint profit by Π_i,[12] and Q_i is the corresponding quantity sold given that the price is chosen optimally. If D installs only B_i, the optimal unit fee is again $w_{B_i} = -r_i$, and the corresponding joint profit and quantity are denoted π_i and q_i.

Notice that in any equilibrium where D installs A and B_i the joint profit must equal Π_i.

We make the following set of assumptions:

Assumption 1. *If $r_i \geq r_j$, we have:*

i $\Pi_i \geq \Pi_j$, $Q_i \geq Q_j$, $\pi_i \geq \pi_j$ *and* $q_i \geq q_j$.

ii $\Pi_i - \pi_i \geq \Pi_j - \pi_j$

iii $\Pi_j \geq \pi_i$ *and* $Q_j \geq q_i$

By part (i), the most efficient component facilitates higher sales and a larger joint profit. Part (ii) means that adding A to the product is more valuable if the chosen B component is the most efficient one. Part (iii) implies that the asymmetry between B_1 and B_2 is not too large compared to the value of installing A.

Our first result is a negative one:

Proposition 2. *Bundling A and B_1 is not profitable if upstream firms can offer two-part tariffs.*

The proofs of this section appear in the online appendix. Intuitively, competition in two-part tariffs leads firms to offer the efficient level of the unit fee, $w_{B_i} = -r_i$ and $w_A = 0$. Competition therefore only takes place with respect to the fixed fees. But this set-up is equivalent to one in which the "single monopoly profit theory" applies: when B_2 is more efficient than B_1, U_1 can charge a higher price for product A if it does not bundle it with B_1.

4.2 Upstream moral hazard

We now discuss the profitability of bundling when some contracting friction prevents firms from designing contracts that achieve the joint first-best. For our purpose, any friction leading to a positive upstream mark-up ($w_{B_i} > -r_i$) would work; we focus on moral hazard.

Suppose that, after D has chosen which B component to install, the selected upstream firm can exert a non-contractible effort that increases the final demand.[13] Such effort could

[11] If $r_i > 0$ the marginal cost of B_i is negative.

[12] i.e., $\Pi_i = \max_p\{(p + r_i)Q(p, \{A, B_i\})\}$.

[13] Only the supplier of the B-component can exert such effort. Later we discuss the possibility of investment by the A supplier.

consist of advertising or product improvement. A two-part tariff such that $w_i = -r_i$ would leave U_i with no incentives to exert effort, because its profit would be independent of the number of units sold. Equilibrium contracts should therefore involve positive upstream markups so as to induce effort.

To keep notations simple, we focus on the following technology: effort is binary with cost $e \in \{0, 1\}$, and a positive effort increases demand by Δ. We assume that a positive level of effort is always desirable.

The timing is the following: at $t = 0$, U_1 publicly announces whether it bundles A and B_1 or not. At $t = 1$, U_1 and U_2 offer two-part tariffs to D. A $t = 2$, D selects the set of components it installs. At $t = 3$ the supplier of the selected B component chooses whether to exert effort. At $t = 4$, D observes the level of effort and chooses a final price p.

Optimal fee and notations If D has opted for component B_i, U_i finds it optimal to exert effort if and only if $(w_{B_i} + r_i)\Delta \geq 1$. Therefore, assuming that it is optimal to induce effort by U_i, the unit fee that maximizes the joint profit of D and its suppliers is $w_{B_i} = -r_i + 1/\Delta$. Any smaller value leads to no effort; larger values exacerbate the double-marginalization problem. After payment of the unit fees, the B supplier is therefore left with a revenue of n/Δ if n units are sold.

We define Π_i, π_i, Q_i, and q_i as the downstream profits (excluding the fixed fees), and the associated quantities, with and without A, when $w_{B_i} = -r_i + 1/\Delta$ and U_i exerts effort.[14] Let $\tilde{\Pi}_i$, $\tilde{Q}_i$, $\tilde{\pi}_i$ and $\tilde{q}_i$ be the corresponding objects when $w_{B_i} = -r_i$ and U_i does not exert effort. We maintain Assumption 1, and assume that the value of component A is not reduced when the B supplier exerts effort.

Assumption 2. *For $i = 1, 2$, $\Pi_i - \pi_i \geq \tilde{\Pi}_i - \tilde{\pi}_i$.*

For the sake of brevity we only present results for the case where $r_2 > r_1$, implying bundling is inefficient.

4.2.1 Bundling

Because $w_{B_i} > -r_i$, upstream profits depend on the number of consumers served. Thus, as in Section 3, bundling limits the slotting fees offered by U_2 by decreasing demand when B_2 is installed.

Lemma 3. *There is a unique equilibrium under bundling, in which U_2 is foreclosed and U_1's profit is $\Pi_1 - \pi_2 + (Q_1 - q_2)/\Delta - 1$.*[15]

[14] i.e. $\Pi_i \equiv \max_p (p + r_i - \frac{1}{\Delta})\,(Q(p, \{A, B_i\}) + \Delta)$, etc.

[15] The term -1 is the cost of effort.

In equilibrium both upstream firms offer the efficient unit fee that induces effort, $w_i = -r_i + 1/\Delta$. U_2's losing bid offers all the joint profit (without A), $\pi_2 + q_2/\Delta$, to D. U_1's offer makes D indifferent between $\Pi_1 - F_1$ and $\pi_2 + q_2/\Delta$, and U_1 gets the mark-up $1/\Delta$ for the Q_1 units sold.

4.2.2 No bundling

There is now a multiplicity of equilibria in the subgame without bundling, some of which deliver outcomes that are similar to the equilibrium under bundling.[16]

Lemma 4. *Suppose that $r_2 > r_1$. In the model with upstream moral hazard and two-part tariffs, there are two types of equilibria.*

1. ***Efficient equilibria**, such that D installs $\{A, B_2\}$, always exist. Firm U_1's profit ranges from $\Pi_1 - \pi_1$ to $\Pi_2 - \pi_2$.*

2. *There also exist **inefficient equilibria**, i.e. such that D installs $\{A, B_1\}$, whenever $(Q_1 - q_2)/\Delta - 1 \geq \Pi_2 - \Pi_1$. U_1's profit ranges from $\Pi_2 - \pi_2$ to $\Pi_1 - \pi_2 + (Q_1 - q_2)/\Delta - 1$.*

In an efficient equilibrium, unit fees are $w_A = 0$ and $w_{B_i} = -r_i + \frac{1}{\Delta}$. The logic is then similar to Lemma 1: U_2 anticipates that D will also install A and is therefore willing to offer a large slotting fee (up to Q_2/Δ). More specifically, the best equilibrium for U_1 has $F_A = \Pi_2 - \pi_2$, $F_{B_2} = \pi_2 - \pi_1 - \frac{Q_1}{\Delta}$ and U_1's rejected offer for B_1 is $F_{B_1} = -\frac{Q_1}{\Delta}$.

In an inefficient equilibrium, U_1 adjusts the unit fees so as to make it unprofitable for D to install B_2 alongside A, while keeping $w_A + w_{B_1}$ at the efficient level. In effect, firm 1 creates a virtual bundle through its choice of contracts. Anticipating this, U_2 is no longer willing to offer a large slotting fee. One strategy profile that sustains U_1's preferred equilibrium is: $w_A = r_2 - r_1$, $w_{B_1} = -r_2 + \frac{1}{\Delta}$, $F_A = \Pi_1 - \pi_2$ and $F_{B_1} = -\frac{q_2}{\Delta}$. U_2's rejected offers are $w_{B_2} = -r_2 + \frac{1}{\Delta}$ and $F_{B_2} = -\frac{q_2}{\Delta}$.[17]

The next Proposition is obtained as a corollary from Lemmas 3 and 4.

Proposition 3. *When $(Q_1 - q_2)/\Delta - 1 > \Pi_2 - \Pi_1$, the unique equilibrium under bundling delivers the same profit to U_1 as the best equilibrium under no bundling.*

When $(Q_1 - q_2)/\Delta - 1 < \Pi_2 - \Pi_1$, bundling is not profitable for U_1.

With two-part tariffs and upstream moral hazard, explicitly bundling A and B_1 is no longer necessary to foreclose B_2. The value of (explicit) bundling comes from the first-mover advantage it gives to U_1, allowing it to select its preferred equilibrium.

[16] The multiplicity of equilibrium payoffs comes from the fact that the binding constraint on the fixed fees paid to D only pins down $F_A + F_{B_i}$.

[17] Off the equilibrium path, if U_2 offers $F_{B_2} < -\frac{q_2}{\Delta}$, D installs B_2 alone even though it is indifferent with installing B_2 and A. In the proof we construct an equilibrium that does not rely on this tie-breaking assumption.

Discussion of moral hazard with A Our assumption that the effort only concerns producers of the B component is less innocuous than our assumption that A does not generate any revenue. Indeed, with moral hazard on both markets there would be an efficiency argument for having B_1 instead of B_2: a mark-up on A (necessary to induce effort on the A component) would reduce the need for a further markup on B_1, but not on B_2, to induce effort. This logic is similar to the logic of double marginalization in the pricing of complements. While it would make the analysis of the game much more intricate, it would not affect the key insight that bundling reduces B_2's willingness to offer slotting fees. In terms of welfare, bundling would be less likely to be inefficient, given that, provided r_2 is not too large compared to r_1, the efficiency gains from having a single upstream provider (outlined just above) would offset the fact that $r_2 > r_1$.

5 Conclusion

Upstream bundling can reduce rivals' willingness to pay slotting fees and thereby enable profitable leverage. This can be achieved as the unique equilibrium through strict bundling, or as one equilibrium among many with appropriately designed contracts.

A motivation for our analysis is the case of smartphone application bundling. In this market consumers can modify the downstream firm's offering by installing alternative applications. It is fairly straightforward to allow this in our model. Bundling can continue to be profitable, provided some consumers will not change the default application configuration (because, e.g., they have high switching costs, they are indifferent between applications, or they suffer from default bias).

Though motivated by the Android case, our model can be applied more broadly. First, observe that other markets share similar institutional features to smartphones. For instance, upstream cable TV networks offer bundles of channels ('components') to downstream cable companies and earn advertising revenue when their channel is viewed. Thus, our work speaks to ongoing policy concerns around wholesale bundling in the pay-TV market (see Crawford, 2015, for a discussion).

Secondly, the model can also be used to study bundling by manufacturers in standard retail supply chains. Recall that our analysis depends on two assumptions: retail complementarity and contractual frictions that give rise to slotting fees. If consumers value one-stop shopping then a downstream retailer attracts more customers by stocking more products; our retail complementarity assumption is then satisfied. Moreover, the analysis of Section 4 is unchanged if we let $r_i < 0$ (interpreted as an upstream manufacturer's marginal cost of production). Thus, positive wholesale mark-ups and slotting fees offered to retailers endogenously arise under contractual frictions as before. Given that our assumptions are satisfied, we again find bundling by a manufacturer can foreclose a rival by denying them the chance to be stocked alongside important products.

Our setup involves a downstream monopolist. With downstream competition, bundling by U_1 has the potential to prevent downstream firms from differentiating by offering different versions of the B product, which may intensify competition. Exploring this issue is a promising research avenue.

References

Adams, William James and Janet L Yellen (1976). "Commodity Bundling and the Burden of Monopoly". *Quarterly Journal of Economics* 90.3, pp. 475–498.

Aghion, Philippe and Patrick Bolton (1987). "Contracts as a Barrier to Entry". *The American economic review*, pp. 388–401.

Amelio, Andrea and Bruno Jullien (2012). "Tying and Freebies in Two-Sided Markets". *International Journal of Industrial Organization* 30.5, pp. 436–446.

Burstein, Meyer L (1960). "A Theory of Full-Line Forcing". *Northwestern University Law Review* 55.1, pp. 62–95.

Calzolari, Giacomo, Vincenzo Denicolò, and Piercarlo Zanchettin (2016). "Exclusive dealing with costly rent extraction".

Caprice, Stéphane and Vanessa Schlippenbach (2013). "One-Stop Shopping as a Cause of Slotting Fees: A Rent-Shifting Mechanism". *Journal of Economics & Management Strategy* 22.3, pp. 468–487.

Carbajo, Jose, David De Meza, and Daniel J Seidmann (1990). "A strategic motivation for commodity bundling". *Journal of Industrial Economics* 38.3, pp. 283–298.

Carlton, Dennis W. and Michael Waldman (2002). "The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries". *RAND Journal of Economics* 33.2, pp. 194–220.

Chen, Yongmin (1997). "Equilibrium product bundling". *Journal of Business* 70.1, pp. 85–103.

Choi, Jay Pil (2010). "Tying in two-sided markets with multi-homing". *The Journal of Industrial Economics* 58.3, pp. 607–626.

Choi, Jay-Pil and Doh-Shin Jeon (2016). "A Leverage Theory of Tying in Two-Sided Markets". Working Paper.

Choi, Jay Pil and Christodoulos Stefanadis (2001). "Tying, investment, and the dynamic leverage theory". *RAND Journal of Economics* 32.1, pp. 52–71.

Chu, Wujin (1992). "Demand signalling and screening in channels of distribution". *Marketing Science* 11.4, pp. 327–347.

Cournot, Antoine-Augustin (1838). *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot*. chez L. Hachette.

Crawford, Gregory S. (2015). "The Economics of Television and Online Video Markets". *Handbook of Media Economics*. Ed. by Simon P. Anderson, Joel Waldfogel, and David Strömberg. Amsterdam: Elsevier, pp. 267–339.

Director, Aaron and Edward Hirsch Levi (1956). “Law and the Future: Trade Regulation”. *Northwestern University Law Review* 51, pp. 281–296.

Edelman, Benjamin and Damien Geradin (2016). “Android and Competition Law: Exploring and Assessing Google’s Practices in Mobile”. *European Competition Journal* 12.2–3, pp. 159–194.

Farrell, Joseph (2001). “Some thoughts on slotting allowances and exclusive dealing”. *Washington, DC: US Department of Justice*.

Foros, Øystein and Hans Jarle Kind (2008). “Do slotting allowances harm retail competition?” *The Scandinavian Journal of Economics* 110.2, pp. 367–384.

Fumagalli, C, M Motta, and C Calcagno (2018). “Monopolization: A Theory of Exclusionary Practices Cambridge University Press”.

Ho, Katherine, Justin Ho, and Julie Holland Mortimer (2012). “The use of full-line forcing contracts in the video rental industry”. *The American Economic Review* 102.2, pp. 686–719.

Ide, Enrique and Juan-Pablo Montero (2016). “Bundled Discounts and Monopolization in Wholesale Markets”. Working Paper.

Jeon, Doh-Shin and Domenico Menicucci (2012). “Bundling and competition for slots”. *The American Economic Review* 102.5, pp. 1957–1985.

Kartik, Navin (2011). “A note on undominated Bertrand equilibria”. *Economics Letters* 111.2, pp. 125–126.

Lee, Robin S (2013). “Vertical integration and exclusivity in platform and two-sided markets”. *The American Economic Review* 103.7, pp. 2960–3000.

Marx, Leslie M and Greg Shaffer (2010). “Slotting allowances and scarce shelf space”. *Journal of Economics & Management Strategy* 19.3, pp. 575–603.

Nalebuff, Barry (2004). “Bundling as an Entry Barrier”. *Quarterly Journal of Economics* 119.1, pp. 159–187.

O’Brien, Daniel P. and Greg Shaffer (2005). “Bargaining, Bundling, and Clout: The Portfolio Effects of Horizontal Mergers”. *RAND Journal of Economics* 36.3, pp. 573–595.

Pouyet, Jérôme and Thomas Trégouët (2016). “Vertical Integration in Platform Markets”. Working Paper.

Rasmusen, Eric B, J Mark Ramseyer, and John S Wiley Jr (1991). “Naked exclusion”. *The American Economic Review*, pp. 1137–1145.

Rey, Patrick and Jean Tirole (2007). “A Primer on Foreclosure”. *Handbook of Industrial Organization, Volume 3*. Ed. by Mark Armstrong and Robert Porter. Amsterdam: Elsevier, pp. 2145–2220.

Schmalensee, Richard (1984). “Gaussian Demand and Commodity Bundling”. *Journal of Business* 57.1, S211–230.

Segal, Ilya R and Michael D Whinston (2000). “Naked exclusion: comment”. *The American Economic Review* 90.1, pp. 296–309.

Shaffer, Greg (1991a). “Capturing strategic rent: Full-line forcing, brand discounts, aggregate rebates, and maximum resale price maintenance”. *The Journal of Industrial Economics*, pp. 557–575.

— (1991b). “Slotting allowances and resale price maintenance: a comparison of facilitating practices”. *The RAND Journal of Economics*, pp. 120–135.

— (2005). “Slotting allowances and optimal product variety”. *The BE Journal of Economic Analysis & Policy* 5.1.

Stigler, George J (1963). “United States v. Loew’s Inc.: A note on block-booking”. *The Supreme Court Review* 1963, pp. 152–157.

Whinston, Michael D (1990). “Tying, Foreclosure, and Exclusion”. *American Economic Review* 80.4, pp. 837–859.

A Proof of Proposition 2

(1) Case with $r_2 > r_1$. Suppose that U_1 bundles A and B_1. Let $T_1 = (w_1, F_1)$ be U_1's offer, with $w_1 = -r_1$.

First, in equilibrium, U_2 must offer $w_{B_2} = -r_2$ and $F_{B_2} = 0$. Indeed, D must be indifferent between $\{A, B_1\}$ and $\{B_2\}$, and if $w_{B_2} \neq -r_2$ than U_2 could profitably deviate and induce D to choose $\{B_2\}$. Given that $w_{B_2} = -r_2$, we obtain $F_{B_2} = 0$ using standard weak dominance arguments.

Given U_2's offer, U_1's accepted offer must then satisfy $\Pi_1 - F_1 = \pi_2$ for D to be indifferent between $\{A, B_1\}$ and $\{B_2\}$. U_1's profit is then $\hat{V}_1 = \Pi_1 - \pi_2$.

Suppose instead that U_1 chooses not to bundle A and B_1 and sets $w_A = 0, w_{B_1} = -r_1$ and $F_{B_1} = 0$ (i.e. it makes the best possible offer for B_1). For D to choose $\{A, B_2\}$, three conditions must hold: (i) $F_{B_2} \leq \Pi_2 - \Pi_1$ (so that D prefers $\{A, B_2\}$ to $\{A, B_1\}$), (ii) $F_A \leq \Pi_2 - \pi_2$ (so that D prefers $\{A, B_2\}$ to $\{B_2\}$), and (iii) $F_A + F_{B_2} \leq \Pi_2 - \pi_1$ (so that D prefers $\{A, B_2\}$ to $\{B_1\}$). The worst configuration for U_1 is when constraints (i) and (iii) are binding. In this case its profit is $V_1 = F_A = \Pi_1 - \pi_1$, which is still larger than $\hat{V}_1$. Bundling is therefore not profitable.

(2) Case with $r_1 > r_2$. Under bundling, B_2's rejected offer must be $w_{B_2} = -r_2$ and $F_{B_2} = 0$. The profit of U_1 is therefore equal to the maximal fee it can charge D, i.e. $\hat{V}_1 = \Pi_1 - \pi_2$.

If U_1 does not bundle its products and offers $w_A = 0$ and $w_{B_1} = -r_1$, then D installs $\{A, B_1\}$ in equilibrium. Again, B_2's rejected offer must be $w_{B_2} = -r_2$ and $F_{B_2} = 0$. The constraints that F_A and F_{B_1} must satisfy are (i) $F_{B_1} \leq \Pi_1 - \Pi_2$ (so that D prefers $\{A, B_1\}$ to $\{A, B_2\}$), (ii) $F_A \leq \Pi_1 - \pi_1$ (so that D prefers $\{A, B_1\}$ to $\{B_1\}$), and (iii) $F_A + F_{B_1} \leq \Pi_1 - \pi_2$ (so that D prefers $\{A, B_1\}$ to $\{B_2\}$). By Assumption 1(2), constraint (iii) is binding, so that $V_1 = \Pi_1 - \pi_2 = \hat{V}_1$.

B Proof of Lemma 3

If U_1 bundles A and B_1, in equilibrium D must be indifferent between $\{A, B_1\}$ and $\{B_2\}$ (otherwise U_1 could demand higher fixed fees). B_2's rejected offer must be $w_{B_2} = -r_2 + 1/\Delta$ and $F_{B_2} = -q_2/\Delta$: $w_{B_2} = -r_2 + 1/\Delta$ maximizes the joint profit, and $F_{B_2} = -q_2/\Delta$ allocates all the profit to D. Lower values of F_{B_2} are dominated strategies, while higher values could not constitute an equilibrium (U_2 could reduce F_{B_2} and profitably induce D to install B_2).

In equilibrium U_1 must offer $w_1 = -r_1 + 1/\Delta$, so that the maximal fixed fee it can charge is given by $\Pi_1 - F_1 = \pi_2 + q_2/\Delta$. U_1's profit is therefore $\hat{V}_1 = F_1 + (r_1 + w_1)Q_1 - 1 = \Pi_1 - \pi_2 + (Q_1 - q_2)/\Delta - 1$.

C Proof of Lemma 4

Efficient equilibria First, in an efficient equilibrium, we must have $w_A = 0$ and $w_{B_2} + r_2 = 1/\Delta$ to maximize the realized joint profit. w_{B_1} is not uniquely pinned down in equilibrium. For our purpose, we can focus on equilibria where the rejected B_1 offer would have induced effort if accepted, i.e. $w_{B_1} = -r_1 + 1/\Delta$. Let F_{B_1} be the rejected offer's fixed fee.

For D to select $\{A, B_2\}$ rather than respectively $\{A, B_1\}$, $\{B_2\}$ or $\{B_1\}$, we must have (i) $F_{B_2} \leq \Pi_2 - \Pi_1 + F_{B_1}$, (ii) $F_A \leq \Pi_2 - \pi_2$ and (iii) $F_A + F_{B_2} \leq \Pi_2 - \pi_1 + F_{B_1}$. By Assumption 1(3), (iii) is always binding. There is then a continuum of (F_A, F_{B_2}) compatible with (i)-(iii). U_1's associated profit ranges from $\underline{V_1}^E \equiv \Pi_1 - \pi_1$ (when (i) also binds) to $\overline{V_1}^E \equiv \Pi_2 - \pi_2$ (when (ii) also binds). Let us check that these constitute equilibria of the subgame without bundling.

Let us take a (F_A, F_{B_2}) compatible with (i)-(iii). Neither D nor U_2 have a profitable deviation from such a strategy profile. Could U_1 profitably deviate? The only possibility would be to make offers such that D chooses $\{A, B_1\}$. One constraint would then be that D prefers $\{A, B_1\}$ to $\{B_2\}$, i.e. $\Pi_1 - F'_A - F'_{B_1} \geq \pi_2 - F_{B2}$. Because (iii) is binding, we have $F_{B2} = \Pi_2 - \pi_1 + F_{B1} - F_A$. Therefore the deviation must satisfy $\Pi_1 - F'_A - F'_{B_1} \geq \pi_2 - (\Pi_2 - \pi_1 + F_{B1} - F_A)$. Now, we know that in an $\{A, B_2\}$ equilibrium, U_1's profit V_1 is equal to F_A. So the previous constraint rewrites as $\Pi_1 - \pi_1 + \Pi_2 - \pi_2 + F_{B1} - V_1 \geq F'_A + F'_{B_1}$. The best deviation by U_1 is therefore to make this constraint binding. Its new profit is then $F'_A + F'_{B_1} + Q_1/\Delta = \Pi_1 - \pi_1 + \Pi_2 - \pi_2 + F_{B1} - V_1 + Q_1/\Delta$. The deviation is not profitable if $\Pi_1 - \pi_1 + \Pi_2 - \pi_2 + F_{B1} - V_1 + Q_1/\Delta \leq V_1$ i.e. if $2V_1 \geq \Pi_1 - \pi_1 + \Pi_2 - \pi_2 + F_{B1} + Q_1/\Delta$.

To sustain $V_1 = \underline{V_1}^E$ as an equilibrium, we must have $F_{B_1} \leq \Pi_2 - \pi_2 - (\Pi_1 - \pi_1) - Q_1/\Delta$. This is not ruled out by weak dominance, since weak dominance only rules out $F_{B_1} < -Q_1/\Delta$. Therefore any $V_1 \in [\Pi_1 - \pi_1, \Pi_2 - \pi_2]$ can be sustained in an efficient equilibrium.

Inefficient equilibria Take ϵ arbitrarily close to zero and consider the following strategy profile: $w_A = r_2 - r_1 + \epsilon$, $F_A = \Pi_1 - \pi_2 - \epsilon q_2$, $w_{B_2} = -r_2 + \frac{1}{\Delta}$, $F_{B_2} \in [\Pi_2 - \Pi_1 - \frac{Q_1}{\Delta} + 1 + \epsilon q_2, \frac{-q_2}{\Delta}]$, $w_{B_1} = w_{B_2} - \epsilon$, $F_{B_1} = F_{B_2}$.

D's profit if it installs $\{A, B_1\}$ is $\Pi_1 - F_A - F_{B_1} = \pi_2 + \epsilon q_2 - F_{B_2}$. If it installs $\{A, B_2\}$, its profit is $\Pi_1 - \epsilon Q_1 - F_A - F_{B_2} = \pi_2 - \epsilon Q_1 - F_{B_2}$. If it installs B_1 alone, its profit is $\pi_2 + \epsilon q_2 - F_{B_2}$. If it installs B_2 alone, its profit is $\pi_2 - F_{B_2}$. So D chooses $\{A, B_1\}$ whatever the value of F_{B_2}.

The key aspect of U_1's strategy is that (w_A, F_A) are chosen such that D always strictly prefers $\{B_2\}$ to $\{A, B_2\}$ for any value of F_{B_2}. Therefore, given that $F_{B_2} \leq \frac{-q_2}{\Delta}$, U_2 is not willing to increase the slotting fee it offers (i.e. to offer $F'_{B_2} < F_{B_2}$) because it would lose money by doing so.

Under this strategy profile, U_1's profit is $V_1 = F_A + F_{B_1} + \frac{Q_1}{\Delta} - 1 = \Pi_1 - \pi_2 - \epsilon q_2 + F_{B_2} + \frac{Q_1}{\Delta} - 1$. The best possible deviation for U_1 would be to induce D to install

$\{A, B_2\}$ by choosing $w'_A = 0$ (so as to maximize the joint profit) and $F'_A = \Pi_2 - \pi_2$ (along with high prices for B_1). The resulting profit would be $V'_1 = \Pi_2 - \pi_2$. When $F_{B_2} \geq \Pi_2 - \Pi_1 - \frac{Q_1}{\Delta} + 1 + \epsilon q_2$ such a deviation is not profitable.

As the possible equilibrium values of F_{B_2} cover the interval $[\Pi_2 - \Pi_1 - \frac{Q_1}{\Delta} + 1 + \epsilon q_2, \frac{-q_2}{\Delta}]$, V_1 goes from $\Pi_2 - \pi_2$ to $\Pi_1 - \pi_2 + \frac{Q_1 - q_2}{\Delta} - 1 - \epsilon q_2$.

Incorporating Cyber-Physical Systems in the Global Cyber Regime Complex

Dr. Mark Raymond and Dr. Laura DeNardis[1]

The complexity and high stakes of global cyber policy problems escalate significantly as the Internet moves out of display screens and information systems and diffuses into the physical world all around us. The so-called Internet of Things (IoT) becomes entangled with even the most seemingly unrelated cyber policy areas. To use one of the most far-afield examples, what does the IoT have to do with cryptocurrency? The original idea behind Bitcoin was to shift the role in administering currency transactions from a trusted financial institution to cryptographic proof, essentially math.[2] The blockchain technology underlying Bitcoin and other cryptocurrencies relies upon computationally intensive calculations distributed over a peer-to-peer network. Solving these math problems keeps the system operational and also allows so-called "Bitcoin miners" to be issued bitcoin in exchange for this service. This computational incentive system has led to cryptojacking, in which hackers infiltrate and take over devices to hijack electrical power, central processing unit (CPU) power and graphics processing unit (GPU) power in order to mine cryptocurrency. Cybercriminals implant cryptocurrency-mining malware on insecure computing devices and steal these resources without the consent or knowledge of the device owners. Because of the relatively insecure features of IoT devices, cryptocurrency-mining malware is targeting these systems.[3] IBM security researchers even uncovered a variant of the Mirai malware (that

[1] Mark Raymond (mraymond@ou.edu), Director, Cyber Governance and Policy Center; Wick Cary Assistant Professor of International Security, Department of International and Area Studies, University of Oklahoma
Laura DeNardis (denardis@american.edu), Professor, School of Communication; Faculty Director, Internet Governance Lab, American University
The authors would like to thank the Carnegie Corporation of New York and the Tech and Policy Initiative at the School of International and Public Affairs at Columbia University for their support of this research. We are also grateful to Josie Smith and Asvatha Babu for their research assistance.

[2] See the 2008 white paper by the person or people known as Satoshi Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System." Accessed at https://bitcoin.org/bitcoin.pdf.

[3] See, for example, Indian Computer Emergency Response Team bulletin, 2018, "Cryptocurrency-mining Malware Targeting IoT Devices." Accessed at https://cert-in.org.in.

previously exploited IoT devices to create what was then the most massive DDoS attack in history) "with a new twist: a built-in bitcoin mining component" (McMillen 2017).

This opaque connection between the IoT and cryptocurrency helps demonstrate a number of contemporary features of the IoT. First, cyber-physical devices, especially inexpensive consumer-oriented IoT devices used in homes, are widely understood to be insecure. Nearly all cybersecurity expert communities – from the National Institute of Standards and Technology to the United States Department of Homeland Security – have warned about the security vulnerabilities in IoT devices. In part because of the rush to bring new products to market in this quickly evolving consumer product sector, adequate security is not designed at product inception. Many devices lack upgradeability and contain vulnerabilities that malware can be exploit. This raises much more consequential problems than cryptojacking because vulnerabilities in physical world objects can affect human safety and basic societal functioning. The cryptojacking example also illuminates another feature: invisibility. As the Internet moves into the everyday background objects of society – from cars to RFID system in packages to home lighting systems – the presence of cyberspace becomes less visible. Screens are no longer the arbiter of whether something is "on the Internet." This embedding into real-world objects complicates every area of cyber policy. For example, how can consent for data collection practices be reasonably accomplished without screen interfaces, or when those in the proximity of the device are not the device owner and may not even be aware of this data collection? This one connection between the IoT and cryptocurrency also lays bare how IoT governance is not only about the IoT but about other critical public policy issues from financial policy to environmental protection.

IoT systems inherently have a physical presence in real world objects, so it is tempting to relegate the public policy issues they raise as squarely within the borders of local or national concern and jurisdiction. In 2017, we published an article arguing that cyber-physical systems present pressing, cross-border global policy issues that cannot be adequately addressed solely at the local and national levels. These issues include critical Internet resource constraints, privacy complications, human security concerns, international security and issues pertaining to global economic and industrial policy (DeNardis and Raymond 2017).

On the basis of this work, in June 2018, we had the opportunity to help organize the State of the Field Workshop on the Digital Transformation at the Columbia University School of International and Public Affairs (SIPA). This workshop, convened by the Tech and Policy Initiative with the support of the Carnegie Corporation of New York, convened leading scholars, policymakers and practitioners to consider how governance arrangements might be adapted to the digital age, and what kinds of new arrangements are needed to cope with rapid developments at the frontier of digital technologies.

The workshop considered a range of policy challenges including the role of online platforms in governance, the effects of digital transformation on the international trade regime, complications associated with cryptocurrencies and digital payment systems, the management of digital identities, and critical questions surrounding elections and democracy in the digital age. While each of these topics is important in its own right, in this paper we will report and extend the workshop's findings on the policy challenges associated with the "Internet of Things". We do so by comparing and contrasting these findings with the argument advanced in our earlier article, in order to present an analysis of areas of agreement and disagreement on the challenges and solutions associated with the rapid adoption and diffusion of the IoT throughout the public and private sectors. This analysis sets the stage for the remainder of the paper, which assesses the stakes and examines the extent to which contemporary cyber norms processes and emerging cyber peace movements provide vehicles for addressing at least some of the most important global public policy challenges.

We note the existence of several high-profile initiatives to develop cyber norms and to pursue various visions of "cyber peace", briefly describing each of these ongoing processes and providing analysis of their potential to contribute to improving global governance of cyber-physical systems in the context of an emerging global cyber regime complex (Nye 2014). The global public policy challenges associated with the Internet of Things are increasingly urgent, and their rising stakes require timely action to mitigate actual and potential damage arising from global governance gaps and failures. The paper concludes by identifying several key issues to be addressed in order to ensure that IoT global governance challenges are adequately addressed within and beyond the broader global cyber regime complex.

Global Public Policy Challenges Associated with Cyber-Physical Systems

Discussion at the SIPA State of the Field Workshop advanced thinking on the global public policy challenges associated with the mass deployment of cyber-physical systems in several respects. The discussion generally built on and extended the assessments we arrived at in our 2017 article. Participants agreed that it is useful to understand IoT policy challenges as problems of global governance rather than via the predominant existing frames, which see them either as a series of consumer harms issues or more broadly as issues for national and subnational legislation and regulation. There are several reasons that a conceptual frame drawn from the study of global governance is particularly useful.

First, while integration of cyber-physical systems in consumer products constitutes an important and growing part of the IoT landscape, it is far from the entirety of the way in which these technologies are being used. Participants in the workshop generally concurred with our assessment that industrial and municipal applications are vital parts of the IoT landscape, and that these kinds of applications present distinct policy challenges. For example, where consumer products often have relatively short lifespans and are replaced relatively frequently, cyber-physical systems in industrial and municipal settings tend to remain in use for longer time periods. There also tend to be few mechanisms to provide continued support for security updates to hardware, firmware and software. This is especially the case in contexts where cyber-physical systems include components purchased off-the-shelf and added to devices that originally lacked this functionality, either by traditional manufacturers or by operators retrofitting legacy systems.

Second, there was clear recognition by participants that cyber-physical systems are being designed, manufactured, deployed and operated in inherently transnational ways. The design and manufacture of cyber-physical systems, as with most other goods in advanced manufacturing sectors, takes place within globally-integrated supply chains. Finished products combine intellectual property and intermediary components from a number of different countries. As such, finished products will be subject to various different actual or potential legal and regulatory requirements that may affect the use of encryption and other security-related technologies. Efforts to mandate minimum standards or otherwise ensure product and supply chain security will therefore be difficult to ensure at the national or subnational levels. Inclusion of components from jurisdictions that mandate backdoors or other means to ensure government access to communications or other forms of data will frustrate efforts to ensure high security standards. Where particular kinds of components are produced only by a small number of suppliers, certain countries may exercise outsized influence on best-available products and services. For these reasons and others, there are clear benefits from international coordination of relevant regulatory and legal frameworks, as well as from coordination in what are often highly privatized technical standard-setting bodies (DeNardis 2009; Büthe and Mattli 2011). Likewise, jurisdictions with globally significant market share like the European Union and the United States may be able to drive product development by setting conditions for market access that either ensure or undermine product security. However, if such large jurisdictions adopt mutually incompatible standards for market access, firms may be compelled to develop separate product lines for different markets or else make difficult choices about whether or not they will maintain market presence in different jurisdictions.

Similar conditions exist for firms wishing to deploy and operate various kinds of cyber-physical systems. While governments are likely to operate these systems primarily within their national territories,[4] multinational firms will increasingly need to ensure that cyber-physical systems they wish to operate across multiple legal jurisdictions are in regulatory compliance in each of those jurisdictions. Beyond these compliance costs, the development of incompatible legal and regulatory requirements may require the deployment of different systems within various jurisdictions. Such requirements may also complicate the formatting, retention and use of data originating from different countries for business operations. At present, the governance landscape for cyber-physical systems is underdeveloped (Weber 2016). However, given the increased emphasis on Internet governance and cyber policy issues on the part of states (Bradshaw et al. 2015; Cowhey and Aronson 2017; Demchak and Dombrowski 2013; DeNardis 2014), this state of affairs is unlikely to continue.

One possibility is that firms engaged in the design, manufacture, deployment and operation of cyber-physical systems across all sectors of the economy will continue to operate according to a key cultural tenet of the Internet economy – to “move fast and break things”, by bringing products and services to market with scant and inconsistent regard for the existence of legal or regulatory frameworks. In doing so, firms should be aware that at least some large states are increasingly willing to subject such business practices to scrutiny under regulatory frameworks relating to privacy, competition, and national security. On the one hand this means that late adopters are likely to face higher initial compliance costs and more complex regulatory scrutiny. On the other hand, early adopters are likely to face investigation and potential fines arising from choices of design, manufacture, deployment and operation of cyber-physical systems that were made in the absence of tailor-made regulations for such technologies. They will face this scrutiny, at first, under the

[4] Military applications are an important exception to this tendency. However, it is likely that states would interpret the operation of such systems as falling within broad latitude provided by international law for self-defense and for collective action with the aim of ensuring international peace and security.

extension of cognate bodies of regulation that are extended to cover cyber-physical systems;[5] over time, the creation of technology-specific legal and regulatory frameworks is more likely.

It is worthwhile to note that the extension of regulatory scrutiny is likely to be more expensive for the designers, manufacturers and operators of the Internet of Things than it is for online service providers. This is simply because legal and regulatory scrutiny may limit the legality of certain kinds of functionality built into these systems. One example is the inclusion of devices that collect certain kinds of information, such as voice recordings or location data attributable to individuals. Change in legal and regulatory frameworks may require not just change in computer code, which can sometimes be executed relatively quickly and deployed at scale in an inexpensive manner in the context of a smartphone or other computer application, but changes in physical devices such as the removal of cameras or microphones to eliminate the possibility that such sensors can be remotely activated in a clandestine manner.

Third, concern with the potential of cyber-physical systems to enable mass surveillance was a consistent theme throughout the SIPA State of the Field Workshop. Beyond the potential for flagrant human rights harms similar to those in the Khashoggi case (Kirkpatrick 2018; Marczak et al 2018), a deeper privacy concern is the effect of mass deployment of cyber-physical systems on everyday privacy, including in public spaces (O'Connor et al 2017; Neisse et al 2017). Such technologies raise important questions about the viability of the notice-and-consent model for privacy protection employed in the European Union's General Data Protection Regulation (European Union 2016). If collection of audio, video and geolocation data becomes even more ubiquitous than it is at present, it will become effectively impossible for individuals to opt out of such surveillance. Given the transnational nature of the Internet economy, it is likely that such systems will collect, transmit and store data about individuals across international borders, subjecting their data to the privacy protection laws and regulations of multiple states – often in ways that individuals cannot be reasonably expected to either understand or consent to in advance.

The deployment of such systems creates various kinds of actual and potential legal problems both for the individuals being tracked and for the executives of the firms that collect and retain the data. One possibility is that states may attempt to enforce their own domestic laws against their own citizens for conduct those citizens have committed in countries where the conduct in question was legal. (Allen-Ebrahimian 2018) An example of such a scenario would be an authoritarian state enforcing laws against political protest when the protest takes place outside the territory of the citizen's home state (Zhongsun, 2018). Another possibility is that states may attempt to utilize their lawful access provisions to mandate the disclosure of data held abroad by foreign firms, and that they may seek to arrest representatives of the firm who enter their jurisdictions. Such a case arose in 2016, when Brazil arrested Facebook's top executive in Latin America in order to compel disclosure of data for the purposes of a criminal investigation (Mastroianni 2016).

The mass adoption of cyber-physical systems creates and exacerbates particular kinds of human rights and human security (Paris 2001) concerns for individuals. Some such harms, such as the violation of the right to privacy affirmed in the International Covenant on Civil and Political Rights (United Nations 1966), are created simply by virtue of certain forms of data collection, acquisition

[5] The extension of pre-existing rule frameworks in this manner is consistent with constructivist scholarship on International Relations. See Sandholtz (2008); Brunnée and Toope (2010).

and retention by state actors. Other such harms, such as political imprisonment, torture and killing, can be enabled by large-scale data collection facilitated by cyber-physical systems.

As the Khashoggi case, the Facebook case, and others make clear, these concerns are not unique to cyber-physical systems. However, while these kinds of legal issues have primarily been the concerns of large global online service providers, the metastasization of data collection practices throughout virtually every sector of the economy via the adoption and deployment cyber-physical systems will create policy concerns surrounding privacy and data collection for a large range of firms that have not previously dealt with these kinds of regulatory, civil and even criminal liability. Furthermore, such firms will have to deal with demands of this kind not only from their own national governments but potentially from the governments of any state in which their technology is deployed and operated. Governments seeking to protect the human rights and human security of their own citizens and permanent residents will also increasingly need to be concerned with the potential for foreign governments to more easily acquire information in violation of privacy rights and to employ that information to enable other violations of rights to physical freedom and safety. Accordingly, these kinds of problems cannot be addressed by any single state acting independently, and will require global coordination. Weber (2016, p. 10) suggests that such coordination will need to take a bottom-up rather than top-down form given the extensive role of private actors in the Internet of Things landscape. While we agree that bottom-up policy development is a necessary part of ensuring adequate governance of cyber-physical systems, we also conclude that there are certain policy areas where states are likely to insist on a primary role for reasons of economic policy and national security, and where more traditional diplomatic and legal cooperation among states are therefore likely to remain important.

Fourth, consistent with the argument in our article, workshop participants recognized that the mass deployment of the Internet of Things creates particular national and international security concerns. One such concern is that the incorporation of cyber-physical systems across a variety of economic sectors massively expands the critical infrastructure base. The United States Department of Homeland Security includes the information technology sector and the communications sectors as two of its sixteen designated dimensions of critical infrastructure.[6] As such, any product or good that includes an Internet-connected sensor or control component appears to qualify as critical infrastructure. Given the likelihood that critical infrastructure sectors will be subject to increased regulatory burdens covering product design and manufacture, as well as deployment and operation, these governance arrangements create responsibilities for the national security state across virtually the entirety of the economy. They should also be expected to create problematic interactions among other existing rule sets for governing disparate policy areas such as human rights and international trade. The invocation of national security exceptions to restrict free trade rules or override human rights protections are examples of these kinds of interactions. The extension of such policies to any product incorporating an Internet-connected sensor or control capability could significantly undermine the openness of the international economy as well as the practical enjoyment of human rights such as privacy and free expression. Apart from such outcomes, the increased density of applicable international rules will create a significant volume of work in interpreting and applying such rules, with concomitant potential for disputes about how to do so that could themselves become sources of international disputes.

[6] See https://www.dhs.gov/cisa/critical-infrastructure-sectors.

Aside from the regulatory and governance burdens associated with these kinds of rules, the deployment of cyber-physical systems at scale also creates other, more familiar security problems. One problem is the expanded size of the attack surface in cyberspace, since any Internet-connected device could become the target of an attack. A second problem is that the mass deployment of cyber-physical systems will vastly increase the number of Internet-connected devices. If not properly secured, these devices are susceptible to recruitment into botnets and can be used to attack other targets. The Mirai botnet demonstrates both the potential for devices comprising cyber-physical systems to become targets of cyberattacks and also the potential for such devices to be enlisted to amplify the scale of attacks on other kinds of targets by incorporating them into botnets.[7]

A fifth kind of global governance concern that arose from the SIPA State of the Field Workshop went beyond the concerns addressed in our 2017 article. Specifically, a participant noted[8] that the increasing reliance of policymaking processes on "Big Data" could undermine the quality of public policy in the event that policymakers are reliant on poor quality data that provide partial and incomplete information or that have particular blindnesses. One such concern centers around the digital divide; the community of global Internet users is growing, and also growing more diverse, but roughly half of all people currently alive are still excluded. Digital data is often understood as representing only the portion of the human population that has Internet access. However, it is important to recognize here that it can more broadly be understood as being of two types: data we volunteer about ourselves on the internet and the traces we leave through our presence online; and data that is collected about us through sensors on cyber-physical systems without our specific or active input. People without access to the internet may be represented only by data of the second type until they become Internet users themselves; and, as such, policy made with an overreliance on data from Internet technologies will continue to under-represent a large number of citizens of emerging economies and the broader Global South for the foreseeable future. Issues of poor data quality go beyond this basic issue of who is and is not represented in the data. Databases are prone to error and this aspect only gets magnified as the database gets larger or if multiple databases are combined. Making big data-driven policy gives rise to other complications as well. Manovich (2011) points to institutional inequalities in the access to big data and the biases this produces in the collection of data. In addition, Boyd and Crawford (2014) discuss ethical issues and the problem of apophenia (seeing patterns where there are none). All of these could be exacerbated by the massive increase in data collected as cyber-physical systems become more commonplace.

Collectively, these global policy challenges associated with the mass adoption of cyber-physical systems complicate established modalities for accomplishing governance according to territorially-demarcated sovereign states (Ruggie 1993), including international cooperation among those states and the agents they delegate to. Cyber-physical systems call these familiar governance modalities into question because the emergence of transnational cyber-physical systems across various sectors of the economy raises fundamental questions about how to establish and divide jurisdiction over processes, institutions and organizations that shape important aspects of individuals' lives. The fundamental reality is that digital governance issues are and will continue

[7] For background on the Mirai botnet, see https://www.csoonline.com/article/3258748/security/the-mirai-botnet-explained-how-teen-scammers-and-cctv-cameras-almost-brought-down-the-internet.html.

[8] Since the workshop was conducted on the basis of the Chatham House Rule, we do not identify the participant in question either by name or affiliation.

to be governed in a decentralized manner by a large number of actors from a variety of classes of actor types (Raymond and DeNardis 2015; Raymond 2016).

Relevant governance actors include a wide array of firms, such as network operators, companies offering online services of various kinds such as large social media platforms, hardware manufacturers, and software developers. These traditional kinds of firms are rapidly being supplemented by firms that we are not used to considering in digital policy conversations, and that have limited experience thinking of themselves in these terms. These non-traditional private sector participants in Internet governance and Internet policy include insurance providers, financial institutions, health services providers, retailers and – most important for our purposes here – a wide range of manufacturers producing products that are beginning to incorporate Internet connectivity as part of cyber-physical systems. These include companies producing wearables and home appliances, but also include firms that manufacture a variety of sensor and control devices intended to improve manufacturing, energy production and transport, and various forms of public sector infrastructure such as street lights, water systems, and public transportation. In the public sector, relevant governance actors include foreign and defense ministries, but increasingly include subnational actors such as local public safety agencies; cities and municipalities; attorneys and judges; and public utilities. The emergence of public interest issues pertaining to information and communications technologies, including cyber-physical systems, has encouraged the emergence of civil society actors that lobby and advocate for their preferred policy solutions and governance arrangements.

The decentralized nature of these governance arrangements is problematic because the nature of the Internet enables decisions made by one actor to create externalities for other actors. In this respect, Internet governance and Internet policy resembles a series of nested clubs (Raymond 2013/2014). In the remainder of this section, we show that this nested club perspective is helpful for understanding the inescapably global nature of Internet governance and Internet policy, including the issues raised by cyber-physical systems. We then briefly sketch how highly decentralized governance arrangements for a global communications and control facility complicate the emerging global cyber regime complex.

At the most basic level, all Internet users[9] are members of three kinds of groups or, in the terminology employed in social science, 'clubs' (Buchanan 1965): the club of all global Internet users, the club of Internet users from a particular country, and the club of Internet users that utilize a specific Internet service provider. In practice, users are also members of a host of other clubs by virtue of their access to workplace or educational computer networks; their accounts with email, social media, retail and other online service providers; and various other kinds of restricted access groups. Each of these clubs has its own rules and rule-making processes, though some are much more formalized than others, and some have much more participatory and democratic governance mechanisms than others.

For the most part, these various clubs operate relatively autonomously on a day-to-day basis; however, this autonomy is in some respects illusory and in any event has tended to diminish over

[9] We employ the word 'user' in an expansive sense here, referring not just to individuals but to firms, governments and their agencies, and non-profits or non-governmental organizations. The same club, or group, might include all of these different kinds of actors, or might be comprised of a more restrictive set of members.

roughly the past decade. It is illusory because access to the Internet requires compliance with the basic technical standards and protocols that define the Internet at the logical layer, as well as access to specific kinds of physical infrastructure and connection agreements with network operators. Virtually all specific clubs of Internet users are rule-takers in these respects – dependent for continued access on a relatively small number of firms and (largely Western) technical experts that provide connection services and core over-the-top services, and who define the technical standards and terms of service applicable to smaller, subsidiary online clubs. The Internet, therefore, inscribes and reproduces certain forms of power relations. While online power has been highly privatized since the mass commercialization of the Internet in the late 1990s, this outcome was the product of state choice rooted in specific ideas about the proper role of the state vis-à-vis the market (Strange 1996) and facilitated by the fact that public authorities tended to have little initial understanding of the technology. Over the last decade, the autonomy enjoyed by most online clubs has declined as states have become less willing to leave Internet policy to firms and the technical community, and as governments have developed more sophisticated capacities online, for example with respect to cyber espionage and the military use of ICTs.

As a result of these trends as well as increased levels of global Internet penetration, many online actors and entities are now more likely to experience various forms of disruption and conflict in the course of their routine operations. These include: (1) obviously malicious activity like phishing campaigns, ransomware and Distributed Denial-of-Service (DDoS) attacks; (2) inadvertent degradation or disruption of service, as a result of damage to physical infrastructure or attempts by specific network operators or other authorities to prevent access to certain online content by the users they govern; and (3) deliberate attempts to influence the operation of other clubs by the exercise of various forms of power. This last category of disruptions and conflicts includes cases where one club is relatively clearly subject to the authority of the other, such as where a state asserts authority over individuals or firms within its jurisdiction, and where a firm exerts the private power of a service provider to set the terms it offers to individuals or groups. It also includes cases where no such clear relationship exists rooted in rules or agreements that entitle one club to make demands on another. This set of cases includes instances where a state demands or forbids access to data held by a foreign government, or by a firm, organization or individual within the territorial jurisdiction of another state. While mutual legal assistance treaties (MLATs) and other mechanisms for international law enforcement cooperation (such as the Budapest Convention on Cybercrime) exist, these mechanisms have important limitations. MLATs are inadequate to handle the current scale and scope of requests for lawful access to data, and are not meant to accommodate requests to limit access to data for example on security grounds, such as a 2018 request by British intelligence that an American scholar restrict access to a personal blog about radical Jihadist groups (Prothero 2018). The Budapest Convention is limited in its geographic reach and, while its approach of committing states to harmonizing domestic computer crime laws is a valuable step toward facilitating the operation of the MLAT regime and extradition treaties, it is not a panacea for resolving conflicts or disputes in which states seek access to data held in other jurisdictions, or other similar kinds of international disputes that are already endemic to the Internet and that will only become more serious as global Internet penetration increases. Similarly, emerging international security norm candidates pertaining to the state military use of ICTs (which necessarily affect other states and their associated clubs of Internet users) remain in flux given the inability of the Group of Governmental Experts to reach consensus on the applicability of the law of armed conflict in this domain and given recent developments in the UN General Assembly's First Committee more broadly (UN 2018). In any case, these norms include only a candidate norm forbidding peacetime attacks on critical infrastructure and do not make direct reference to cyber-physical systems as such (UN 2013; UN 2015).

The mass adoption of cyber-physical systems will amplify these potential sources of disruption and conflict that are already inherent in highly networked societies tied together by a global communications and control facility that is governed in a highly decentralized manner. At the most basic level, it will create a multitude of new clubs, such as the users and administrators of public transit or electrical grid or municipal water systems that are dependent on IoT devices for key regulation and control functions. Integrating these kinds of core municipal government functions more closely with the Internet will involve new actors in navigating the global cyber regime complex, for example in seeking assistance for incident response or simply in managing network security and network traffic. The rapid involvement of a large number of often poorly-resourced, novice players in the operation of the Internet ecosystem creates risks of instability and accidents. It also creates a large number of vulnerable critical infrastructure targets. Beyond these kinds of harms associated with the integration of cyber-physical systems throughout the real economy and critical infrastructure, the mass adoption of cyber-physical systems also creates harms associated with vast increases in the scale and scope of surveillance, since it entails the collection of data in more intimate and less visible ways, and means that this data is held in the hands of a large number of players that may be unaccustomed to roles as collectors and stewards of highly personal information. In this sense, the mass adoption of cyber-physical systems has the effect of creating, and widely distributing, large amounts of what Bruce Schneier has called a “toxic asset” (Schneier 2016).

In short, the mass adoption of cyber-physical systems will greatly expand the global cyber regime complex. In addition to increasing the number of players, it will more tightly enmesh legacy arrangements for global Internet governance and for traditional international security matters with an array of subnational actors traditionally far-removed from these kinds of concerns. The global cyber regime complex, then, will be not simply a collection of formerly disparate *international* regimes;[10] rather, it will increasingly become a multilevel governance arrangement spanning the individual through global levels of analysis. The diffusion of cyber-physical systems will result in the metastasization of the global cyber regime complex. In the process, the global cyber regime complex will become critical *governance* infrastructure in the same way that the Internet itself has become critical physical infrastructure. If the global cyber regime complex is compromised as a legitimate and/or effective venue for global rule-making, interpretation and application (Raymond 2019), states and other public authorities will increasingly find themselves unable to meet their citizens’ demands for the effective delivery of public goods not simply at the global level but also at the domestic level.

The Rising Stakes of Cyber-Physical Systems for Democracy, Society and the Economy

In the contemporary era, most cyber policy attention is still directed at information and communication systems. How should social media companies and law respond to Russian social media influence campaigns preceding the United States Presidential election in 2016? Can state

[10] This is the traditional conception of a regime complex in the IR literature; see Raustiala and Victor (2004).

voter rolls, and voting infrastructure itself, be adequately secured in democratic elections? Given massive data breaches that continue to become public – from Equifax to Marriott – how can individual privacy be protected and risk mitigated? How can content-centric issues such as hate speech, cyberbullying, and propaganda be addressed and by whom? What are the prospects for free speech and global media freedom in light of censorship campaigns and cybersecurity attacks that disrupt speech platforms. Given that all economic sectors depend upon digital information systems, what cybersecurity practices are necessary to preserve economic stability? Who governs technical infrastructure? All of these are critical societal concerns. None of them on their surface, appear to have anything to do with the Internet of Things.

We argue that IoT policy concerns are not only entangled with but actually escalate all of these traditional cyber governance concerns. Indeed, the so-called Mirai botnet demonstrated how the security and stability of these information systems is only as strong as the security and stability of the IoT. This botnet in the fall of 2016 caused a massive distributed denial of service (DDoS) attack that disrupted some of the popular information platforms including Twitter and Reddit. The attack was carried out by a network of insecure consumer IoT devices that were infected and hijacked by a piece of malware called Mirai. The majority of embedded physical devices connect – either directly or indirectly via a gateway – to the public Internet. They exist on the same networks as information servers, smart phones, and other communication devices. Hacking into a factory's HVAC system or a farm's automated infrastructure system can provide access to customer data and other critical information if they are stored on the same network. The security of all networks is reliant on the resilience of connected objects that have both cyber and physical components. (Magrani and Lemos 2018)

Cyber-physical systems similarly escalate the stakes of international cyber conflict. While some may still view cyber conflict and real-world war as distinct, international security is an area that has long laid bare connections between information systems and cyber-physical systems. Ukrainian industrial sites, and in particular energy systems, have for years experienced debilitating infrastructure attacks attributed to Russian hackers (ICS-CERT 2016).[11] Nearly a decade ago, the Stuxnet worm targeting the Iranian nuclear program demonstrated this connection between digital and material elements. The worm was extremely sophisticated code designed to sabotage the Siemens control systems operating Iranian nuclear centrifuges (Zetter 2014).

One irony is that the further cyber devices move away from people, and away from human display interfaces, the more human rights issues these devices raise. As one of the participants in the Columbia IoT session explained, these are simultaneously more intimate and also further away. The diffusion of the Internet into everyday objects raises significant privacy questions about the data collected and the surveillance enabled within the most intimate spheres of human existence (Rosner 2016). How is consent gained and what are the limits to data collection in systems that rely on massive data collection and feedback systems for their very operation? Privacy becomes more complicated and data collection more invasive. An even greater human rights concern relates to consumer safety, with foreign adversaries and hackers able to reach across borders to sabotage

[11] See generally Indus. Control Sys. Cyber Emergency Response Team (ICS-CERT), *Alert (IR-Alert-H-16-056-01): Cyber-Attack Against Ukrainian Critical Infrastructure*, U.S. Department Homeland Security, (Feb. 25, 2016), https://ics-cert.us-cert.gov/alerts/IR-ALERT-H-16-056-01.

systems. The human safety dimensions of all cyber-embedded objects are only as strong as the cybersecurity of these devices, known to be, as Bruce Schneier has described it, "wildly insecure."

The stakes already mentioned – the stability of the public sphere and the digital economy, international security, and human rights including privacy and safety – are already real concerns. It is easy to speculate into the future about how these points of vulnerability and influence could potentially threaten systems of democracy. It may no longer be necessary to disrupt systems of election infrastructure. Disrupting power systems or municipal "smart city" transportation systems on election day in micro-targeted districts would have the same disruptive effects. Email phishing attacks, such as the ones that led to hacking into DNC emails, rely upon subversion operations that trick the target into believing in the legitimacy of the email. This form of subversion becomes much more believable when it includes highly personal information gleaned from cyber-physical systems touching every part of human existence. Highly targeted, believable, and personal information raises the potency of phishing attacks. There are no publicly known examples of IoT-connected attacks on democracy infrastructure, but it is inevitable that this will become a potential mechanism of disruption, similar in effect to social media influence campaigns, hacking of voter rolls, and infiltration of political email accounts.

The stakes could not be higher for securing and protecting cyber-physical systems. In the same way these consequences and the technologies themselves cross borders, so must the solutions cross borders. Cyber-physical governance requires coordination and cooperation at the global level, as these problems cannot simply be solved at the national and subnational levels by any particular country.

Global Cyber Governance and Norms Initiatives

The ongoing process of establishing the global cyber regime complex remains in an early stage, and while the existence of regime complexes has been documented in a number of substantive issue-areas,[12] relatively little is known either about the process of regime complex formation or about regime complex dynamics. Accordingly, care should be taken in identifying ways to deal with global policy challenges arising from the mass adoption of cyber-physical systems within the context of the global cyber regime complex. However, it is equally true that given the current pace of adoption for these technologies and their global implications outlined above, there will ultimately be no choice – participants in the global cyber regime complex will need to confront these challenges to make the continued operation of the Internet possible, and the nature of the technology will enmesh an extremely large number of individuals and organizations in that regime complex.

For these reasons, we argue that existing and future processes for clarifying, creating and changing norms and institutions in the global cyber regime complex must deal more explicitly with the global public policy challenges emanating from the mass adoption of cyber-physical systems. The first requirement in doing so is to adopt a conceptual understanding of the Internet that extends

[12] Raustiala and Victor (2004); Betts (2010); Keohane and Victor (2011); Colgan, Keohane and Van de Graaf (2012); Nye (2014);

beyond received wisdom portraying it as an information and communications system to encompass its role as a vital global system for the management and control of a range of physical devices and machines extending far beyond traditional IT systems. We identify key governance issues pertaining to cyber-physical systems in more detail in the final section of the paper. In this section, we identify a number of current processes and venues in which these issues might be more fully and directly considered.

A complete survey of cyber norms processes is beyond the scope of this paper. Rather, we identify five very recent and ongoing examples that we believe to have potential in addressing these issues. Additionally, we have selected examples to illustrate processes from various regions of the world and that include a variety of stakeholder types. Some processes are state dominated, others are largely driven by firms, while some are best described as civil society and multistakeholder processes. Crucially, we believe that these processes have value over the medium to long term that is partially independent of their ability to generate immediate results or agreements on specific norms. This is because the process itself is crucial to the ongoing management of Internet and cybersecurity governance (Finnemore and Hollis 2016), very much including governance issues pertaining to the mass adoption of cyber-physical systems.

The first notable cyber norms process that could profitably take cyber-physical systems into more direct consideration is the long-standing effort in the United Nations General Assembly's First Committee on Disarmament and International Security. The First Committee has been engaged in deliberation and study on the state military use of ICTs since the introduction of a Russian resolution in 1998 (Maurer 2011; Tikk-Ringas 2012). Russia and a growing group of co-sponsors initially maintained that international law had no mechanisms to govern the state military use of ICTs, in order to justify a multilateral treaty that took an expansive view of "information security" understood to encompass an asserted right of governments to take drastic surveillance and blocking measures in order to ensure regime stability. These views continue to inform action by Russia, China and a group of like-minded states both in the UN and via the Shanghai Cooperation Organization (SCO), which has released a voluntary Code of Conduct pertaining to cybersecurity.

In 2013 and 2015, however, the Group of Governmental Experts (GGE) process overseen by the First Committee released consensus reports that included all permanent members of the UN Security Council as well as other key states. In these reports, states acknowledged that international law applies to the state military use of ICTs, specifically enumerating several bodies of international law including the law of sovereignty, human rights law, and the law of state responsibility which deals with breaches of international obligations falling below the threshold of the law of armed conflict (Raymond 2019). In addition, these reports acknowledged the utility and importance of confidence-building measures for cybersecurity and advanced several candidate norms such as immunity from cyberattacks for critical infrastructure and for incident response entities (UN 2013; UN 2015).

While the 2017 iteration of the GGE ultimately did not produce a consensus report, the process did not invalidate the earlier reports, which survive as important indicators of state thinking on the applicability of international law to cybersecurity. Unfortunately, in its 2018 work process, the First Committee passed two competing resolutions on the basis of recorded votes. An American-led resolution endorses the 2013 and 2015 GGE reports in their entirety and calls for the creation of a new, enlarged GGE with augmented opportunities for consultation with other states and with outside experts (United Nations 2018b). In contrast, the Russian and Chinese-led resolution appears to more selectively endorse the outcomes of the 2013 and 2015 GGE processes in a manner intended to narrow the UN referential baseline for cyber norms going forward. This second resolution also calls for an open working group to negotiate, rather than study, norms for state military use of ICTs. These differences reflect growing confidence and familiarity on the part of Russia, China and their allies with procedures for making, interpreting and applying global rules (Raymond 2019). They are also likely to increase contention over cybersecurity governance issues within the UN in an unhelpful and counterproductive fashion, since the shift from study to negotiation is likely to raise perceptions of sovereignty costs associated with adverse outcomes, leading to harder bargaining.

No matter which path the First Committee takes in its deliberations on cybersecurity issues, it is important that states carefully consider the implications associated with the mass adoption of cyber-physical systems. Existing norms in the GGE work product pertaining to critical infrastructure and supply chain security provide potentially useful starting points for such deliberations. Again, it is worth noting that the possibility of immediate agreement is not crucial to determining whether these discussions should take place, if for no other reason than that there is no possible set of rules that will definitively solve the global policy challenges associated with cyber-physical systems. Rather, the challenge is to ensure adequate mechanisms for ongoing processes of rule-making, interpretation and application. Such processes must also include robust conflict-resolution mechanisms to deal with all-but-inevitable disputes about what the appropriate rules are and how to properly apply them to specific cases.

States have also begun to engage with cyber norms processes at the regional rather than the global level. In addition to the highly problematic regional effort underway in the SCO, the Association of Southeast Asian Nations (ASEAN) has also made strides in norm development pertaining to cybersecurity (ASEAN 2018). ASEAN efforts are noteworthy as they provide a potential regional counterweight to the SCO effort that may be at least modestly more likely to ensure key human rights protections and Internet openness.[13] Asia is also globally significant in terms of the current and especially the future Internet user base, as well as in terms of the ICT industry. Existing ASEAN efforts, as indicated by the statement issued at the 32nd ASEAN summit in April 2018, emphasize the need for regional cooperation on cybersecurity as well as the importance of information sharing among states and the inherently transboundary nature of cyber issues (ASEAN 2018). However, the statement contains no specific commitments and no specific references to cyber-physical systems or to the Internet of Things. Within ASEAN, as within other regional efforts, there is a need to go beyond broad statements of principle and to attempt to arrive at more specific measures to deal with cybersecurity governance challenges, including those pertaining to the mass adoption of cyber-physical systems.

The Global Commission on the Stability of Cyberspace (GCSC) was created by The Hague Center for Strategic Studies and the EastWest Institute. It is most accurately described as a civil society venture, though its website acknowledges support from the Dutch, Singaporean, French and Estonian governments, as well as from firms including Microsoft and from other civil society

[13] It is worth noting here, however, that ASEAN itself has a track record of deference to state sovereignty, and has little independence from its member-states compared to other international organizations. See Acharya (2014).

actors such as the Packet Clearing House, the Internet Society, and Black Hat USA.[14] The GCSC has provided a platform for advancing a number of potentially beneficial norms, perhaps most notably the notion of a norm against interfering in the "public core" of the Internet (Broeders 2017). This public core includes physical and logical resources that are important to the stability and security of cyber-physical systems as well as to other parts of the Internet ecosystem. In a recent "norm package" released in November 2018, the GCSC advances a norm against interference with products prior to their release, in order to protect supply chain security, and a norm against states commandeering ICT assets outside their borders for use in botnets (Global Commission on Internet Governance 2018). Both of these norms are relevant to global policy challenges associated with the mass adoption of cyber-physical systems, and both would have positive effects if states could be persuaded to comply with them. While the GCSC's roster of commissioners includes several former high-ranking policymakers from a number of states, these individuals are now acting in their private capacities and therefore have limited influence over state policy. Accordingly, expectations for these kinds of processes should be modest.[15] Entities like the GCSC can play useful complementary roles but should not be expected to generate major progress on their own.

In addition to state and civil society efforts to develop and publicize norms within the global cyber regime complex, there are also a set of notable processes emerging from the private sector. In this regard, Microsoft stands out for its innovative leadership efforts though it is increasingly attempting to lead by cultivating broader support for its ideas within the private sector. In February 2017, Microsoft executive Brad Smith took the unusual step of publicly calling for a multilateral treaty that would act as a "Digital Geneva Convention" (Smith 2017). An associated policy paper outlined several core commitments that Microsoft called upon states to adopt. These included commitments to refrain from attacking critical infrastructure and global economic systems. These aspects of the Internet ecosystem are notable for their inclusion of cyber-physical systems. The paper also called on states to show restraint in developing and proliferating cyber weapons. Though the paper did not define the term, most common understandings of the category would include malicious code aimed at disabling or gaining unauthorized access to cyber-physical systems (Microsoft 2017).

Along with its call for a multilateral treaty, Microsoft has also played a leading role in the creation and expansion of the Cybersecurity Tech Accord, a private sector agreement that commits participating firms to refusing to assist governments (including their own) from conducting cyberattacks against civilian targets (Cybersecurity Tech Accord 2018). Though the agreement has attracted considerable attention and the subsequent adherence of a large number of other major Internet firms, a range of crucial questions remain. Notably, it is not clear how the participating firms will define assistance, cyberattacks or civilian targets. Presumably simply providing IT hardware and software to governments will not be interpreted by participating firms as violating the commitment should governments use those assets to conduct such attacks, since government procurement is an important segment of the global IT market. It is similarly unclear whether efforts to 'prepare the battlefield' by implanting malicious code in case of armed conflict would count as conducting an attack since there would be no manifest effect prior to the onset of hostilities. And,

[14] See https://cyberstability.org/.

[15] Here we note that we speak from experience having played key contributing roles in the Secretariat of a similar venture, the Global Commission on Internet Governance, with which the GCSC shares several of its commissioners.

finally, while there are well-established rules and procedures in international law for distinguishing civilian from combatant targets, technology firms lack experience or expertise in making such determinations and it is not clear that participating firms envision acquiring these kinds of capabilities. Finally, and most important, it is not clear how states will respond to efforts by major technology firms to resist their requests for assistance (Raymond and Smith 2018). Several countries have passed or are contemplating the passage of laws that would enable public authorities to compel firms to assist them in conducting cyber operations (Karp 2018). While the accord itself does not specifically mention cyber-physical systems, the Cybersecurity Tech Accord website is being used to host a variety of webinars to educate the public about cybersecurity issues. Some of these webinars pertain to Internet of Things issues and illustrate the potential of the Cybersecurity Tech Accord to consider cyber-physical systems policy and governance issues more explicitly.[16]

Finally, Microsoft has partnered with the French government to develop and promote the multistakeholder Paris Call for Trust and Security in Cyberspace, which was presented to the Internet Governance Forum in November 2018. The document has attracted more than 370 signatories within a month of its initial release.[17] Its principles align with those advanced by many of the other initiatives we have identified. Broadly, they call for stakeholders of various kinds to protect the Internet, critical and electoral infrastructure, and intellectual property. They also call for actors to pursue improved cyber-hygiene and to comply with cyber peace norms. Like the other processes identified here, however, the Paris Call includes little direct reference to issues associated with cyber-physical systems beyond a reference to the importance of ensuring the security of devices and processes throughout their lifecycle and at all points in the supply chain (Paris Call 2018).

Looking to the Future

Cyber governance is often reactive rather than proactive. Edward Snowden's disclosures about NSA surveillance prompted global attention action around privacy and the limits of private data collection and aggregation. Russian attempts to influence foreign elections have prompted attention to social media influence campaigns, misinformation, fake news, and deep fakes. It should not take a shocking incident involving cyber-physical systems to prompt similar global attention and action. This paper has explained the high public policy stakes of the cyber-physical systems. Systems with both cyber and physical components raise human safety issues far beyond digital-only devices, and they radically expanding the national security threat plane. Adversarial nations and networks can easily reach across borders to disrupt or surveil the physical world. The types of privacy concerns that the GDPR addresses in regard to cyber systems are far more complicated and pervasive in cyber-physical systems. We have also argued that the Internet of Things, even while involving physically local devices, is not solely a local concern, but a cross-border global policy concern. The technical complexity of IoT devices, and their integration of both cyber and physical components, significantly complicates governance questions around

[16] For access to webinars, see https://cybertechaccord.org/webinar-series/.

[17] For ongoing list of signatories, see https://www.diplomatie.gouv.fr/en/french-foreign-policy/digital-diplomacy/france-and-cyber-security/article/cybersecurity-paris-call-of-12-november-2018-for-trust-and-security-in.

everything from globally integrated supply chain security to technically embedded intellectual property rights. A minimal first step is for global norms processes to address the cyber-physical disruption. There is not yet adequate attention to these issues in global norms processes.

However, it is also important to recognize that global norms processes, while important, are insufficient because they are divorced from the lower level and smaller scale layers of governance in practice, across all actor classes. To date, the tendency has been to treat the global cyber regime complex as a global level phenomenon. A world in which local and municipal governments, and firms across every industrial sector, operate large numbers of Internet-connected sensors and control devices and manage large pools of sensitive personal data fundamentally challenges this assumption, particularly given the relatively oligopolistic nature of critical industrial sectors such as cloud-based data storage. Subnational governments and small firms from countries across the world will have little choice but to engage with large foreign firms in order to maintain basic functionality, and so will become increasingly dependent upon the operation of the global cyber regime complex that is essential to the ongoing stability and operation of the Internet. As a result, to the extent it is still employed as an organizing concept, it is necessary to treat the global cyber regime complex as an instance of multilevel, rather than simply global, governance.

It is also critical to acknowledge that there are multiple views about the normative and institutional architecture appropriate for Internet governance and Internet policy, with cyber sovereignty approaches on the rise and western liberal approaches on the decline. There has long been a tension between private sector-led, multistakeholder Internet governance approaches and cyber-sovereignty approaches that favor multilateral or authoritarian governance. This has manifested in everything from questions about Internet interconnection at the ITU World Conference on International Telecommunications to administration of domain names and numbers. At present, there is rising interest in multilateral approaches to Internet governance and a surge of authoritarian digital information practices, particularly (but not only) in China and Russia. An open question is what cyber sovereignty will mean in the IoT space. The public policy complexities inherent in cyber-physical systems may have to tip the scale toward multistakeholder solutions to adequately solve security, privacy, and consumer safety problems. But the mass deployment of cyber-physical systems will also enable new forms of authoritarian control.

Critically, cyber-physical systems also blow open what is now understood to be the cyber-regime complex in a third way. Already blurred policy distinctions between the cyber world and the physical world will increasingly become erased. It no longer makes sense to speak of the digital economy, but only the economy in which everything embeds digital components. In the same way, it will be difficult to speak about Internet governance issues as distinct from other areas of governance. The cyber-physical disruption will make them one and the same. Cyber norms processes and the practice of Internet governance will have to draw in experts in physical world processes and include actor classes from all industries, and from every level of social scale, from the local to the global. Alternatively, instead of a cyber-regime complex thought to “handle” governance in cyber space, cyber governance issues become integrated in real world policy dimensions.

Mass deployment of cyber-physical systems necessitates the incorporation of an enormous number of new, novice and often poorly-resourced players with divergent views about appropriate governance modalities into the global cyber regime complex. This new phase of global Internet penetration therefore places the global cyber regime complex under vastly increased strain at the same time as it makes that regime complex essential to the ongoing maintenance of crucial social, political and economic institutions and systems at every level of social scale from the local to the global. That is, the global cyber regime complex is becoming more fragile and prone to failure precisely as (and in some ways because) it is becoming increasingly necessary.

References

Abbott, Kenneth and Duncan Snidal. 2000. "Hard and soft law in international governance." *International Organization* 54(3): 421-456.

Acharya, Amitav. 2014. *Constructing a Security Community in Southeast Asia: ASEAN and the Problem of Regional Order*, 3rd ed. London: Routledge.

Allen-Ebrahimian, Bethany. 2018. "Chinese Police Are Demanding Personal Information From Uighurs in France." *Foreign Policy*, March 2, 2018. https://foreignpolicy.com/2018/03/02/chinese-police-are-secretly-demanding-personal-information-from-french-citizens-uighurs-xinjiang/.

ASEAN. April 27, 2018. *ASEAN Leader's Statement on Cybersecurity Cooperation.* https://asean.org/wp-content/uploads/2018/04/ASEAN-Leaders-Statement-on-Cybersecurity-Cooperation.pdf.

Betts, Alexander. 2010. "The refugee regime complex." *Refugee Survey Quarterly* 29(1): 12-37.

Boyd, Danah and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15 (5): 662–79.

Bradshaw, Samantha, Laura DeNardis, Fen Osler Hampson, Eric Jardine, and Mark Raymond. 2015. "The Emergence of Contention in Global Internet Governance." Global Commission on Internet Governance Paper Series No. 17. Available at https://www.cigionline.org/sites/default/files/no17.pdf.

Broeders, Dennis. 2017. "Aligning the International Protection of 'the Public Core of the Internet' with State Sovereignty and National Security." *Journal of Cyber Policy* 2 (3): 366–76.

Brunnée, Jutta, and Stephen J Toope. 2010. *Legitimacy and Legality in International Law: An Interactional Account*. Vol. 67. Cambridge: Cambridge University Press.

Buchanan, James M. 1965. "An economic theory of clubs." *Economica* 32 (125): 1-14.

Büthe, Tim, and Walter Mattli. 2011. *The New Global Rulers: The Privatization of Regulation in the World Economy*. Princeton, NJ: Princeton University Press.

Colgan, Jeff, Robert O. Keohane, and Thijs Van de Graaf. "Punctuated equilibrium in the energy regime complex." *The Review of International Organizations* 7(2): 117-143.

Cowhey, Peter and Jonathan Aronson. 2017. *Digital DNA: Disruption and the Challenges for Global Governance.* Oxford: Oxford University Press.

Cybersecurity Tech Accord. April 17, 2018. https://cybertechaccord.org/accord/.

Demchak, Chris and Peter Dombrowski. 2013. "Cyber Westphalia: Asserting state prerogatives in cyberspace." *Georgetown Journal of International Affairs*: 29-38.

DeNardis, Laura. 2009. *Protocol Politics: The Globalization of Internet Governance*. Cambridge, MA: MIT Press.

———. 2014. *The Global War for Internet Governance*. New Haven, CT: Yale University Press.

DeNardis, Laura and Mark Raymond, "The Internet of Things as a Global Policy Frontier," *UC Davis Law Review* 51.2 (2017): 475-497.

Department of Homeland Security. n.d. "Critical Infrastructure Sectors" dhs.gov Accessed December 16, 2018. https://www.dhs.gov/cisa/critical-infrastructure-sectors

European Union. 2016. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)." *Official Journal of the European Union* L119 (May): 1–88.

Finnemore, Martha and Duncan Hollis. 2016. "Constructing norms for global cybersecurity." *American Journal of International Law* 110(3): 425-479.

Fruhlinger, Josh. 2018. "The Mirai Botnet Explained: How Teen Scammers and CCTV Cameras Almost Brought down the Internet." *CSO*. March 9, 2018. https://www.csoonline.com/article/3258748/security/the-mirai-botnet-explained-how-teen-scammers-and-cctv-cameras-almost-brought-down-the-internet.html.

Global Commission on Internet Governance. *Norm Package Singapore*. November 2018. https://cyberstability.org/wp-content/uploads/2018/11/GCSC-Singapore-Norm-Package-3MB.pdf.

ICS-CERT. 2016. U.S. Department of Homeland Security, Industrial Control System Cyber Emergency Response Team. *Alert IR-Alert-H-16-056-01,* "Cyber-Attack Against Ukrainian Critical Infrastructure." Accessed at https://ics-cert.us-cert.gov/alerts/IR-ALERT-H-16-056-01.

Indian Computer Emergency Response Team bulletin. May 31, 2018. "Cryptocurrency-mining Malware Targeting IoT Devices." Accessed at https://www.cert-in.org.in

Karp, Paul. 2018. "Coalition's Surveillance Laws Give Police Power to Access Electronic Devices." *The Guardian*, August 13, 2018. https://www.theguardian.com/australia-news/2018/aug/14/coalitions-surveillance-laws-give-police-power-to-access-electronic-devices

Keohane, Robert O. and David Victor. 2011. "The regime complex for climate change." *Perspectives on politics* 9(1): 7-23.

Kirkpatrick, David D. 2018. “Israeli Software Helped Saudis Spy on Khashoggi, Lawsuit Says.” *The New York Times*, December 2, 2018. https://www.nytimes.com/2018/12/02/world/middleeast/saudi-khashoggi-spyware-israel.html.

Magrani, Eduardo, and Ronaldo Lemos. 2018. “Governance of Internet of Things and Ethics of Intelligent Algorithms.” The State of Responsible IoT. ThingsCon. https://medium.com/the-state-of-responsible-iot-2018/governance-of-internet-of-things-and-ethics-of-intelligent-algorithms-b88b565e126.

Manovich, Lev. 2011. “Trending: The Promises and the Challenges of Big Social Data.” *Debates in the Digital Humanities* 2: 460–75.

Microsoft. 2017. *A Digital Convention to Protect Cyberspace.* https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW67QH.

Marczak, Bill, John Scott-Railton, Adam Senft, Bahr Abdul Razzak, and Ron Deibert. 2018. “The Kingdom Came to Canada: How Saudi-Linked Digital Espionage Reached Canadian Soil.” T*he Citizen Lab*. https://citizenlab.ca/2018/10/the-kingdom-came-to-canada-how-saudi-linked-digital-espionage-reached-canadian-soil/.

Mastroianni, Brian. 2016. “Facebook Executive Arrested in Brazil over Data Access.” CBS News. March 1, 2016. https://www.cbsnews.com/news/facebook-executive-arrested-in-brazil-after-refusing-to-share-data-with-police/.

Maurer, Tim. 2011. “Cyber Norm Emergence at the United Nations – An Analysis of the UN‘s Activities Regarding Cyber-security?” Discussion Paper 2011-11, Cambridge, Mass.: Belfer Center for Science and International Affairs, Harvard Kennedy School.

McMillen, Dave. 2017. “Mirai IoT Botnet: Mining for Bitcoins?,” *Security Intelligence* 10. Accessed at https://securityintelligence.com/mirai-iot-botnet-mining-for-bitcoins/.

Nakamoto, Satoshi. 2008. “Bitcoin: A Peer-to-Peer Electronic Cash System.”

Neisse, Ricardo, Gianmarco Baldini, Gary Steri, and Vincent Mahieu. 2016. “Informed Consent in Internet of Things: The Case Study of Cooperative Intelligent Transport Systems.” In , 1–5. IEEE.

Nye Jr., Joseph S. 2014. “The Regime Complex for Managing Global Cyber Activities,” Global Commission on Internet Governance Paper Series, No. 1. Available at http://www.cigionline.org/publications/regime-complex-managing-global-cyber-activities.

O’Connor, Yvonne, Wendy Rowan, Laura Lynch, and Ciara Heavin. 2017. “Privacy by Design: Informed Consent and Internet of Things for Smart Health.” *Procedia Computer Science* 113: 653–58.

Paris, Roland. 2001. “Human security: paradigm shift or hot air?.” *International security* 26(2): 87-102.

Paris Call for Trust and Security in Cyberspace. November 12, 2018. https://www.diplomatie.gouv.fr/IMG/pdf/paris_call_text_-_en_cle06f918.pdf.

Prothero, Mitch. 2018. “Now Academics Studying ISIS Are Feeling The Heat Of An Internet Crackdown.” BuzzFeed News. December 8, 2018. https://www.buzzfeednews.com/article/mitchprothero/isis-researchers-have-become-the-collateral-damage-of-the.

Raustiala, Kal, and David G Victor. 2004. “The Regime Complex for Plant Genetic Resources.” *International Organization* 58 (2): 277–309.

Raymond, Mark. 2013/14. “Puncturing the Myth of the Internet as a Commons.” *Georgetown Journal of International Affairs* (International Engagement on Cyber III): 53-64.

———. 2016. "Managing Decentralized Cyber Governance: The Responsibility to Troubleshoot." *Strategic Studies Quarterly* 10(4): 123-149.

———. 2019. *Social Practices of Rule-Making in World Politics*. New York: Oxford University Press.

Raymond, Mark and Josie Smith. 2018. “Can Cybersecurity Tech Accord Really Curb State Actions?” *Center for Democracy & Technology* (blog). April 25, 2018. https://cdt.org/blog/can-cybersecurity-tech-accord-really-curb-state-actions/.

Raymond, Mark and Laura DeNardis. 2015. “Multistakeholderism: Anatomy of an Inchoate Global Institution.” *International Theory* 7 (3): 572–616.

Rosner, Gilad. 2016. *Privacy and the Internet of Things.* Sebastopol, CA: O’Reilly Media, Inc. https://www.oreilly.com/library/view/privacy-and-the/9781492042822/ch01.html.

Ruggie, John Gerard. 1993. "Territoriality and beyond: problematizing modernity in international relations." *International Organization* 47(1): 139-174.

Sandholtz, Wayne. 2008. “Dynamics of International Norm Change: Rules against Wartime Plunder.” *European Journal of International Relations* 14 (1): 101–31.

Schneier, Bruce. 2016. “Data Is a Toxic Asset.” Blog. *Schneier on Security*. March 4, 2016. https://www.schneier.com/blog/archives/2016/03/data_is_a_toxic.html.

Smith, Brad. 2017. “The Need for a Digital Geneva Convention.” *Microsoft On the Issues (blog)*. February 14, 2017. https://blogs.microsoft.com/on-the-issues/2017/02/14/need-digital-geneva-convention/.

Strange, Susan. 1996. *The retreat of the state: The diffusion of power in the world economy*. Cambridge: Cambridge University Press.

Tikk-Ringas, Eneken. 2012. "Developments in the Field of Information and Telecommunication in the Context of International Security: Work of the UN First Committee 1998-2012." *Cyber Policy Process Brief*. Geneva: ICT4Peace Publishing. Available at https://citizenlab.ca/cybernorms2012/ungge.pdf.

United Nations. 1966. *International Covenant on Civil and Political Rights.* Resolution 2200A, Article 17. Full text available at https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx.

———. June 24, 2013. "Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security." UN Doc. A/68/98. http://www.un.org/ga/search/view_doc.asp?symbol=A/68/98.

———. July 22, 2015. "Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security." UN Doc. A/70/174. http://www.un.org/ga/search/view_doc.asp?symbol=A/70/174.

———. 2018a. *First Committee Approves 27 Texts, Including 2 Proposing New Groups to Develop Rules for States on Responsible Cyberspace Conduct.* GA/DIS/3619. https://www.un.org/press/en/2018/gadis3619.doc.htm.

———. 2018b. "Advancing responsible State behaviour in cyberspace in the context of international security." Resolution A/C.1/73/L.37. https://undocs.org/A/C.1/73/L.37.

Weber, Rolf H. 2016. "Governance of the Internet of Things – From Infancy to First Attempts of Implementation," *Laws* 5(3): 28.

Zetter, Kim. 2014. "An Unprecedented Look at Stuxnet, the World's First Digital Weapom." *Wired*, November 3, 2014. https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/.

Zhongsun, Qiu. 2018. "Chinese Students Protest in America, Face Danger at Home." *Foreign Policy*, May 28, 2018. https://foreignpolicy.com/2018/05/28/chinese-students-in-america-arent-free-xi-jinping-communist-party-repression/.

How Internet Infrastructure Emerges in the Global South: Sociotechnical Aspects of an Internet Exchange Point

Fernanda R. Rosa[1]

Abstract

This research paper examines the formation of an Internet Exchange Point (IXP) in a country with a low level of telecommunications competition in the Global South. IXPs are internet nodes that work as meeting points for networks to exchange routes and traffic shaping the global internet since its formation. In this paper I apply the Actor-Network Theory (ANT) framework to the policy dynamics around the first IXP formed in Mexico, contributing to unveil the materiality of internet infrastructure and the imaginary that surrounds it. Following Michel Callon (1984)'s three principles: *agnosticism* – to be impartial with the parts of a controversy; *symmetry* – to analyze different perspectives with the same lens; and *free association* – to break the divide between society and technological artifacts, the analysis, based on ethnographic field work and qualitative in-depth interviews, shows how technical and political aspects are completely interlaced in the design and implementation of an IXP, defined here as a network of relationships. Organizations—characterized by their design and governance—along with individuals, documents, laws, and technology artifacts, are significant actors in the scenario where social, political and economic goals are delegated to the IXP's technical functions. The lack of convergence among these actors, however, prevents the project from succeeding in the first years of deployment as its implementers expected, while regulatory documents work as the supporters of a dynamic equilibrium to keep the project ongoing.

Key words: Internet Exchange Point (IXP), Internet Governance, Internet Interconnection, Infrastructure, Actor-Network Theory, Science and Technology Studies (STS).

[1] PhD Candidate, American University - This work was supported by the American University Doctoral Student Research Award, Columbia University's School of International and Public Affairs and Carnegie Corporation of New York Research Award, and Google Policy Fellowship. Special thanks to the Red en Defensa de los Derechos Digitales (R3D)'s team that was a vital support during my period as a Google Fellow in Mexico City; and also Philipp Richter for useful comments on a previous version of this paper, a recipient of the TPRC Student Paper Award at the 46th Research Conference on Communication, Information and Internet Policy (TPRC46).

Introduction

In May of 2014, a cadre of companies, policymakers, and journalists convened at an event for launching a new part of the internet architecture in Mexico—an Internet Exchange Point (IXP). An IXP can be initially understood as an internet node, a physical facility where different networks can interconnect and make private agreements for the purpose of optimizing their respective resources to exchange traffic and routes on the internet. Networks, in this context, are mainly Internet Service Providers (ISPs) (e.g. Comcast, AT&T, Telmex) and content providers (e.g. Google, Facebook, Amazon), but can be also banks, universities and other organizations which have an autonomous system number (ASN) to uniquely identify their networks on the internet, known as the "network of networks." In terms of internet infrastructure, autonomous systems *run* networks, but here both terms will be used interchangeably.

Commercial agreements between networks shape the way that they will interconnect with each other, if by "peering" or by "transit." Peering is a collaborative relationship that is beneficial for the networks involved due to the capacity to share resources. In this relationship, autonomous systems, such as ISPs allow the networks with whom they are peering to have access to both their own routes and clients' routes to have the same benefits in return. Commonly this relationship does involve monetary payment, because it is based on the assumption of parity between peers; however, depending on the amount of imbalance in the traffic, ISPs can establish paid-peering (Faratin et al., 2008; Metz, 2001). Transit, in turn, is a customer-provider relationship established between two autonomous systems, whereby access to the larger internet is provided through a financial settlement (Faratin et al., 2008; Metz, 2001). In this relationship, one party wants to buy connectivity while the other one has the infrastructural resources to sell it.

When peering involves two autonomous systems only, it is known as "bilateral peering," or a Bilateral Peering Arrangement (BLPA). When autonomous systems connect to multiple peers at once, this is known as "multilateral peering" or a Multilateral Peering Arrangement (MLPA) (Giotsas et al., 2013; Metz, 2001; Richter et al., 2014). IXPs facilitate any of these arrangements. They are considered *public* peering facilities, as opposed to private facilities where two networks interconnect directly without an IXP—a constant occurrence among big players. Importantly, being considered *public* should not be confused with IXP models of administration, which can be public or private, conducted by government, for profit or not-for-profit organizations.

The benefits of IXPs are generally recognized, particularly in that they allow the joining of many networks at the same place, facilitating private agreements and interconnection arrangements in order to reduce international traffic and traffic costs, and improve, as a consequence, the internet quality for final users by keeping local content locally, diminishing latency and leveraging speed (Fanou, Valera, Francois, & Dhamdhere, 2017). And although there are divergences on the number of IXPs in the world depending on the source considered, there are likely more than five hundred IXPs, which are unequally distributed and more numerous in affluent areas of the globe (Klöti, Ager, Kotronis, Nomikos, & Dimitropoulos, 2016). Intimately related to this unequal distribution is the traffic pattern on the internet. It is estimated that in Latin America (LATAM) a great amount of data goes from LATAM countries to the United States (CAF, 2014), even when emitter and recipient are from that same region. This is known as boomerang traffic, detour or trombone.

Although IXPs are considered a critical part of the internet architecture, many countries do not have an IXP (e.g. Uruguay and Venezuela in Latin America; Algeria and Libya in Africa; Afghanistan and Iraq in Asia; Tonga and the Solomon Islands in Oceania, to name a few) (PCH, 2018). In Mexico, at the time of writing, the IXP under study is responsible for a low amount of internet traffic, with a speed of 10 to 20 Gigabits per second (Gbps) according to interviewees. For a rough comparison, as information about the internet traffic per country is privatized and not publicly available, the main IXP in Latin America, located in Sao Paulo, Brazil, has an average of more than 2 Terabits per second (Tbps). As populations and internet penetration in both countries are different, to contextualize these numbers it helps to know that Mexico is responsible for 1,5% of the global web traffic while Brazil, located in the same region, responds to 3,5% (Akamai, 2018). Another metric to contextualize the volume of data passing through the first Mexican IXP is the number of autonomous systems in the country connected to that facility: of the 366 autonomous system numbers assigned within Mexico, fewer than 10 are connected to the IXP. Together, these sources can be considered an indication that the internet traffic in the country continues to go through bilateral agreements in private facilities that precede the relatively new available IXP in Mexico City.

Nonetheless, with regard to the launching of an IXP in Mexico, Carlos Casasús, who is the president of the committee formed to coordinate the new facility, mentioned to a journalist some benefits that would justify the implementation of the first IXP in the country (Rivera, 2014). His considerations encompass four key issues:

a) Leveraging the quality of the internet, through the “decrease of latency between connections” and the “improvement of the internet traffic”;

b) Strengthening sovereignty, through avoiding unnecessary international routes, “enriching the country’s technological infrastructure,” enabling the country to join others “that are at the forefront of technology”;

c) Leveraging market competition, helping to establish “a healthier competition among telecommunications operators,” and “attract more foreign investment”; and

d) Generating social benefits, “narrowing the digital divide by making the internet more accessible to more people,” and “encouraging further development of national content online.”

While these reasons reflect local motivations, they incorporate components of a prevailing dialogue among international organizations. Many different agencies, including the Organisation for Economic Co-operation and Development (OECD), the Inter-American Development Bank (IDB) and the World Bank, have produced reports on broadband development, emphasizing the role of IXPs in improving connectivity rates in “developing” countries (Agudelo et al., 2014; Blackman & Srivastava, 2011; Intven & Tétrault, 2013; OECD & IDB, 2016; Weller & Woodcock, 2013).

By unveiling a very opaque technology underlying the internet architecture, what this paper elucidates is that the expectations about the first IXP in Mexico are based on assumptions that depend on different sociotechnical processes and actors intertwined, and are not—as can be understood by a technology deterministic approach—sustained by the IXP "affordances" themselves. Affordances are 'the possible actions a person can perform upon an object' (Norman, 2010, p. 228), or yet, the 'promise and permission' of artifacts, which, in action, merge their characteristics with who handles them, supporting new actions that emerge in a process named "translation" (Latour, 2002)—only conceivable if object and subject are considered altogether. In this paper, I analyze the incomplete realization of such a translation process in the case of the Mexican IXP, or the reasons for the expectations of some groups involved in broadband discussions and in the deployment of the IXP to be frustrated.

This research is guided by Actor-Network Theory (ANT), "a method for mapping how every object or actor is shaped in its relations" (Law, 2016, p. 10, emphasis in the original). In this framework, there is a call to look at micro structures where social relations are built between humans and non-humans, society and technological artifacts. Despite the scale difference, ANT theorists understand that the macro structure of society is not distinct from its micro structure. In fact, from an ethnographic approach, ANT seeks to avoid not only technology determinism, but also social determinism, dismissing the existence of a social structure ruling life, without disregarding the existence of enduring patterns that may be identified (Law, 2015). This is why, in this framework, power relations are expected to be unveiled only after a certain web of relations is understood (Callon, 1984; Latour, 1991).

Authors from this stream of thought assume not only a symmetric relation between society and artifacts, but also understand that there is a continuous interchange between humans' goals and artifacts' functions. This happens in such a way that a speed bump, for instance, can be understood as a "delegation" of engineers' goals in pavement and concrete, and a "translation" of an action—the speed law enforcement—into a technique (Latour, 1999). In other words, the desired action of making drivers slow down, in this case, is not only *expressed* by a "negotiable" speed limit sign, in which the driver has the opportunity to ignore it. Instead, the action is *provoked* by "unnegotiable speed bumps" (Latour, 1999). Thus, mediating human goals, the technique influences human behavior with its own functions and characteristics.

I apply this approach to the study of IXPs using three principles presented by Michel Callon (1984): *agnosticism* – to be impartial with the parts of a controversy; *symmetry* – to analyze different perspectives with the same lens; and *free association* – to break the divide between society and technological artifacts. In the following sections I examine the dynamics of IXP formation and the actors that emerge from it. I then discuss the translation process throughout four moments that are shown to be embedded by social, political and economic factors. Finally, I conclude defending an IXP sociotechnical definition as a way to illuminate the complex dynamics that characterize Internet Exchange Points.

Materials and Methods

The material analyzed for the present work comes from ethnographic research conducted in the states of Mexico, Oaxaca and Chiapas in Mexico between June and September of 2017. Participatory observation of events, including IXP activities, the Forum on Indigenous and Communitarian Media, international organization report presentation and more than twenty in-depth interviews with indigenous communities', policymakers', not-for-profit organizations', academics', and Internet Service Providers' representatives compose the primary sources analyzed.

The Formation of an IXP

In 2012, the Organisation for Economic Co-operation and Development (OECD) released an influential report on Mexico, one of its few member-countries from the Global South, stating that "The welfare loss attributed to the dysfunctional Mexican telecommunication sector is estimated at USD 129.2 billion (2005-2009) or 1.8% GDP per annum" (OECD, 2012, p. 9). Among its recommendations, the report stated that the telecommunications regulator, the Federal Telecommunications Institute (IFT), should have the power to impose regulations and sanctions to leverage competition. With regard to infrastructure specifically, the report says that "The inability to mandate, or at least set out, reasonable conditions for infrastructure sharing is arguably one of the main bottlenecks that prevent competition" (OECD, 2012, p. 12). Since then, the report has been a respected voice in policymakers' circles discussing infrastructure-sharing projects and the intensification of asymmetric regulation applied to the preponderant economic agent, Telmex. Carlos Casasús' story of conversations about creating an Internet Exchange Point in the country at the regulatory agency is an example:

> We were already talking about having an IXP. I was the chairman of COFETEL's Advisory Board [currently IFT]. I had a meeting with the COFETEL's president [Mony Sacha de Swaan] and I said 'Why do not we do that? It is an OECD recommendation.' He said: 'Do you think we can do that? We have been working for many years…' So, we managed to get [some] partners to start.[2]

Casasús is known for his efforts within the not-for-profit organization Corporación Universitaria para el Desarrollo de Internet (CUDI), whose goal is to congregate and escalate resources among higher education institutions in Mexico.[3] It is in this context that he and colleagues thought about building an IXP first in the beginning of the 2000s to improve universities' internet connectivity, keep the country's content local and decrease dependence on the United States' infrastructure. Hans Ludwing Reyes Chávez, one of the engineers who work for CUDI and who is currently responsible for the IXP in México, remembers that: "[The idea] did not prosper because there were not enough fiber networks to do it." In fact, network interconnection depends on numerous infrastructure resources including optical fiber and broadband links.

According to Carlos Casasús, an inspiration for CUDI and the IXP project has been the Brazilian National Research and Educational Network (RNP), a network of universities in Brazil whose goal is to integrate academic institutions with the support of a backbone fiber network running since

[2] This and other verbatim quotes come from interviews with the author.

[3] Previously to this role, he was the Financial Director of Telmex, when it was a state company, and worked in the front of the Federal Law of Telecommunications discussions, approved in 1995 (www.diputados.gob.mx/LeyesBiblio/abro/lftel/LFTel_abro.doc). He was then the first COFETEL president in 1996, the regulatory agency that since 2014 is called IFT.

1992. At the time of writing, RNP has access points in all 27 Brazilian states, facilitating the interconnection of networks in different regions, and serving as points of interconnection of some IXPs within the country. Unlike RNP, though, CUDI does not have a fiber network in Mexico. The organization depends on an agreement between the Ministry of Communications and Transportation (SCT) and the Federal Electricity Commission (CFE), which interconnects approximately 40 universities, but constantly presents technical problems, according to the interviewees.

This is an important context to understand, that the first IXP initiative in Mexico was led by an educational organization with clear purposes, but devoid of internet infrastructure resources. In 2014, CUDI, and more specifically its president, put together five companies to start the exchange point in Mexico City: Kio Networks, Megacable, Nextel, redIT, and Transtelco. These organizations constituted the IXP's founding partners, which envisioned some benefits for themselves, including sharing infrastructure and exchanging traffic among the parties, and in the case of Kio Networks—a prominent data center within the country—the opportunity to become the host of new networks. Interestingly, the group of the IXP founders does not comprise Telmex, the telco incumbent, and other academic institutions than CUDI, which would be required to have autonomous system numbers to interconnect, and is reported to have difficulties in receiving ASNs from NIC Mexico.

While the participation of a player like Telmex cannot guarantee the success of an IXP, Telmex competitors and the IXP founders defend that it is a crucial contributor to it, given that Telmex not only has the biggest number of clients, concentrating 57.7% of the internet market,[4] but it also has the largest infrastructure to reach different parts of the country, with more than 190,000 km of optic fiber (Telmex, n.d.). For instance, an Internet Service Provider (ISP), which needs to deliver data packets in places where its own optic fiber mesh does not reach, has two possibilities: buy transit or do peering with another company to deliver it. However, an incumbent agent has very few incentives to share its own infrastructure and peer with potential competitors. Economically, it can conclude that it is more advantageous to sell transit to some ISPs than to peer with them. In Mexico, due to its reach, Telmex would be one of the most likely companies from which this supposed ISP would buy transit. Thus, for Telmex, it is reasonable to think that an Internet Exchange Point would likely to reduce its clients and would not benefit its business.

Taking part on this controversy, and guided by the purpose of leveraging competition in the country, two months after the beginning of the IXP operation, the law that marks the reform in the telecommunications sector in 2014 determined that the preponderant agent, Telmex, should: "Have a physical presence in the Internet exchange points in the national territory, as well as to enter into agreements that allow Internet service providers the internal exchange of traffic in a more efficient and less expensive way according to the terms that the Institute define" (Mexico, 2014, Art. 138, VIII, own translation). As of the time of writing, though, Telmex was not yet an IXP member, but the expectations were that it would happen soon, if the company does not appeal to the guidelines issued to enforce the law in 2017.

[4] The other big players are Grupo Televisa, with 21,5% of the market, and Megacable-MCM, with 13.5% (IFT, 2017, p. 27).

According to these guidelines, the preponderant agent or the agent with substantial market power "must establish Connectivity through the deployment of fiber optic links to IXPs that request it, and where there is at least one Internet Service Provider with which [it] does not have a traffic exchange agreement [peering agreement] (…)." (Mexico, 2017, Cap. III, own translation). Moreover, it "(…) must advertise the Routes of [its] clients and accept the Routes of the ISP members of the IXP. The Routes must be kept constantly updated in the Routing Table" (Mexico, 2017, Cap. III, own translation). With such rules, all the costs for the incumbent to be connected to any IXP are supposed to be covered by the company. Furthermore, by establishing peering agreements with any ISP connected to an IXP within the country, Telmex will lose the possibility of selling transit to its competitors when they want to reach Telmex's own networks and Telmex clients' networks.

This type of regulation to require interconnection, while it has reflected significant lobbying from CUDI, has not been received unanimously among players in the market and specialists. In the illustrative opinion of a content provider representative, who is responsible for interconnection issues at a company that already maintains private peering agreements with Telmex, he points out that an IXP is useless in a market where there is a low level of competition, and not an ISP ecosystem to benefit from interconnecting publicly at an exchange point. He defines the Mexican IXP as a "party where all the guests already have relationships with each other," so pay for a "ticket" to participate in such party is a waste of money. In other words, for this interconnection specialist, in a market like Mexico, the equipment necessary to build an IXP and the structure necessary to maintain it becomes costly, and will not solve the competition problem by itself: "A switch helps the small players, but if they do not exist (…) an IXP will not generate small players." In this vein, he sees the regulation to require a player to participate at an IXP as an unwelcome interference: "When there are no commercial reasons [to interconnect], one makes the law," he says.

On the other side, IXP defenders and pro-regulation actors shift the focus to the challenges faced by both small players that do not have the market power for interconnecting directly via private facilities, or other Telmex medium-size competitors, which depend exclusively on transit services because of Telmex business strategies and policies. It is not uncommon that a Telmex competitor who wants to reach Telmex networks needs to send its traffic to an IXP in the United States, where global internet networks—also known as Tier 1 networks—which keep agreements with Telmex will redirect such traffic to return to Mexico. IXP defenders and pro-regulation actors will say that this boomerang route raises cost issues for the companies, internet quality issues for the users and sovereignty issues for the country.

Julio César Bravo, an incumbent competitor representative whose company is one of the IXP founders, believes that there are viable business opportunities to raise in an IXP in México, but Telmex needs to be part of it to make it attractive to Content Delivery Networks (CDNs). CDNs are services provided by third-party companies or by big content providers (e.g. Google, Netflix) that cache highly accessed web content to make it easier and quicker for users to reach. They have an interest in becoming a member of an IXP if a great amount of traffic is expected to circulate through its facilities. Although Julio César Bravo would agree that the IXP is currently a party with guests that are already linked among them, his company accepted to be an IXP founder based on future business perspectives, such as providing connection for the IXP to the United States. "In the end it is business (…) There is no altruistic issue. Everything is totally and completely business," he admits.

Thus, for the IXP team and participants, the state regulation to require the incumbent to be part of the exchange point is positive, and generates expectations that other important players will interconnect to the exchange point in the near future. On the other side, consequential players, including the incumbent and the ones that already have interconnection agreements with it, such as big content providers, do not see benefits from connecting to an IXP in the present conditions. In fact, although the IXP has already been working for some years, its outcomes have not been measured or made public, which generates critics: "I have no elements to know if I can trust the IXP operator or not. In theory, yes, because I'm in a university and I have to rely on CUDI, right? But I do not even know where IXPs' performance measures are, if I do not have numbers I cannot have confidence," says Luis Miguel Martínez Cervantes, a professor and also the Internet Society Chair in Mexico, an organization that has supported the creation of IXPs around the world.

Luis Martínez argues further that building an IXP in Mexico at that moment was "a political and not a technical decision," meaning that the IXP was a government response to the OECD report agenda, while his academic colleague, Judith Mariscal, a professor and specialist in telecom and digital divide issues (Flores-Roux, Mariscal, & Aldama, 2009; Galperin & Mariscal, 2016), argues that the IXP was Carlos Casasús' and CUDI's agenda, indicating lack of involvement in the discussion. Clearly, CUDI's IXP lobby was directed to government and some companies and did not incorporate other academics and civil society organizations in its process.

To finalize this examination session, it is important to explain the governance and design of the IXP, here understood as two sides of the same phenomenon (DeNardis, 2014; Musiani, 2013). The IXP governance is under the auspices of the not-for-profit organization Consortium of Internet Exchange Traffic (CITI, A.C.), which is led by the CUDI president, Carlos Casasús, and complies with the partner organizations of the IXP that meet every three months. As of the time of writing, organizations connected to the IXP are Akamai, Cloudfare, CUDI, Enlace TPE (TotalPlay Empresarial), Google, KIO Networks, NIC Mexico, Megacable, y Transtelco. Interestingly, some companies that were connected to the IXP in its beginning are not anymore. This happens because, there is a merging trend among businesses (e.g. AT&T bought Nextel Mexico and KIO Networks bought redIT), an expression of technology convergence that may reduce the number of IXP participants in a small market.

KIO Networks is the company that owns the data center which hosts the IXP's equipment, being responsible for the co-location and the building infrastructure—electricity, cooling and security. It has an important role in IXP governance, once its policies are crucial in the design of the IXP and its geographic location.

To be part of the IXP consortium, the organizations need to pay $810 or $2,430 monthly to have a port of 1Gbps or 10Gbps, respectively, but companies such as Content Delivery Networks may negotiate these terms due to the perceived importance of caching highly accessed content locally for the economy of IXP participants. To be connected to the IXP, a network—owned by a company, a community or the government—needs to be an autonomous system, which means having an autonomous system number assigned by NIC Mexico, and to be physically connected to the IXP in Santa Fé, Mexico City, where the KIO Networks data center is located. If an interested network is already based in this data center, it will purchase a "cross-connection" service from KIO to have its cables connected to the IXP. If this is not the case, a point-to-point link from the company headquarters to the IXP is necessary. In this scenario, one of the challenges is that the usage cost for local fiber lines is expensive and wireless lines are not abundant in the country, contributing to preventing IXP attractiveness. As Luis Martínez exemplifies:

> What happens is that for [my network] to arrive from a town 10 km from the IXP, I have to use the Telmex network. And in this case, I find it cheaper to use the Telmex internet service than what the IXP is going to give me. Because what Telmex will charge [for a fiber line] to take me to IXP is going to be more than what Telmex will charge to provide me the internet service without having to go to the IXP.

Part of this scenario is due to the access that the incumbent has to passive infrastructure throughout the country, including antennas, posts, and right-of-way—the legal possibility of passing cables through public spaces. Telmex used to be a public company and has kept better negotiations with supporting infrastructure historically. In contrast, small players have more difficulty to have access to right-of-way according to interviews, making competition even more problematic.

The Incomplete Translation Process

The analysis that follows is based on the actors that stood out in the dynamics of the IXP formation: the OECD report, CUDI's president, the telecommunication regulator, the telco incumbent, the telecommunications law and IFT guidelines, NIC Mexico, the fiber networks, the passive infrastructure (posts, optical fiber, right-of-way), big content providers/CDNs, incumbent competitors, global networks (Tier 1 network), civil society (including academics) and the core actor, the IXP, that from the narrative goes beyond its equipment—cabinet, switch, router, cables—, and includes the data center, the networks connected to it, and the governance consortium team. Independently of being human or non-human, actors are considered symmetrically, including individuals, networks, supporting infrastructure and documents, who have had an active role in the dynamics. In ANT, action is conceived not as an exclusive "property of humans," but as a result of a combination of agents or "actants," including technical artifacts (Latour, 1999). Regarding documents, the very argument to consider them more than sources of information is that text transcends authors and their intentions. They can instigate actions and can "be considered as actors in their own right" (Prior, 2008, p. 822).

The IXP formation is a result of numerous social, political and economic goals that are delegated to this artifact, in a translation process in which actors' identities and characteristics are negotiated in relation to the others. Michel Callon (1984) suggests four moments of observation to understand this translation development: problematization, interessement, enrolment and mobilization. These moments are not independent of each other, though, they are dynamics that can overlap.

Problematization

The problematization moment is when certain actors "establish themselves an obligatory passage point in the network of relationships they [are] building (…) [, an actor] indispensable in the network." (Callon, 1984, p. 204). In the present case, this actor is the IXP, voiced by CUDI's president who can be considered IXP's "spokesman," in Michel Callon's terms.

The goal of building the first IXP of the country required CUDI's president to negotiate with several actors. In this context, the OECD report worked as a catalyst for the interconnection facility formation once it recommended reducing market concentration. The IXP promises, echoed by CUDI's president, conveyed this possibility, which was in accordance with the telecommunications regulator interest. The document worked as both a symbolic and material supporter for CUDI's president to resort to it in his dialogues to enable a group of supporters.

Notably, even when the IXP was just a project, it was already an actor in terms of the outcomes expected. The question was if there would be enough support to physically build it. CUDI's president starred the problematization phase, defending that it was the best moment for joining efforts to build an IXP, and that such a technological artifact was the best answer to address not only economic disparities in the market, but also social and political issues.

Interessement

The interessement moment arises when the IXP project needs to attract enablers and distance them from other alternative responses to the existent problems. CUDI's president defended that, once formed, an IXP would improve internet traffic and quality; avoid international routes and strength sovereignty; leverage market competition; narrow digital divide and encourage development of national content online. Interestingly, CUDI has for a long time been interested in improving Mexican universities' connectivity, so sharing infrastructure in the IXP was seen as an alternative to their difficulty in negotiating effective fiber networks connections given that the organization is devoid of an academic backbone network.

The regulator, IFT, heard CUDI's president voice parallel to the OECD report repercussions and moved to delegate its policy goals to law and guidelines requiring the telco incumbent to participate in the incoming IXP. The law issued in 2014 worked as a guarantee for companies to invest and engage in the project, even if in a small number. The promise of making peering agreements and sharing infrastructure with the incumbent in the near future supported such private investments. The players interested were in unison, understanding that without the law, the telco incumbent would not integrate the project.

Advertised outcomes of the IXP showed technical, political and economic purposes completely intertwined. Beyond the government collaboration, they attracted companies interested in optimizing their costs and leveraging their profits based on the belief that in a certain period of time the IXP would deliver what had been promised, especially traffic exchange with Telmex. The specificity of the networks attracted to the IXP project is that they were at a disadvantage in the market in comparison with the incumbent infrastructure and the dependence on Tier 1 networks in the United States to connect to the Telmex network. The regulator's law and guidelines requiring

Telmex to be part of the IXP give the reasons necessary for them to join the project, and more importantly, keep the project ongoing even after some years of no expected results. The law and the guidelines, which the telecommunications regulator issued to enforce the law, are key actors for keeping the IXP live in a fragile equilibrium.

Remarkably, companies read the expected outcomes with an economic lens. For instance, reducing the international traffic means saving money in traffic costs and decreasing latency, while CUDI and government would defend that it means strengthening Mexico sovereignty. Thus, the association of broad social and political benefits to the IXP does not have the same significance or attractiveness for different actors involved.

Enrolment

Callon points out that "To describe enrolment is (…) to describe the group of multilateral negotiations, trials of strength and tricks that accompany the interessements and enable them to succeed" (Callon, 1984, p. 211). The IXP formation depended on actors not always visible and ready to support the project: a data center designed to securely host its equipment—servers, switches, routers, cables, fiber internet links—and autonomous system numbers. For the networks to interconnect using an IXP they need to "negotiate" (Callon, 1984) with these actors; otherwise, they become barriers for networks to effectively be part of the IXP. For instance, the difficulty faced by some universities to be assigned an autonomous system number by the NIC Mexico has kept them apart from the IXP. Legal and economic constraints that restrict the offer of affordable fiber links to Santa Fé can reduce interest of regional networks based far from Mexico City in connecting to that internet node, as well as induce them to continue buying internet from the incumbent as exemplified earlier by an interviewee. Additionally, to keep the IXP equipment functional and sustain its colocation at the data center, IXP members are asked to pay a monthly contribution in dollars, which also becomes a barrier for small internet service providers. In the end, the design and governance of the data center are altogether crucial for IXP performance, not only for what they allow, but also for what they constrain. While the IXP itself is considered to be physically formed by a cabinet with switches, routers, servers and cables, it is in fact intertwined with the attributes of the data center where it is colocated, the networks that are successfully connected to it, and the ones that are not connected due to failed negotiations with other infrastructure actors marked by legal and economic barriers.

There are certain actors that were not involved in the formation of the Mexican IXP, although the social outcomes that the IXP spokesman advertised to attract supporters are of great interest to them. These include civil society groups who advocate for affordable internet and are responsible for building community networks in places where internet service providers are not willing to serve, as well as academics who are important voices in the area of telecommunications and the digital divide. Considering that, despite CUDI's president and the telecommunication regulator, the other active actors engaged in the formation of the IXP who voiced their interests do not mention concerns with the digital divide or with sovereignty. Such promised outcomes seem to be primarily a rhetoric tool for the IXP spokesman, and not a mobilizer used to aggregate actors previously interested in these issues around the IXP. In this scenario, modeling the IXP in this direction is thus unlikely to happen, once such outcomes are restricted to the desire of some actors.

Interestingly, in this case, IXPs become similar to other infrastructures in which beyond their technical functions, their form, “or the poetics of infrastructure” (Larkin, 2013, p. 329) shows political facets through the “imaginary” and the “fantasy” created around them. Furthermore, as Cynthia Cockburn argues, if based on the way and by whom they were built, technologies are masculine and cannot be seen in a sexless mode (Cockburn, 1983), as a technology, the Mexican IXP is also a commercial entity, used to facilitate commercial agreements, and based further on its governance and design, cannot be seen differently even if led by a not-for-profit organization.

Mobilization

Convergence and a certain level of consensus around a proposition mark the success of the mobilization moment. In the case of the IXP formation project and the actors that emerged in the dynamics, the mobilization results can be considered only partial. The project was formulated based not only on the affordances of an IXP—or what it can do—but on the successful translation of organizations’ goals into technology functions. Yet the fact is that after some years since the IXP formation, that didn’t happen. IXP development has maintained the interest of new networks in connecting to the first Mexican IXP low, keeping the number of its members less than ten. Lack of abundant and affordable links to connect networks in other regions to the data center shows the role of fiber networks and passive infrastructure as actors that constrain such interest. Companies that founded the IXP were acquired by other businesses and the IXP stage was not enough to initially attract new big players to the project, such as AT&T who bought Nextel, a previous IXP member.

Public information about IXP performance is not available, but the reported IXP traffic in interviews is modest. Thus, there is no evidence that key promised outcomes, such as reducing international traffic and latency, leveraging competition and access to the internet for more people, have been addressed. Some academics are skeptical and still not engaged in the project.

On the other hand, it is not a trivial outcome that, despite all the frustrated expectations, the IXP in Mexico is still running while there are numerous defunct IXPs in the world.[5] The mobilization moment that started with the formation of the group that would support the IXP formation, including the telecommunication regulator and some companies, has been continuously sustained. For this to happen, the most important actors in this scenario seem to still be the law and the guidelines that require the telco incumbent to adhere to the IXP. They generate the expectancy that after the Telmex connection, IXP traffic will exponentially increase, networks will not need to use Tier 1 networks in the United States to connect to the incumbent, and new networks will be attracted to the IXP, contributing to the likelihood that the IXP will prosper. Such results, however, are not a given. They are part of the infrastructure imaginary around the IXP and will depend on negotiations among actors when Telmex changes its position in the scenario. The regulatory documents thus support a dynamic equilibrium based on this imaginary that allows the project to continue.

[5] A filter at the Packet Clearing House database (www.pch.net/ixp/dir) shows 112 defunct IXPs in the world.

Conclusions

In internet network scholarship, authors have defined Internet Exchange Points as "a network infrastructure with the purpose to facilitate the exchange of Internet traffic between Autonomous Systems and operate below layer three" (Chatzis et al., 2013, p. 20), or "a shared layer-2 switch fabric environment, with three or more participants, where new participation is not rigorously constrained, and over which the members peer with each other, exchanging customer routes" (Fanou et al., 2017, p.4).[6] Such definitions have focused on highlighting the IXP affordance of conducting node-to-node communication and are guided by the industry definition that is compiled in an European association of IXPs report where: "An Internet Exchange Point (IXP) is a network facility that enables the interconnection and exchange of Internet traffic between more than two independent Autonomous Systems" (Euro-IX, 2015, p. 3).

From a sociotechnical vantage point, the evidences obtained with the present research enable understanding IXPs as relationships of players with goals and functions that mesh to become an interconnection facility in the internet. Such networks of relationships are dynamic and are defined relative to each player in the scenario, which includes organizations—characterized by their design and governance—, individuals, documents, laws, and technology artifacts, such as IXP equipment and the passive infrastructure. Negotiations are continuous, and as players' strategies and characteristics change, the relationships also change, strengthening or weakening IXP equilibrium.

IXPs may have different deployments and pathways depending on where they are built. In fact, it is unlikely that one can just transfer an IXP from one country to another, given that actors will likely to be different in each territory, and will require, in consequence, adaptations of other players, including in terms of design and governance when appropriate. Because of that, definitions in which there is a locked understanding of an IXP such as in Fanou et al. (2017), who state that an IXP is "where new participation is not rigorously constrained," are clearly normative and not a generalizable conceptualization as the Mexican IXP demonstrates.

"The Internet is only virtually stable" (Star & Bowker, 2010, p. 237). The study of IXPs reiterates that. It is not that IXPs are formed and then expected to be perennial. Incomplete translation processes can generate discontinuation provoked by a chain of actors. As affordances are learned and not static or given, a continuous interpretation of an IXP, based on local meanings, needs to be in action to understand the lively ties established among players involved with the IXP deployment, design and governance. With that, an IXP may be de-blackboxed and its materiality is unveiled.

[6] "Layers" are abstractions used by the internet community to conceptually describe a network ecosystem. For that, there are two basic references: the Open Systems Interconnection model, also known as OSI model, which comprises seven layers, and the TCP/IP Protocol Architecture Model (acronym to Transmission Control Protocol/Internet Protocol), which comprises five layers (Oracle, n.d.-a, n.d.-b). The layer-2 switch that is below layer-3, as mentioned by the authors above, corresponds to the Data Link layer in both models. It provides point-to-point data transfer (Shaw, 2017).

References

Agudelo, M., Katz, R., Flores-Roux, E., Botero, D., Cristina, M., Callorda, F., & Berry, T. (2014). *Expansión de infraestructura regional para la interconexión de tráfico de internet en América Latina*. CAF. Retrieved from http://scioteca.caf.com/handle/123456789/522

Akamai. (2018, April). Real-Time Internet Monitor | Akamai. Retrieved November 30, 2018, from https://www.akamai.com/us/en/solutions/intelligent-platform/visualizing-akamai/real-time-web-monitor.jsp

Blackman, C., & Srivastava, L. (Eds.). (2011). *Telecommunications Regulation Handbook* (Tenth Aniversary Edition). Washington, DC.: The International Bank for Reconstruction and Development / The World Bank, InfoDev, and The International Telecom munication Union.

Callon, M. (1984). Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay. *The Sociological Review*, *32*(1_suppl), 196–233. https://doi.org/10.1111/j.1467-954X.1984.tb00113.x

Chatzis, N., Smaragdakis, G., Feldmann, A., & Willinger, W. (2013). There is More to IXPs Than Meets the Eye. *SIGCOMM Comput. Commun. Rev.*, *43*(5), 19–28. https://doi.org/10.1145/2541468.2541473

Cockburn, C. (1983). Caught in the Wheels. *Marxism Today*, 16–21.

DeNardis, L. (2014). *The global war for Internet governance*. New Haven: Yale University Press.

Euro-IX. (2015). *euro-ix Internet Exchange Points: 2015 Report* (p. 21). European Internet Exchange Association (Euro-IX). Retrieved from https://www.euro-ix.net/media/filer_public/11/2c/112c450a-7f3c-4947-9c0b-6104b3a38659/euro-ix-ixp-report-2015-final.pdf

Fanou, R., Valera, F., Francois, P., & Dhamdhere, A. (2017). Reshaping the African Internet: From scattered islands to a connected continent. *Computer Communications*, *113*, 25–42. https://doi.org/10.1016/j.comcom.2017.09.006

Faratin, P., Clark, D. D., Bauer, S., Lehr, W., Gilmore, P. W., & Berger, A. (2008). The Growing Complexity of Internet Interconnection. *Communications & Strategies*, *72*(4th Quarter), 51–71.

Flores-Roux, E. M., Mariscal, J., & Aldama, F. A. (2009). Propuesta de licitación de la fibra oscura propiedad de la CFE: solución que genera escasez artificial, tanto presente como futura. *DIRSI, TELECOM-CIDE, IDRC-CRDI*, 1–45.

Galperin, H., & Mariscal, J. (2016). *Internet y pobreza: Evidencia y nuevas líneas de investigación para América Latina*. CIDE.

Giotsas, V., Zhou, S., Luckie, M., & claffy, kc. (2013). Inferring multilateral peering (pp. 247–258). Presented at the ACM CoNEXT, ACM Press. https://doi.org/10.1145/2535372.2535390

IFT. (2017). *Tercer Informe Trimestral Estadístico 2016* (p. 91). Mexico City: Instituto Federal de Telecomunicaciones.

Intven, H., & Tétrault, M. (Eds.). (2013). *Telecommunications Regulation Handbook*. Washington, DC: World Bank. Retrieved from http://hdl.handle.net/10986/15249

Klöti, R., Ager, B., Kotronis, V., Nomikos, G., & Dimitropoulos, X. (2016). A Comparative Look into Public IXP Datasets. *ArXiv:1611.02624 [Cs]*. Retrieved from http://arxiv.org/abs/1611.02624

Larkin, B. (2013). The Politics and Poetics of Infrastructure. *Annual Review of Anthropology*, *42*(1), 327–343. https://doi.org/10.1146/annurev-anthro-092412-155522

Latour, B. (1991). Technology Is Society Made Durable. *The Sociological Review Monograph*, (38), 103–131.

Latour, B. (1999). *Pandora's Hope: Essays on the Reality of Science Studies* (1 edition). Cambridge, Mass: Harvard University Press.

Latour, B. (2002). Morality and Technology. *Theory, Culture & Society*, *19*(5–6), 247–260. https://doi.org/10.1177/026327602761899246

Law, J. (2015). *STS as Method*. Retrieved from http://heterogeneities.net/publications/Law2015STSAsMethod.pdf

Lievrouw, L. A. (2014). Materiality and media in communication and technology studies: An unfinished project. In T. Gillespie (Ed.), *Media technologies: Essays on communication, materiality, and society* (pp. 21–51).

Metz, C. (2001). Interconnecting ISP Networks. *IEEE Internet Computing*, *5*(2), 74–80. https://doi.org/10.1109/4236.914650

Mexico. (2014, July 14). Ley Federal de Telecomunicaciones y Radiodifusión.

Mexico. (2017, July 24). DOF - Diario Oficial de la Federación. Retrieved from http://www.dof.gob.mx/nota_detalle.php?codigo=5491665&fecha=24/07/2017

Musiani, F. (2013). Network architecture as internet governance. *Internet Policy Review*, *2*(4), 1–9.

Norman, D. A. (2010). *Living with Complexity*. Cambridge, Mass: The MIT Press.

OECD. (2012). *OECD Review of Telecommunication Policy and Regulation in Mexico*. OECD Publishing. https://doi.org/10.1787/9789264060111-en

OECD, & IDB. (2016). *Broadband policies for Latin America and the Caribbean: a digital economy toolkit*. Paris: OECD Publishing.

PCH. (2018, April). Packet Clearing House Report on Internet Exchange Point Locations | PCH. Retrieved November 30, 2018, from https://www.pch.net/ixp/summary

Prior, L. (2008). Repositioning Documents in Social Research. *Sociology*, *42*(5), 821–836.

Richter, P., Smaragdakis, G., Feldmann, A., Chatzis, N., Boettger, J., & Willinger, W. (2014). Peering at Peerings: On the Role of IXP Route Servers. In *Proceedings of the 2014 Conference on Internet Measurement Conference* (pp. 31–44). New York, NY, USA: ACM. https://doi.org/10.1145/2663716.2663757

Rivera, A. (2014, May 2). Forman en KIO el primer IXP de México. Retrieved August 1, 2017, from http://boletin.com.mx/datacenter-almacenamiento-y-respaldo-boletin/item/755-forman-en-kio-el-primer-ixp-de-mexico/755-forman-en-kio-el-primer-ixp-de-mexico

Shaw, K. (2017, December 4). The OSI model explained: How to understand (and remember) the 7 layer network model. Retrieved April 15, 2018, from https://www.networkworld.com/article/3239677/lan-wan/the-osi-model-explained-how-to-understand-and-remember-the-7-layer-network-model.html

Star, S. L., & Bowker, G. C. (2010). How to Infrastructure. In *Handbook of New Media: Social Shaping and Social Consequences of ICTs, Updated Student Edition* (pp. 230–245). London: SAGE Publications Ltd. https://doi.org/10.4135/9781446211304

Telmex. (n.d.). Retrieved from http://blog.telmex.com/?s=fibra+optica

Weller, D., & Woodcock, B. (2013). *Internet Traffic Exchange: Market Developments and Policy Challenges* (OECD Digital Economy Paper No. 207). OECD Publishing. Retrieved from https://econpapers.repec.org/paper/oecstiaab/207-en.htm

Section 2:
Cyber Conflict

COLUMBIA | SIPA
School of International and Public Affairs

CCSA
Cyber Conflict Studies Association

Cyber Conflict History

AUTHORS: Max Smeets and Jason Healey[1]
SERIES EDITOR: Justin Key Canfil EXECUTIVE EDITOR: Jason Healey

About the State of the Field Series

This article is part of the 2017 Cyber Conflict State of the Field (SOTF) paper series, under the auspices of the Cyber Conflict Studies Association and Columbia University's School of International and Public Affairs.

The conference, held annually since 2016, brings together experts from various academic disciplines, including political science, law, economics, and policy research, to define key questions and map the research frontier in the emerging field of cyber conflict studies. The conference is cumulative: each year builds upon past discussions. As a result, discussions have necessarily matured at different rates as new topics are added.

The papers in this series are meant to capture the findings of the 2017 conference. Together, the papers represent the conference attendees' understanding of the present state of the field in the academic study of cyber conflict.

Introduction

The study of cyber history can provide insight into the often complex and obscure dynamics of cyber conflict. An exclusive focus on the present unnecessarily handicaps efforts to understand, categorize and qualify past behavior. Cyber history provides a storehouse of information on past cyber activity, including case studies and datasets that can fuel the formulation of theories and testable hypotheses. In addition to serving as a laboratory—however imperfect—to explain cyber behavior, the study of cyber history also helps academics, policy makers, and practitioners to interpret ongoing developments. To understand why a cyberattack—like the 2016 Democratic National Committee (DNC) email hack[2]—occurred, it is necessary to examine the historical context in light of present conditions. Was the targeting of the DNC espionage, an escalatory attack or criminal behavior? Was it a continuation of long-standing practices? Have similar methods been utilized in the past? Why did the attacker conduct this activity? A relatively recent history often suffices to explain key cyber conflict events and trends, but in some cases, it is necessary to delve further back in time to fully understand the underlying causes.

There is simultaneously too much and not enough cyber history. The secrecy surrounding government organizations and their capabilities as well as the anonymity of attackers complicate the documentation of cyber history. Moreover, the recent and rapid growth of the field means the subject has yet to receive adequate attention from professional historians.[3] Vast amounts of raw data, case study reports, and other documents still await analysis and therefore leave much research left to be undertaken.

It is hardly surprising that SOTF devoted a panel to cyber history, given its importance and the work that remains to be done. SOTF first included a cyber history panel, with Jason Healey moderating, in 2016. Karl Grindal summarized the key views of the panelists and provided a comprehensive overview of canonical works in the field.[4] The 2017 panel built on that of the previous year with only minor changes. The table below provides an overview of the topics on the agenda at the State of the Field conferences in 2016 and 2017.

	TOPICS	
	STATE OF THE FIELD 2016	STATE OF THE FIELD 2017
I	Origins of the Cyber Domain	Conceptual History
II	Development of the Field	History of Cyber Conflict Discourse
III	Eras in Cyber Conflict History	Eras in Cyber Conflict History
IV	Organizational History	Organizational History
V	Operational History	Operational and Strategic History*
VI	History of Non-State Actors	

* On agenda but not discussed during workshop due to lack of time.

Major Takeaways from SOTF 2017

Discerning the continuities and discontinuities of cyber conflict formed the central and overarching theme of the panel. Although the two perspectives were not explicitly compared, elements of each came up throughout the discussion. Participants discussed several "turning" and "tipping" points throughout cyber history as potential points of discontinuity. These tipping points also offer potential qualitative shifts and likely differ across countries/regions, making periodization complex and inherently spatially bounded. Complicating matters further, cyber events (e.g., the Morris Worm, Stuxnet, and Operation Orchard) and non-cyber events (e.g., the Oklahoma City Bombing, the Asian Financial Crisis, and the November 2015 Paris terrorist attacks) alike were identified as potential roots of decisive change in this field.[5] But beyond the many spatial and temporal discontinuities identified were signs of remarkable continuity in the nature of cyber conflict, including an ever-evolving relationship between "cyber" and "info" warfare.

I. Conceptual History

QUESTION(S):	GAP(S):
• How has our perception of cyber-related concepts changed? • How does conceptual ambiguity affect governance?	• The relationship between the prefix "cyber" and other terms (e.g., "info," "computer," etc.).

The workshop session started with a discussion on the historical semantics of cyber-related terms. The term "cyberspace" has long been attributed to William Gibson, who first used it in the 1982 short story "Burning Chrome" and again in his 1984 novel *Neuromancer*.[6] However, it has now been traced back to an earlier etymology. An article in a Norwegian art magazine suggests the term may have first appeared, without gaining currency, in a painting collage produced by Susanne Ussing and Carsten Hof between 1968–1970.[7] John Perry Barlow is credited with introducing "cyberspace" to political discourse in 1996.[8]

Regardless of its origins, the term has been interpreted in numerous ways and embodied a variety of meanings over the years. In assessing the history of cyber-related concepts,[9] panelists identified two primary questions:

- How has our understanding of cyber-related concepts changed?
- How does conceptual ambiguity affect governance?

Following the discussion, panelists made the key observation that it is only possible to understand the nature of cyberspace by assessing its pre-history. They referenced a brief history by several chief architects of the Internet on the technical foundation of the Advanced Research Projects Agency Network (ARPANET).[10] This account emphasizes the decentralized, trust-based nature of the project. The panelists also reiterated a point raised in 2016, that the prefix "cyber" is still often conflated with other terms, such as "info," "computer," or "the Internet." This ongoing conflation and ambiguity of terms hinders policy efforts to establish "rules of the road."

One participant noted the importance of understanding how "cyber" became a military domain, guiding our organizational and strategic thinking. In 2011, the U.S. military forces officially expanded the traditional domains of warfare—air, sea, land, and space—to include cyber.[11] A number of countries have subsequently adopted a similar approach, and in 2016 NATO also officially declared cyberspace a warfare domain.[12] Both the historical origins and the implications of conceptualizing cyberspace in this manner remain ill understood.

Finally, participants noted that the conceptualization of cyberspace and its relevant terminology varies across countries. For example, a report from the East-West Institute states, "Unlike Americans, Russians saw cybersecurity as an inextricable part of a larger discussion on information security."[13] Regional differences are also evident in the interpretation of "cyber sovereignty," which describes a government's goal of exercising control over cyber activities within its own borders.[14]

II. History of Cyber Conflict Discourse

QUESTION(S):	GAP(S):
• How has the discourse surrounding cyber conflict (and the cyber threat) developed over time?	• U.S.-centric. • Limited group of actors analyzed.

The study of discourse and narratives is becoming increasingly important to the field of conflict resolution. Scholars have been analyzing the cyber security discourse since the early 2000s.[15] Myriam Dunn Cavelty makes a compelling and comprehensive argument that cyber threats have been inflated by numerous policymakers.[16] Limiting his scope to the United States, Ralf Bendrath reaches a similar conclusion: "there is no link at all between the cyber threat perception and the real world."[17] Although the field has evolved around the literature on securitization, scholars have also addressed the consequences of cyber threat inflation. Thomas Rid and Robert M. Lee contend that "cyber-angst" is damaging and self-serving and that a more nuanced debate is needed.[18]

Workshop participants addressed a key gap in the current academic discussion on cyber discourse: the lack of scholarship into how cyber threat assessment and general threat assessment affect each other. They highlighted the November 2015 Paris attacks as a potentially interesting case warranting analysis. Historically, the primary focus of cyber conflict has centered on attacks originating in Russia and China. The 2015 terrorist attacks in Paris, while not cyber conflict, catapulted "cyber terrorism" back into focus as a key threat to the general public.

In addition, participants noted that the ongoing discussion of the nature of cyber conflict—often post-cyber incident—takes place in a number of different forums. They pointed out that many excellent insights on recent cyber activity have appeared on Twitter instead of in the mainstream media. The research community needs to think carefully about how best to capture these views to ensure that they will not be lost to reports published in future case studies.

Finally, participants noted that the tendency to attach the prefix "cyber" to other terms, though ongoing, may slow or cease in the future. As one participant observed, we no longer talk about the "digital economy"—it's just the "economy." "Cyber warfare" may someday mirror this trend; "cyber" will come to seem inherent to, and implicit in, "warfare."

III. Eras in Cyber Conflict History

QUESTION(S):	GAP(S):
• How can we divide cyber conflict history into eras? • Which incidents or moments serve as transition points between these eras? • How did institutions develop around cyber conflict in the early era? • How has the balance changed between military operations and intelligence as a matter of doctrine, organization, and practice? • How has cyber conflict history already been divided into "eras"?	• What unique technical and political attributes are linked to these eras? • How do different levels of granularity overlay when we outline the history of cyber conflict?

The third topic on the agenda was "Eras in Cyber Conflict History," with "eras" viewed as frameworks that we impose over events to make sense of them. In his "pre-history" of cyber security, Michael Warner argues that the U.S. government's insights can be categorized into four phases:

- Computers can spill sensitive data and must be guarded (1960s)
- Computers can be attacked, and data can be stolen (1970s)
- We can build computer attacks into military arsenals (1980s and 1990s)
- Others might do the same to us—and perhaps already have (1990s)[19]

Healey, by contrast, identifies three phases of cyber conflict history: realization (1980s), takeoff (1990s–), and militarization (2003–).[20] Awareness of the potential of cyber conflict grew through various events in the 1980s and 1990s, including the Morris Worm (1988), the Wank Worm (1989), the Cuckoo's Egg (1989), Michelangelo (1992), Eligible Receiver (1997), and Solar Sunrise (1998). Post-2000, several events, including the Code Red Worm (2001), the SQL Slammer Worm (2003), JSF espionage (2009), and Stuxnet (2010), increased the sense that targeted cyber activity would only intensify.[21]

Workshop participants echoed two points from earlier cyber history discussions: first, that non-cyber events have, at times, been instrumental in shaping cyber policy. For example, the murder of 168 and injuring of hundreds more in the 1995 Oklahoma City bombing likely impacted the United States' formal response to cyber threats. Following the attack, the Clinton administration formed the President's Commission on Critical Infrastructure Protection, which released a report stressing the need to implement new measures around cyber security.[22] Second, that periodization differs significantly across regions. For example, as was noted, Chinese experts published several articles in the mid-1990s on the United States' growing interest in information warfare.[23] The Chinese also likely learned important lessons from Operation Orchard, Israel's 2007 use of electronic warfare to neutralize Syrian radar systems, facilitating an airstrike on a suspected nuclear reactor.

In the Middle East, rapid change occurred from 2009–2012, when Stuxnet was revealed and popular uprisings partially fueled by online mobilization hit multiple autocratic regimes. The events of the Arab Spring, though often forgotten, are arguably as relevant as Stuxnet to the course of cyber policy in this region. To quell protests, governments across the Arab world severely tightened Internet controls, arrested bloggers, stole passwords for social media accounts, and in several countries (Egypt, Libya, and Syria), even attempted to shut down the Internet completely. The perspective of many autocratic regimes in the region was that the cyber threat was coming from multiple vectors simultaneously.[24]

While these examples of periodization in China and the Middle East are illuminating, participants noted that a broad overview of which events were essential for different countries or regions is still missing.

IV. Organizational History

QUESTION(S):	GAP(S):
• How have legislation, rules, and doctrines evolved to address cyber threats? • How have major cyber incidents impacted organizational policies or structures? • Have doctrinal and organizational developments abroad been secondary or primary factors for domestic organizational change? • How have organizations adopted and incorporated offensive cyber capabilities? • How can non-state actors help to establish and cascade cyber norms?	• Weighted toward institutions that have either defended or threatened the United States. • The impact and evolution of non-governmental organizations is underexplored. • Interstate cooperation (particularly on offensive).

Scholarly discussion of organizational responses to cyber conflict has taken place on several levels. Most research has focused on the relationship *between* the public and private sectors in *defending* cyberspace. The specific conditions under which information should be shared and the extent to which non-voluntary standards should guide the public-private relationship form a central consideration within organizational responses.[25]

Workshop participants identified the high level of secrecy as a barrier to discussion about how the links *within* government relate to the conduct of *offensive* operations.[26] Future research is necessary on the relationship between offensive intelligence operations and offensive military operations—or what the U.S. military calls Computer Network Exploitation (CNE) and Computer Network Attack (CNA).

Another often-overlooked area of study within the field is the role of informal "trust networks." Milton Mueller writes, "there is a strong and persistent tension between state sovereignty, which is territorially bounded, and the nonterritorial space for social interaction created by networked computers."[27] This tension, and the non-nation-centered arrangements that may follow from it, deserve further analysis.

From an organizational perspective, the intertwining of cyber warfare and information warfare has been critical to some governments while an anathema to others. Following the DNC hack, many experts have prioritized a reorientation toward countering information operations. As one participant observed, this focus is nothing new: "In the United States, there was a debate about it 20 years ago. Interestingly, the term 'cyber' was purposefully separate to make sure it becomes 'something' on its own; it showed that it went beyond psychological warfare. Yet, some now go down this road, which back then was seen as a dead end." The voluminous documentation on information warfare following the Kosovo War and the liberation of Kuwait in the 1990s supports this point, as does the outlining of the elements underlying "information warfare" in a 1976 paper by Boeing engineer Thomas P. Rona.

Finally, there remain open questions about how cyber conflict relates to issues of global order. Despite ongoing initiatives from the Global Commission on Cyberspace, it is unclear how we can embed cyber regimes into broader global stability. Note that even though the current reading lists include references on "cyber norms," several participants argued that it would be better to address that issue separately.

V. Operational and Strategic History

QUESTION(S):	GAP(S):
• How have operators detected, identified, responded to, and recovered from major cyber incidents? • Can current cyber defenders or policymakers draw any lessons from past operational incidents? • Is there fundamental continuity or discontinuity in cyber operations? • Which early works helped to shape strategic thinking on cyber conflict?	• Still lacking comprehensive case studies. • Analysis of incidents using historical datasets is still limited to a few scholars. • There is no clear sense of how pre-cyber operations link to current-day cyber activity.

There are many open questions when it comes to operational and strategic cyber history. Unfortunately, the discussion was cut short due to time limitations.

Concluding Remarks

As the panel discussion emphasized, cyber history is simultaneously characterized by "continuous change" as well as "turning points." While time constraints prevented participants from providing a comprehensive overview of the field, the identified gaps establish a natural starting point for discussion at the next SOTF. At the next workshop, it may be especially worthwhile to focus on the operational and strategic history of cyber conflict. This could include an overview of known cases that have received insufficient attention.

Important Works

Conceptual History: Meaning of Cyber(space)

PRIMARY READING

Cornish, Paul. (2015) Governing Cyberspace through Constructive Ambiguity. *Survival*, 57(3): 153-176.

Gibson, William. (1984) *Neuromancer*. London: Victor Gollancz.

Hayden, Michael. (2011) The Future of Things "Cyber." *Strategic Studies Quarterly*, 5(1): 3-7.

Rid, Thomas. (2016) *Rise of the Machines*. New York: W.W. Norton & Company.

Wolff, Josephine. (2016) What we Talk About when we Talk About Cybersecurity: Security in Internet Governance Debates. *Internet Policy Review*, 5(3).

SECONDARY READING

Betz, David and Tim Stevens. (2011) *Cyberspace and the State: Toward a Strategy for Cyber-Power*. London: Routledge.

Ebert, Hannes and Tim Maurer. (2013) Contested Cyberspace and Rising Powers. *Third World Quarterly*, 34(6): 1054-1074.

Giles, Keir and William Hagestad II. (2013) Divided by a Common Language: Cyber Definitions in Chinese, Russian and English. In *NATO CCD COE Publications*, K. Podins, J. Stinissen and M. Maybaum (eds.), 5th International Conference on Cyber Conflict.

Herrera, Geoffrey L. (2006) *Technology and International Transformation: The Railroad, the Atom Bomb, and the Politics of Technological Change*. Albany: State University of New York Press.

Kello, Lucas. (2013) The Meaning of the Cyber Revolution: Perils to Theory and Statecraft. *International Security*, 38(2): 7-40.

Maurer, Tim and Robert Morgus. (2014) Compilation of Existing Cybersecurity and Information Security Related Definitions. New America Foundation.

History of Cyber Conflict Discourse

PRIMARY READING

Bendrath, Ralph. (2003) The American Cyber-Angst and the Real World — Any Link? In *Bombs and Bandwidth: The Emerging Relationship Between Information Technology and Security*, R. Latham (ed.). New York: The New Press.

Dunn Cavelty, Myriam. (2008) *Cyber-Security and Threat Politics: U.S. Efforts to Secure the Information Age*. Abingdon: Routledge.

Hansen, Lene and Helen Nissenbaum. (2009) Digital Disaster, Cyber Security, and the Copenhagen School. *International Studies Quarterly*, 53: 1155-1175.

SECONDARY READING

Walt, Stephen M. (2010) Is the Cyber Threat Overblown? *Foreign Policy*, http://walt.foreignpolicy.com/posts/2010/03/30/ is_the_cyber_threat_overblown.

Eras in Cyber Conflict History

PRIMARY READING

Bamford, James. (2002) *Body of Secrets: Anatomy of the Ultra-Secret National Security Agency*. New York: Anchor Books.

Healey, Jason, ed. (2013) *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012*. Vienna, VA: Cyber Conflict Studies Association.

Kahn, David. (1996) *The Codebreakers: The Story of Secret Writing*. New York: Scribner.

Kaplan, Fred M. (2016) *Dark Territory: The Secret History of Cyber War*. New York: Simon & Schuster.

Rid, Thomas. (2016) *Rise of the Machines*. New York: W.W. Norton & Company.

SECONDARY READING

Aldrich, Richard. (2010) *GCHQ: The Uncensored Story of Britain's Most Secret Intelligence Agency*. London: Harper Collins.

Bamford, James. (1983) *The Puzzle Palace: A Report on America's Most Secret Agency*. New York: Penguin Books.

Harris, Shane. (2015) *@War: The Rise of the Military-Internet Complex*. New York: Houghton Mifflin Harcourt.

Hayden, Michael. (2016) *Playing to the Edge: American Intelligence in the Age of Terror*. New York: Penguin Books.

Warner, Michael. (2012) *Cybersecurity: A Pre-History. Intelligence and National Security*, 27(5): 781-799.

Yardley, Herbert O. (2004) *The American Black Chamber*. Annapolis: Naval Institute Press.

Organizational History

PRIMARY READING

Johnson, David R. and David G. Post. (1996) Law and Borders: The Rise of Law in Cyberspace. *Stanford Law Review*, 48b: 1367-1402.

Lessing, Lawrence. (1998) The Laws of Cyberspace. Working paper, Harvard Law School. https://cyber.law.harvard.edu/works/lessig/laws_cyberspace.pdf.

Lipner, Steven B. (2015) The Birth and Death of the Orange Book. *IEEE Annals of the History of Computing*, 37(2): 19-31.

Maurer, Tim. (2011) Cyber Norm Emergence at the United Nations - An Analysis of the Activities at the UN Regarding Cyber-Security. Belfer Center, Discussion Paper #2011-11, Explorations in Cyber International Relations Discussion Paper Series.

Ruffini, Joseph. (1999) 609 IWS Chronological History. Department of the Air Force.

SECONDARY READING

Carr, Madeleine. (2016) Public-Private Partnerships in National Cyber-Security Strategies. *International Affairs*, 92(1): 43-62.

Choucri, Nazli, Stuart Madnick, and Jeremy Ferwerda. (2014) Institutions for Cyber Security: International Responses and Global Imperatives. *Information Technology for Development*, 20(2): 96-121.

DeNardis, Laura. (2015) The Internet Design Tension between Surveillance and Security. *IEEE Annals of the History of Computing*, 37(2): 72-83.

Finnemore, Martha and Duncan B. Hollis. (2016) Constructing Norms for Global Cybersecurity. *The American Journal of International Law*, 110(3): 425-479.

Hurwitz, Robert. (2014) The Play of States: Norms and Security in Cyberspace. *American Foreign Policy Interests*, 36(5).

SECONDARY READING

Nye, Joseph. (2014) *The Regime Complex for Managing Global Cyber Activities*. Harvard Kennedy School Belfer Center. Available at http://belfercenter.hks.harvard.edu/files/global-cyber-final-web.pdf.

Rosensweig, Paul. (2010) The Organization of the United States Government and Private Sector for Achieving Cyber Deterrence. In National Research Council, *Proceedings of a Workshop on Deterring CyberAttacks: Informing Strategies and Developing Options for U.S. Policy*. Available at sites.nationalacademies.org/cs/groups/cstbsite/ documents/webpage/cstb_059443.pdf.

Yost, Jeffrey R. (2015) The Origin and Early History of the Computer Security Software Products Industry. *IEEE Annals of the History of Computing*, 37(2): 46-58.

History of Operational and Strategic Thinking

PRIMARY READING

Arquilla, John and David Ronfeldt. (2001) *Networks and Netwars: The Future of Terror, Crime, and Militancy*. Santa Monica: Rand Corporation.

Buchanan, Ben and Michael Sulmeyer. (2016) Hacking Chads: The Motivations, Threats, and Effects of Electoral Insecurity. Paper, Cyber Security Project. Harvard Kennedy School Belfer Center. Available at www.belfercenter.org/sites/default/files/legacy /files/hacking-chads.pdf.

Denning, Dorothy E. (1998) *Information Warfare and Security*. New York: Addison-Wesley Professional.

Harknett, Richard J. (1996) Information Warfare and Deterrence. *Parameters*, 26(Autumn): 93-107.

Libicki, Martin C. (1995) What is Information Warfare? Strategic Forum, No. 28, Washington: National Defense Univ., Institute for National Strategic Studies.

Lindsay, Jon. R. (2013) Stuxnet and the Limits of Cyber Warfare. *Security Studies*, 22(3): 365-404.

Ottis, Rain. (2008) Analysis of the 2007 Cyber Attacks Against Estonia from the Information Warfare Perspective. Proceedings of the 7th European Conference on Information Warfare, 163.

Rattray, Gregory J. (2001) *Strategic Warfare in Cyberspace*. Boston: MIT Press.

Schwartau, Winn. (1996) *Information Warfare: Second Edition*. New York: Thunder's Mouth Press.

Zetter, Kim. (2014) *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon*. New York: Crown Publishers.

SECONDARY READING

Denning, Peter. (1990) *Computers Under Attack: Intruders, Worms, and Viruses*. New York: Addison-Wesley.

Farwell, James P. and Rafal Rohozinski. (2011) Stuxnet and the Future of Cyber War. *Survival: Global Politics and Strategy*, 53(1): 23-40.

Herzog, Stephen. (2011) Revisiting the Estonian Cyber Attacks: Digital Threats and Multinational Responses. *Journal of Strategic Security*, 4(2): 49-60.

Hollis, David. (2011) Cyberwar Case Study: Georgia 2008. *Small Wars Journal*.

Oder, Joseph E. (1994) Digitizing the Battlefield: The Army's First Step to Force XXI. *Army*: 36-42.

Rid, Thomas. (2017) Disinformation: A Primer in Russian Active Measures and Influence Campaigns. Hearing before the Select Committee on Intelligence, United States Senate.

Toffler, Alvin and Heidi Toffler. (1995) *War and Anti-War: Making Sense of Today's Global Chaos*. New York: Grand Central Publishing.

About the Authors

Dr. Max Smeets is a cybersecurity postdoctoral fellow at Stanford University Center for International Security and Cooperation (CISAC). He is also a non-resident cybersecurity policy fellow at New America, and Research Associate at the Centre for Technology & Global Affairs, University of Oxford.

Jason Healey is a Senior Research Scholar at Columbia University's School for International and Public Affairs specializing in cyber conflict, competition and cooperation.

End Notes

[1] The authors would like to thank Aaron Brantly and Karl Grindal for their helpful comments on earlier drafts of this report.

[2] Thomas Rid, "How Russia Pulled Off the Biggest Election Hack in U.S. History," *Esquire*, (20 October 2016), retrieved from: www.esquire.com/news-politics/a49791/russian-dnc-emails-hacked/

[3] As one participant noted, "Federal historians are a dying breed—they get cannibalized before something else [which directly impacts mission execution]."

[4] Jason Healey and Karl Grindal, "The Cyber Conflict State of the Field Workshop Report 2016," The Cyber Conflict Studies Association.

[5] The workshop participants, however disagreed on what "decisive change" would entail in this field.

[6] William Gibson, Burning Chrome, *Omni*, 46 (1982, July); Gibson, *Neuromancer*, (London: Victor Gollancz, 1984)

[7] Jacob Lillemose and Mathias Kryger, "The (Re)invention of Cyberspace," *Kunstritikk*, (24 August 2015), retrieved from: www.kunstkritikk.no/kommentar/the-reinvention-of-cyberspace/

[8] John Perry Barlow, "A Declaration of the Independence of Cyberspace," *Electronic Frontier Foundation*, (1996), retrieved from: https://projects.eff.org/~barlow/Declaration-Final.html

[9] Note that previous research efforts—e.g., Thomas Rid's *Rise of the Machines*—show that there is an (academic) market for this type of work.

[10] Barry M. Leiner, Vincent G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Larry G. Roberts, and Stephen Wolff, "A Brief History of the Internet," *ACM SIGCOMM Computer Communication Review*, 39:5 (2009): 22–31. First published 1992

[11] Keith B. Alexander, "Warfighting in Cyberspace," *Joint Force Quarterly*, 46 (2007): 58–61

[12] Tomáš Minárik, "NATO Recognises Cyberspace as a 'Domain of Operations' at Warsaw Summit," NATO Cooperative Cyber Defence Centre of Excellence, (21 July 2016), retrieved from: https://ccdcoe.org/nato-recognises-cyberspace-domain-operations-warsaw-summit.html

[13] Karl Frederick Rauscher, "Russia-U.S. Bilateral on Cybersecurity: Critical Terminology Foundations," East-West Institute, (26 April 2011), retrieved from: www.eastwest.ngo/idea/russia-us-bilateral-cybersecurity-critical-terminology-foundations

[14] Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*, (New York: W.W. Norton Company, 2015)

[15] Helen Nissenbaum, "Hackers and the Contested Ontology of Cyberspace," *New Media & Society*, 6:2 (2004): 195–217; Nissenbaum, "Where Computer Security Meets National Security," *Ethics and Information Technology*, 7:2 (2005): 61–73; Rachel Yould, "Beyond the American Fortress: Understanding Homeland Security in the Information Age." In *Bombs and Bandwidth: The Emerging Relationship Between Information Technology and Security*, ed. Robert Latham (New York: The New Press, 2003); James Der Derian, "The Question of Information Technology in International Relations," *Millennium*, 32:3 (2003): 441–456

[16] But also notes that, so far, the cyber issue has *not* been securitized. Myriam Dunn Cavelty, *Cyber-Security and Threat Politics: U.S. Efforts to Secure the Information Age*, (Abingdon: Routledge, 2008)

[17] Ralf Bendrath, "The American Cyber-Angst and the Real World – Any Link?" In *Bombs and Bandwidth*, ed. Latham; also see, Bendrath, "The Cyberwar Debate: Perception and Politics in U.S. Critical Infrastructure Protection," *Information & Security*, 7 (2001): 80–103

[18] Robert M. Lee and Thomas Rid, "OMG CYBER! Thirteen Reasons why Hype Makes for Bad Policy," *The RUSI Journal*, 159:5 (2014): 4–12

[19] Michael Warner, "Cybersecurity: A Pre-history," *Intelligence and National Security*, 27:5 (2012): 781–799

[20] Jason Healey (ed.), *A Fierce Domain: Conflict in Cyberspace*, 1986 to 2012, (Vienna, VA: Cyber Conflict Studies Association, 2013)

[21] One participant observed that a lot of low-hanging fruit remains when it comes to data gathering and analysis. There is still much more to be done in terms of coding these events.

[22] President's Commission on Critical Infrastructure Protection, "Critical Foundations: Protecting America's Infrastructures," (13 October 1997), retrieved from: www.fas.org/sgp/library/pccip.pdf

[23] Warner, "Cybersecurity: A Pre-history"

[24] It was noted that hybrid warfare, critical infrastructure attacks, and surveillance characterized this regional perspective.

[25] Sue Eckert, "Protecting Critical Infrastructure: The Role of the Private Sector," University of Pittsburg, 2006, available at: www.ridgway.pitt.edu/Portals/1/pdfs/Publications/Eckert.pdf; Kenneth Neil Cukier, Viktor Mayer-Schoenberger, and Lewis M. Branscomb, "Ensuring (and Insuring?) Critical Information Infrastructure Protection," Working Paper, John F. Kennedy School of Government, Harvard University, 11 October 2005

[26] Max Smeets, "Organisational Integration of Offensive Cyber Capabilities: A Primer on the Benefits and Risks." In H. Rõigas, R. Jakschis, L. Lindström, and T. Minárik (eds.), *Defending the Core*, 9th International Conference on Cyber Conflict, (Tallinn: NATO CCD COE Publications, 2017). The workshop participants did not discuss in any detail the nature of these relationships, whether that interpersonal, agency, or operational.

[27] Milton F. Mueller, *Network and States: The Global Politics of Internet Governance*, (Cambridge: The MIT Press, 2010), p. 1

The Cyber Conflict Studies Association (CCSA) promotes and leads international intellectual development efforts to advance the field of cyber conflict research. These activities include workshops that bring together professionals from industry, academia and government to discuss strategic issues surrounding cyber conflict and the publication of insightful research articles and position papers and books. CCSA also plays an important role in our national cyber-readiness strategy, serving as a resource for national security decision-makers and helping to frame and promote national cyber conflict policy. CCSA brings together the best and the brightest individuals in the field of cyber conflict study to further the goals of national security and the field of cyber.

This work was supported in part by the Minerva Research Initiative. The Minerva Research Initiative, administered jointly by the Office of Basic Research and the Office of Policy at the U.S. Department of Defense, supports social science research aimed at improving our basic understanding of security.

Funded in part by the Office Naval Research under the OSD Minerva program, grant number: N00014-17-1-2423.

COLUMBIA | SIPA
School of International and Public Affairs

Tactical and Operational Dynamics of Cyber Conflict

AUTHORS: Trey Herr and Roberta Stempfley
SERIES EDITOR: Justin Key Canfil

Introduction

A recurring panel in the two-year history of the conference focuses on the tactical and operational dynamics of cyber conflict. The tactical and operational dynamics of cyber conflict deal with the mechanics of engagement, hewing away from issues like coercion and escalation towards the mechanics of capabilities development, employment, and maintenance. The operational focuses on the relationship between engagements and higher order processes like targeting, command & control, and training. The tactical emphasizes the minutiae of a single engagement, like the mechanics of a DNS reflection attack. Related panels covered Strategic Dynamics, as well as the Law, History, Intelligence processes, and Economics of cyber conflict.

The first conference in 2016 did as much to stoke discussion as organize the literature of cyber conflict. On tactical and operational dynamics, those assembled discussed the impact cyber operations would have on the tactical and operational levels of war as well as differences in the doctrinal development of different states and the relative offensive or defensive dominance of cyberspace. While there was mention of non-state actors, one of the biggest areas of expansion in the second year of the panel was to critically examine the role of the private sector in provisioning the infrastructure on which many of these engagements take place.

About the State of the Field Series

This article is part of the 2017 Cyber Conflict State of the Field (SOTF) paper series, under the auspices of the Cyber Conflict Studies Association and Columbia University's School of International and Public Affairs.

The conference, held annually since 2016, brings together experts from various academic disciplines, including political science, law, economics, and policy research, to define key questions and map the research frontier in the emerging field of cyber conflict studies. The conference is cumulative: each year builds upon past discussions. As a result, discussions have necessarily matured at different rates as new topics are added.

The papers in this series are meant to capture the findings of the 2017 conference. Together, the papers represent the conference attendees' understanding of the present state of the field in the academic study of cyber conflict.

Transition from 2016

In setting up the 2017 panel, we tried to bring together groups of questions that addressed similar topics while also focusing on the core of what tactics and operations would entail. We built on the topics above from the 2016 panel and modified them in expanding to the workshop outline for this year. The 2016 panel was broken into four sections: Cyber Power at the Tactical and Operational

Key Cross-Cutting Questions

What aspects of tactical operations are physics, and cannot be changed, such as “speed of light” or “scale-free networks.” What can we aspects change like “difficulty of attribution” or “offense dominance?” Indeed, what is the full set of all such aspects?

How do these dynamics drive tactical and operational behavior i.e. effective defense requiring timely intelligence collection?

How do these dynamics differ from the tactical and operational dynamics in other domains?

Levels, Legal and Ethical Considerations, Command & Control, and Organizational Considerations as well as discussion the formation of a cyber service. To adapt this organization, rather than create something new from whole cloth, we made two notable modifications to last year’s organization. First, we reorganized some categories, breaking things down along more process-oriented lines so that concepts like Organizational Process didn’t obscure meaningful internal distinctions. Second, we moved some literature to other panels; the discussion of both norms, a largely strategic issue, and law, self-evidently legal, fit better in the context of other panels.

Takeaways from 2017

This year’s panel on tactical and operational dynamics built on discussion from the previous year, taking key questions and literature and adding to them. In the discussion, we saw three overarching themes which reflect in our comments about how to structure this area of cyber conflict research going forward.

1. There were a multitude of questions about the nature of cyberspace and the how the environment of cyber conflict could impact everything from building offensive capabilities to attributing attacks. This addresses a dicey line between disciplines as social scholars are asking about the nature of the computing and networked environment. This is a valuable area of work but one that emphasize careful literature review and resist the temptation to reinvent basic concepts like Benkler’s articulation of the infrastructure, logical, and content layers of the internet or Clark’s control point analysis.[1]
2. This panel and Strategic Dynamics still share some overlapping concerns. Some of this is due to the fluidity of operations in cyberspace. Much like the strategic corporal concept in insurgency, much of what can take place on a computer system can have outsize political impact. Ideas like anarchy, while theoretically intriguing, are largely structural in nature and thus beyond a reasonable discussion of tactical behaviors. In the interests of research coherence, more can be done to specify where concepts lie at the Strategic level vice the Operational or Tactical. We have endeavored to do so here.
3. There is a tremendously intimate mix of technically focused work out of computer science and operations research with political science, doctrinal analysis, and military science. The resulting amalgamation of methods, questions, and topics is difficult to hone into a coherent research area but the full diversity of work deserves our attention as it often talks to each other, if unknowingly. It would benefit future discussion on these topics if they more directly included computer science alongside the other represented disciplines. Friction will be inevitable but both instructive and valuable.

The rest of the paper covers the major questions and literature brought up in this year’s discussion and closes with several recommendations for next year. Each section outlines some of the topics included in the category then proceeds to summarize the key questions, a combination of those carried forward from the 2016 panel and those generated for and during the discussion in 2017. Based on the discussion at the 2017 event and a recent paper on core readings in cyber conflict course syllabi, the following have emerged as three canonical works in cyber conflict:

- Herbert S. Lin, Kenneth W. Dam, and William A. Owens, eds., *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities* (Washington, D.C.: National Academies Press, 2009)

- Thomas Rid. "Cyber War Will Not Take Place." Journal of Strategic Studies. 35:1 (2012): 5- 32. Or Rid, Thomas, 2013. *Cyberwar Will Not Take Place* (London: Hurst and Co.)
- Libicki, Martin C., *Cyberdeterrence and Cyberwar*, RAND Report, Santa Monica: Rand Corp., 2009

2017 State of the Field: Tactical and Operational Dynamics in Cyber Conflict

Structural Issues and Predicate Questions

Across the discussion there were repeated questions about the structure of cyberspace and conflict that weren't well addressed by other sections of the panel, for example, the relative offensive dominance of cyberspace. What ties the structural issues topic together is their consideration of the underlying rules and phenomena in which tactics and operations exist. While not a discrete section in the original design of the 2017 panel, these questions cropped up repeatedly and deserve explicit treatment. Technical Foundations is split apart to emphasize the role that the computing and network environments play in enabling and constraining cyber operations.

- What new metaphors could be useful to describe the environment of cyber conflict? What are the advantages and disadvantages of these different approaches?
  - humanitarian intervention
  - insurgency/counter-insurgency
  - climate change
  - public health
- What advantages or disadvantages extend from using offense vs. defense as a metaphor in studying cyber conflict?
  - Will this be true in the next 5, 10, 15 years?

Structural Issues and Assumptions

- What would be included in a taxonomy of actors, threats, events, capabilities, and key processes for cyber conflict at the tactical and operational levels?
- How have scholars and/or practitioners differentiated and operationalized the varying levels of war in this research area?[2] What are the merits of different approaches?
- Is cyberspace offense dominant?[3] Will it remain so and what are the implications?
- How have scholars assessed that non-Western perspectives, especially Chinese and Russian, differ on the structure and fundamental behavior of cyberspace?[4]
  - How have the differences between these perspectives changed in the last decade?
  - Where is this understanding best applied in US defensive efforts? Technical foundations? Doctrine? Strategy?
- How have or could scholars evaluate the assumptions underlying these questions, to understand which might change and under what conditions that change might take place?
- What is the relationship between the structure of the environment of cyberspace and the structure of organizations which adapt or are created to operate within it?[5]
  - How is this relationship changing with new technology like cloud-computing and more accessible machine learning resources?

Technical Foundations

- Is cyberspace subject to basic physics or engineering rules scholars need to consider, common dynamics that can be agreed to?
  - How might certain versions of these rules better complement some metaphors describing cyberspace over others?
- Cyberspace is often fungible and can be shaped by the participants; what are the structure, roles, and varying strategies of engagement for these actors?
- How have scholars weighed or measured the competing interests of the private sector and the US Government?
  - Where have or could these interests align to present opportunities for leverage by one party or the other?

Techniques and Technology

What are the key technologies and techniques that underpin tactical and operational cyber conflict? This section covers the specific processes used to identify, develop, deploy, maintain, and defend against cyber capabilities. Attribution deals with the range of methods used to identify and trace cyber operations, including the norms around conducting and publicizing this attribution. Generating and Maintaining Capabilities looks how the tools and material of cyber conflict is born, lives, and dies. This is a particularly expansive set of topics but combined here to promote the idea of these processes and technology interacting in a combined lifecycle.

Attribution[6]

- Under what conditions is attribution hard or easy?[7]
    - What are the implications from this?
- How do actors distinguish between espionage and OPE?[8]
    - Under what circumstances are these distinctions useful, or not useful? And what are the implications?
- What is the role of the private sector in attribution vis a vis the state, particularly the US Government?
    - Does the state have primacy?
    - Are there underexplored benefits or costs to private sector attribution?[9]
- Are there different norms, rules of behavior, for attribution?
    - Do these norms differ between states and non-state actors?
- How many of these concepts are true now, only in this moment, vs. true for all time?
    - How could attribution change in the event of conflict e.g. between US and China in South China Sea?

Generating and Maintaining Capabilities

- What is necessary to generate a cyber capability?[10]
- What sort of offensive or defensive capabilities can be generated without advance preparation?[11]
- What knowledge and resources are required to create high levels of effect? What are barriers to entry?[12]
- What is the minimal amount of effort or resources for an organization to operate effectively?
    - How do the malware market and the behavior of non-state actors impact the generation of these capabilities by states?[13]
- How does the process to generate information or influence as an offensive cyber capability differ from the development of software?
    - What impact do these differences have on theorizing around either the process to generate capabilities or outcomes from that process?[14]
- What is necessary to sustain a cyber operation?[15]
- How can actors manage a stockpile/arsenal of capabilities in a way that differentiates between activities that need new engineering input and those that don't? [The Great Kitchen Analogy]
    - Like cooking—some things needed fresh, for some there are substitutes
    - UK and French use similar ingredients, but everyone prefers French
    - Key is that integration, not a black box → integration is where the human factor comes in
- What are our adversaries and allies learning from the US about the modularity in offensive capabilities and how does this learning advantage overall global stability?
- What distinguishes the generation of offensive or defensive cyber capabilities from their maintenance or regeneration?[16]
- What resources, skills, or incentives influence the process of code or exploit development, reuse or proliferation?[17]
    - How quickly do offensive cyber capabilities decay in value/utility?[18]
    - What is the relative importance of single vulnerabilities versus a chain of such flaws to techniques to discovery or exploit them?

- How do we define and categorize offensive capabilities in cyber conflict?[19]
- How can organizations engineer/employ cyber capabilities with consideration for proportionality and proliferation?[20]
- What assumptions about the process to develop or employ offensive capabilities in cyberspace will be voided or altered in a crisis vs. in peacetime?
- What characteristics define damage in cyberspace?[21]
  - Does a definition of damage include reversible effects? Can it be quantified?
- What would be included in a "strategic toolkit" for cyber conflict?

Doctrine

This section covers issues of force employment and doctrinal development. There is rich potential for comparative work to evaluate the relative development, overlap, and key distinctions between the cyber operations doctrine of major cyber powers. This is of particular value in distinguishing between Western conceptions of tactics and operations and those of other states like Russia and China in answering many of the questions posed above. As cyber evolves both as a domain for operations and a domain to be integrated to achieve national objectives, there is emerging opportunity to understand how and when cyber and other operating domains are best employed and aligned.

Force Employment

- What are prominent taxonomies for effects and capabilities in cyber conflict?[22]
  - What are advantages or disadvantages to each of these approaches?
  - What would a universal taxonomy for these effects and capabilities include? How would it be structured?
- How are cyber operations integrated with conventional military capabilities?[23]
  - How could they in future?[24]
  - How do these approaches vary between different national doctrine?[25]
- Should/does the central role of infrastructure providers and software vendors change our conception of what "warfighting" is?
- How do conventional military or cyber operations combine with information operations?[26]
- What is the relationship between tactical engagements and a campaign?[27]
- How does secrecy impact the development of doctrine for, and exercise of, offensive cyber operations?[28]
- How do researchers theorize about intermediaries or vendors, e.g. Amazon, as a combatant?
  - Are vendors more immediately in the 'line of fire' on offense or defense?
- What factors would influence the distribution of forces on the battlefield/what are the peculiarities of battlefield use?[29]
- What are the core differences in cyber operations doctrine between the major powers?[30]

Organizations and Process

Who are the organizations involved in cyber conflict? This covers much more than what might be found on the battlefield given an environment built, shaped, and provisioned by people—largely the private sector. Considering the interests and behavior of actors like Microsoft and Akamai is as important for many tactical and operational questions of conflict. This section also contains more specifically battlefield issues including aspects of how to design and execute a Command & Control apparatus for cyber conflict at the tactical and operational levels.

Process and Categorization Questions

- What are the advantages and disadvantages of different approaches to categorizing the actors in cyber conflict?[31]
  - What can be learned from these varying approaches?
- How do scholars differentiate between information infrastructure firms involved in specific sectors vs. those with cross-

cutting impact? What are the advantages or disadvantages of these approaches?

  - How do the interests of these actors differ?[32]
  - How do these differences have regulatory or governance significance?

- Where do the interests of infrastructure providers differ from those of combatants?
- What are the mechanics and effects of information sharing e.g. in the ISAC/ISAO model?[33]
  - Under what conditions does information provide value to organizations? How does this value manifest and why?
- What is the influence of the intelligence community/culture on assumptions, beliefs, and expectations around cyber operations?
- How do states share cyber capabilities, offensive or defensive?
  - Under what circumstances would they want to?
  - How would the incentives, or mechanics, of sharing capabilities change in crisis vs. in peacetime?

Command and Control (C2) Considerations

- How should authorization for cyber operations be structured? Offense? Defense?[34]
- Are command and control constructs for cyber operations the same as they are in physical operations?[35]
- How should planning and targeting for cyber operations be conducted?[36]
- How will states and other actors integrate autonomous and rapid, automated, systems into their C2 process?[37]
- What confidence levels or probabilities are necessary for decision-making by national authorities?[38]
- What models exist for coordination in C2 between the private sector and the state?
  - How might this influence offensive private sector activities in response to an attack?[39]
- What distinctions exist in theorizing over the C2 of cyber operations vs. the integration of cyber operations in C2 of existing military operations?

Training and Skills

What are the educational processes required to develop skilled cyber operators? The key questions here are still developing but important enough to merit a distinct sub-category. These topics address the workforce involved in conflict. Some of the questions can go on to inform other topics like managing organizations responsible for cyber operations, like looking at the overlap of skills required for offense and defense.

- Are the skills and organizational capabilities required for offense and defense different?[40]
  - To what extent? How and where are skills different?
- What is the minimum level of training or skillset necessary for an individual to operate on offense, on defense?

Summary & Recommendations

The conversation during the panel was wide-ranging. Specific questions on how a piece of malware might be built of proliferated quickly spiraled into the norms around attributing such software's use and how to distinguish between its various potential effects. There was a recurring theme of mixing what seemed to be strategic questions, issues of escalation and deterrence for instance, into the minutia of organizational process issues like authorizing cyber operations in the United States. There was also a repeated emphasis on highly securitized topics, perhaps owing political science playing a prominent role in the intellectual development of the panel and many in the room. There is a tremendous influence of technology on many of these questions and even more so the companies and individuals that develop and maintain it. Stemming from these observations, we make two recommendations for researchers generally and next year's State of the Field in particular.

Split Tactical and Operational Research into Separate but Related Camps

The content covered by this paper should be split into two research areas—Tactical and Operational Cyber Conflict. Tactical should encompass all those activities and questions dealing with the conduct of a single engagement or activity while operational grapples with the ramifications and complexity of multiple engagements linked together. Deploying cyber capabilities on a battlefield works as a good example. The higher order planning issues, for example the question of how to integrate these capabilities within the planning process for a maneuver unit and matching cyber effects with supporting or direct fires—these are operational questions. Questions looking at lower level issues like the performance characteristics of deploying this capability over the Bluetooth protocol vs. 802.11g or the potential rate of decay in the utility of the vulnerability the capability depends on, are questions at the tactical level. Tactical thus encompasses a larger portion of the direct technical questions and conceptually sits closer to the metal. Operational deals with a higher level of abstraction and includes organizational, process, and many doctrinal issues.

There is a potential organizing logic between strategic, operational, and tactical levels.

- *How do I build it? → Tactical*
- *How should I manage or employ it? → Operational*
- *Why should I build or use it? → Strategic*

Push Discussion Beyond the Battlefield

There is a major body of work to be done on the integration of cyber capabilities on the battlefield but these questions are not the whole or even most questions in tactical and operational cyber conflict. During the 2017 workshop, there was a recurring challenge to integrate private sector actors in the discussion as something more than a structural oddity. Cyber conflict as a broad and interdisciplinary area of research should not be limited to understanding how the military will behave in the "cybered" era. Cyberspace is a man-made collection of technologies and standards, highly mutable compared to sea or space, and conceptually more complex with the asymmetrical power yielded by such groups as the IETF and criminal groups distributing ransomware. To study these topics, we argue that the technology vendors, non-governmental organizations, and intermediaries like cloud computing providers need to be made a more explicit topic of study. To this end, we have explicitly captured this dimension in the questions above.

Change over Time

How might these dynamics change over time? For example, will additional automation and AI drive transformational changes in defense and offense (as the radio, airplane, and tank did)? Or will they lead to more incremental changes (such as the switch from fourth- to fifth-generation fighters)?

Further, the rapidly evolving nature of the technology and introduction of machine learning and AI will require continued focus on these questions as it is unclear how the dynamics will change across the tactical, operational and strategic dimensions.

The tactical and operational dynamics of cyber conflict remain under-studied and broadly misunderstood by much of mainstream academia. Dismissing these topics as too grounded in technological minutia is a mistake for the social sciences. There is ample ground here for work by younger faculty and graduate students. Cyberspace is man-made and conflict over its boundaries or to kinetic effects through it must take that into consideration.

Acknowledgements

The authors would like to thank JD and the Cyber Conflict Documentation Project for their support as well as Jay Healey and Karl Grindal for comments and feedback. Lastly, thank you to all of the participants of the 2017 State of the Field workshop on Tactical and Operational Dynamics of Cyber Conflict for their participation, input, and insight.

About the Authors

Trey Herr, Ph.D., is a postdoctoral fellow with the Belfer Center's Cyber Security Project at the Harvard Kennedy School.

Roberta Stempfley is the Director of the CERT Division at the Carnegie Mellon University Software Engineering Institute.

End Notes

1. www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=1242&context=fclj & https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2032124
2. Murat Balci et. al — "Defining Military Levels for Cyber Warfare by Using Components of Strategy/ Ends, Ways, and Means", *21st ICCRTS — C2 in a Complex Connected Battlespace*, www.researchgate.net/publication/307923231_Defining_Military_Levels_for_Cyber_Warfare_by_Using_Components_of_Strategy_Ends_Ways_and_Means; Trey Herr and Drew Herrick — "Understanding Military Cyber Operations", *Cyber Insecurity — Navigating the Next Information Age*, https://books.google.com/books?id=NAp7DQAAQBAJ&pg=PA13&source=gbs_toc_r&cad=3#v=onepage&q=herrick&f=false
3. Rebecca Slayton — "What is the Cyber Offense-Defense Balance? Conceptions, Causes and Assessment", International Security, www.mitpressjournals.org/doi/abs/10.1162/ISEC_a_00267
4. Timothy Thomas. Russia's Information Warfare Strategy: Can the Nation Cope in Future Conflicts. The Journal of Slavic Military Studies. Vol 27 No 1. March 2014; Timothy Thomas. Russian IW and an Analysis of Dr. Igor Nikolaevich Panarin. InfowarCon. Nashville, Tennessee. April 2015; Sergei A. Medvedev. Offense-Defense Theory Analysis Of Russian Cyber Capability. Naval Postgraduate School. March 2015; JD Work. Russian cyber operations in the current strategic landscape. Cambridge Intelligence Seminar. May 2017; Ammilee A. Oliva. China: Paper Tiger in Cyberspace. School of Advanced Military Studies, United States Army Command and General Staff College. March 2012.
5. Milton Mueller et. al — "Internet Security and Networked Governance in International Relations", *International Studies Review*, http://onlinelibrary.wiley.com/doi/10.1111/misr.12024/abstract
6. Panayotis a. Yannakogeorgos. Strategies for Resolving the Cyber Attribution Challenge. Air Force Research Institute. May 2013; Wylie McDade. Attribution, Delayed Attribution and Covert Cyber-Attack. Naval Postgraduate School. June 2014; Eric F. Mejia . Act and Actor Attribution in Cyberspace: A Proposed Analytic Framework. Strategic Studies Quarterly. January 2014; Jeff Wozniak and Samuel Liles. Political and Technical Roadblocks to Cyber Attack Attribution. IO Journal. Vol 1 Issue 1. April 2009; Brian Bartholomew & Juan Andres Guerrero-Saade. Wave Your False Flags! Deception Tactics Muddying Attribution in Targeted Attacks. Virus Bulletin Conference. October 2016 ; Don Cohen & K. Narayanaswamy. Survey/Analysis of Levels I, II,, and III Attack Attribution Techniques. ARDA. December 2004.
7. Thomas Rid and Ben Buchanan — "Attributing Cyber Attacks", *Journal of Strategic Studies*, www.tandfonline.com/doi/abs/10.1080/01402390.2014.977382
8. Gary D. Brown — "Spying and Fighting in Cyberspace: What is Which?" *Journal of National Security Law and Policy*, http://jnslp.com/2016/03/29/spying-fighting-cyberspace/; Aaron F. Brantly. Aesop's wolves: the deceptive appearance of espionage and attacks in cyberspace. Intelligence and National Security. September 2015; Ramberto A. Torruella, Jr. Determining Hostile Intent in Cyberspace. Joint Forces Quarterly. 4th Quarter 2014.
9. Office of Director of National Intelligence. Public - Private Analytic Exchange Program. Cyber Attribution Using Unclassified Data. September 2016. Unclassified.
10. Max Smeets — "Transitory Nature of Cyber Weapons", *Journal of Strategic Studies*, www.tandfonline.com/doi/abs/10.1080/01402390.2017.1288107
11. iSIGHT Partners. Rapid Botnet Acquisition Within the Russian Underground Marketplace. February 2010.
12. Dorothy Denning. Barriers to Entry: Are They Lower for Cyber Warfare. IO Journal. Vol 1 Issue 1. April 2009; Christos Siaterlis and Béla Genge. Cyber- Physical Testbeds. Communications of the Acm. Vol 57 No 6. JUNE 2014; Cormac Herley. The Plight of the Targeted Attacker in a World of Scale. Workshop on the Economics of Information Security. June 2010; Giancarlo Pellegrino, Christian Rossow, Fabrice J. Ryba, Thomas C. Schmidt,

Matthias Wählisch. Cashing Out the Great Cannon? On Browser-Based DDoS Attacks and Economics. USENIX Workshop on Offensive Technologies. August 2015.

13. Trey Herr — "Malware Counter-Proliferation and the Wassenaar Arrangement", *Proceedings of the 8th International Conference on Cyber Conflict*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2711070; Neil Robinson, Agnieszka Walczak, Sophie-Charlotte Brune, Alain Esterle, Pablo Rodriguez. Stocktaking study of military cyber defence capabilities in the European Union. RAND. Unclassified (summary); George Danezis and Bettina, The Economics of Mass Surveillance and the Questionable Value of Anonymous Communications. Workshop on the Economics of Information Security. June 2006.

14. Christopher W. Weimer. Forecasting Effects of Influence Operations: A Generative Social Science Methodology. Air University. March 2012.

15. Paul D. Williams. USAF Cyber Capability Development. Air Command and Staff College Air University. April 2009; Paul H. Orth. Measuring the Operational Readiness of an Air Force Network Warfare Squadron. Air University. June 2008.

16. Kristen Dennesen. Hide and Seek: How Threat Actors Respond in the Face of Public Exposure. SANS Cyber Threat Intelligence Summit. Alexandria, Virginia. 3-4 February 2016.

17. Lillian Ablon, Martin C. Libicki, Andrea A. Golay. Markets for Cybercrime Tools and Stolen Data. RAND. 2014; Rainer Böhme. Vulnerability markets — What is the economic value of a zero-day exploit. 22C3. Berlin, Germany. 2005; Michael Sutton and Frank Nagle. Emerging Economic Models for Vulnerability Research. Workshop on the Economics of Information Security. June 2006; Charlie Miller. The Legitimate Vulnerability Market: Inside the Secretive World of 0-day Exploit Sales. Workshop on the Economics of Information Security. Carnegie Mellon University. 7-8 June 2007; Stefan Frei,; Francisco Artes. International Vulnerability Purchase Program. NSS Labs. December 2013; Thomas Maillart, Mingyi Zhao, Jens Grossklags, and John Chuang. Given Enough Eyeballs, All Bugs Are Shallow? Revisiting Eric Raymond with Bug Bounty Programs. Workshop on the Economics of Information Security. UC Berkeley School of Law. Berkeley, CA. 13-14 June 2016; Art Manion. A Survey of Vulnerability Markets. FIRST Conference. Boston, MA. 26 June 2014; Abdullah M. Algarni, Yashwant K. Malaiya. Software Vulnerability Markets: Discoverers and Buyers. International Journal of Computer, Information Science and Engineering Vol:8 No:3, 2014; Sam Ransbotham, Sabyaschi Mitra, and Jon Ramsey. Are Markets for Vulnerabilities Effective? MIS Quarterly. Vol 36 Issue 1. 2012; Andy Ozment. Bug Auctions: Vulnerability Markets Reconsidered. Workshop on the Economics of Information Security. University of Minnesota. Minneapolis, MN. 13-14 May 2004; Trey Herr, Bruce Schneier, and Christopher Morris, "Taking Stock: Estimating Vulnerability Rediscovery" *Belfer Cyber Security Project*, www.belfercenter.org/publication/taking-stock-estimating-vulnerability-rediscovery

18. Ashish Arora, Ramayya Krishnan, Anand Nandkumar, Rahul Telang, and Yubao Yang, Impact of Vulnerability Disclosure and Patch Availability — An Empirical Analysis. Workshop on the Economics of Information Security. May 2004; Rainer Böhme. A Comparison of Market Approaches to Software Vulnerability Disclosure. In: Müller G. (eds) Emerging Trends in Information and Communication Security. Lecture Notes in Computer Science, vol 3995. Springer, Berlin, Heidelberg. 2006. Lillian Ablon and Andy Bogart, "Zero Days, Thousands of Nights" (Santa Monica, CA: The RAND Corporation, 2017); Trey Herr and Bruce Schneier, "Taking Stock: Estimating Vulnerability Rediscovery," Cyber Security Project Paper (Cambirdge, MA: Belfer Center, Harvard Kennedy School, May 2017),

19. Trey Herr, "PrEP: A Framework for Malware and Cyber Weapons", *Journal of Information Warfare*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2343798; Trey Herr and Amy Armbrust, "Milware: Identification and Implications of State Authored Malicious Software" *New Security Paradigms Workshop*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2569845; Anita DíAmico, Laurin Buchanan, John Goodall1 and Paul Walczak. Mission Impact of Cyber Events: Scenarios and Ontology to Express the Relationships Between Cyber Assets, Missions and Users. Proceedings of the Conference on Information Warfare & Security. 2010; Vitaly Tsygichko: Cyber Weapons as a new means of Combat. Classification of Cyber Weapons. Cyberwar, Netwar and Revolution in Military Affairs. International School on Disarmament and Research on Conflicts. Trento. 2002; Dale Peterson. Offensive Cyber Weapons: Construction, Development, and Employment. Journal of Strategic Studies. February 2013; James Morris-King, Hasan Cam. Ecology-inspired cyber risk model for propagation of vulnerability exploitation in tactical edge. Military Communications Conference MILCOM 2015. 26-28 October 2015.

20. Bellovin, Steven M., Susan Landau, and Herbert S. Lin — "Limiting the Undesired Impact of Cyber Weapons: Technical Requirements and Policy Implications." *Journal of Cybersecurity*, https://papers.ssrn.com/abstract=2809463; Kehler, Robert, Herbert S. Lin, and Michael Sulmeyer — "Rules of Engagement for Cyberspace Operations", Journal of Cybersecurity, https://academic.oup.com/cybersecurity/article/doi/10.1093/cybsec/tyx003/3058505/Rules-of-engagement-for-cyberspace-operations-a; Robert Fanelli and Gregory Conti. A Methodology for Cyber Operations Targeting and Control of Collateral Damage in the Context of Lawful Armed Conflict. in C. Czosseck, R. Ottis, K. Ziolkowski (Eds.) 24th International Conference on Cyber Conflict. NATO CCD COE. 2012

21. Gregory J. Kula. Assessing the Effects of Computer Network and Electronic Attack. Naval War College. May 2009; John Tokar. Assessing Operations: MOP and MOE Development. IO Journal. Vol 2 Issue 3. August 2010. Carrie Gray and Edwin Howard. IO MOE Development and Collection: A Paradigm Shift. IOSphere. Spring 2005; Shirazi, Reza. "Botnet takedown initiatives: A taxonomy and performance model." Technology Innovation Management Review. Vol 5 Issue 1. 2015;

Yacin Nadji, Manos Antonakakis, Roberto Perdisci, David Dagon, Wenke Lee. Beheading hydras: performing effective botnet takedowns. Proceedings of the ACM SIGSAC conference on Computer & communications security. 2013; Hongxu Yin, Rui Xiao, Fenfei Lv. Analysis of Causes and Actual Events on Electric Power Infrastructure Impacted by Cyber Attack. Journal of Power and Energy Engineering. 2015; Dave MacEslin. Methodology for Determining Electronic Warfare Joint Munitions Effectiveness Manual. IOSPhere. Spring 2006; L. Scott Johnson and Toni Whyte. Lessons to be Learned from a Recent Network Infrastructure Attack. IO Journal. Vol 1 Issue 2. September 2009; Larry W. Fortson, Jr. TOWARDS The Development of a Defensive Cyber Damage and Mission Impact Methodology. Air University. March 2007; Richard A. Martino. Leveraging Traditional Battle Damage Assessment Procedures to Measure Effects from a Computer Network Attack. Air Force Institute of Technology. June 2011; Daniel BILAR. On nth Order Attacks. NATO CCD COE. 2009; Deborah Bodeau, Richard Graubart. Characterizing Effects on the Cyber Adversary: A Vocabulary for Analysis and Assessment. Mitre. Technical Report 130432. November 2013; Christian Rossow, Dennis Andriesse, Tillmann Werner, Brett Stone-Gross, Daniel Plohmann, Christian J. Dietrich, Herbert Bos. SoK: P2PWNED — Modeling and Evaluating the Resilience of Peer-to-Peer Botnets. IEEE Symposium on Security and Privacy. 2013; Scott Musman, Aaron Temin, Mike Tanner, Dick Fox and Brian Pridemore. Evaluating the Impact of Cyber Attacks on Missions. Proceedings of the Conference on Information Warfare & Security. 2010.

22. Lin, Herbert S., Kenneth W. Dam, and William A. Owens, editors — *Technology, Policy, Law, and Ethics Regarding US Acquisition and Use of Cyberattack Capabilities*, www.nap.edu/catalog/12651/technology-policy-law-and-ethics-regarding-us-acquisition-and-use-of-cyberattack-capabilities; Colin F. Jackson — "Information Is Not a Weapons System", *Journal of Strategic Studies*, www.tandfonline.com/doi/abs/10.1080/01402390.2016.1139496; National Security Agency. Computer Virus Infections: Is NSA Vulnerable? Cryptologic Quarterly. Declassified in February 2008; Robert Majoris. Cyber Warfare as an Operational Fire. Naval War College. March 2010; Exploitation of Blue Team SATCOM and MILSAT Assets for red Team Covert Exploitation and Back-Channel Communications. David Rohret and Jonathan Holston. Proceedings of the Conference on Information Warfare & Security. 2010; Christopher Bronk and Eneken Tikk-Ringas. The Cyber Attack on Saudi Aramco. Survival. Vol 55 No 2. April-May 2013; David A. Rickards. No Air: Cyber Dependency and the Doctrine Gap. Naval War Collge. March 2010; Marc Romanych. Objectives in the Information Environment. IOSphere. Winter 2006. Per Kjellns. The Role of Computer Network Exploration (Active Sigint) in Information Warfare. Military Technical Section of The Royal Swedish Academy of War Sciences. 8 May 2001.

23. Michael Klipstein and Michael Senft, "Cyber Support to Corps and Below: Digital Panacea or Pandora's Box?" *Small Wars Journal*, http://smallwarsjournal.com/jrnl/art/cyber-support-to-corps-and-below-digital-panacea-or-pandora%E2%80%99s-box; Christopher R. Eidman, Gregory Scott Green. Unconventional cyber warfare: cyber opportunities in unconventional warfare. Naval Postgraduate School. June 2014; Steven Zielechowski. The Commanding Officer's Perspective on Protecting Shipboard IT Networks. Naval Postgraduate School. September 2014

24. Drew Herrick and Trey Herr — "Combatting Complexity", *Working Paper*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2845709; Max Smeets — "Organisational Integration of Offensive Cyber Capabilities: A Primer on the Benefits and Risks", *Proceedings of the 9th International Conference on Cyber Conflict*, https://ccdcoe.org/sites/default/files/multimedia/pdf/Art%2002%20Organisational%20Integration%20of%20Offensive%20Cyber%20Capabilities.pdf;

25. Booz Allen Hamilton. When the Lights Went Out: A Comprehensive Review of the 2015 Attacks on Ukrainian Critical Infrastructure. September 2016; Dan Fayutkin. Russian-Chechen Information Warfare 1994-2006. RUSI Journal. Vol 151 No 5. October 2006.

26. Martin Libicki — "The Convergence of Information Warfare", *Strategic Studies Quarterly*, www.airuniversity.af.mil/Portals/10/SSQ/documents/Volume-11_Issue-1/Libicki.pdf; Drew Herrick — "The social side of 'cyber power'? Social media and cyber operations", *Proceedings of the 8th International Conference on Cyber Conflict*, https://ccdcoe.org/sites/default/files/multimedia/pdf/Art%2007%20The%20Social%20Side%20of%20'Cyber%20Power'.%20Social%20Media%20and%20Cyber%20Operations.pdf; Information Operations by the British in the War of 1812 During the Maryland Campaign. Defense Intelligence Journal. Vol 12 No 2. 2003; Jonathan Reed Winkler. Information Warfare in World War I. The Journal of Military History. Vol 73 No 3. July 2009; David Acosta. The Makara of Hizballah: Deception in the 2006 Summer War. Naval Postgraduate School. June 2007. Carl Anthony Wege. Hezbollah's Communication System: A Most Important Weapon. International Journal of Intelligence and CounterIntelligence. Vol 27 No 2. March 2014

27. Martin Libicki — "Second Acts in Cyberspace", *Journal of Cybersecurity*, https://academic.oup.com/cybersecurity/article/3/1/29/3056957/Second-acts-in-cyberspace

28. Lin, Herbert S. and Taylor Grossman. "The Practical Impact of Classification Regarding Offensive Cyber Operations," in eds. Richard Harrison and Trey Herr, *Cyber Insecurity: Navigating the Perils of the Next Information Age*

29. Brian Thompson and Richard Harang — "Identifying Key Cyber Physical Terrain", *International Workshop on Security and Privacy Analytics (IWSPA)*, https://arxiv.org/abs/1701.07331; J. W. Mickens and Brian Noble — "Analytical Models for Epidemics in Mobile Networks", *Third IEEE International Conference on Wireless and Mobile Computing, Networking, and Communications*, https://experts.umich.edu/en/publications/analytical-models-for-

epidemics-in-mobile-networks; Douglas H. Dearth. Applying Maneuver Warfare to Infrastructure Protection. InfoWarCon. 1999; L. M. Marvel et. al "A Framework to Evaluate Cyber Agility", *Military Communications Conference*, MILCOM, http://ieeexplore.ieee.org/document/7357414/

30. Amos C. Fox and Andrew J. Rossow — "Making Sense of Russian Hybrid Warfare: A Brief Assessment of the Russo–Ukrainian War", *The Institute of Land Warfare*, www.ausa.org/publications/making-sense-russian-hybrid-warfare-brief-assessment-russo%E2%80%93ukrainian-war; Keir Giles — *Handbook of Russian Information Warfare*, NATO Defense College, www.css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/resources/docs/NDC%20fm_9.pdf

31. See Executive Order 13636, Section 9

32. Laura DeNardis — *The Global War for Internet Governance*, Yale University Press, https://books.google.com/books?id=jfxfAgAAQBAJ&pg=PR7&source=gbs_selected_pages&cad=3#v=onepage&q&f=false

33. Jason Healey. Threat and Warning for the Financial Sector. InfoWarCon. 2002; JD Work. Understanding information sharing in cyber intelligence communities of practice: Evidence from collaborative analytic exchange. Intelligence and the Cyber Environment. Brunel University, Uxbridge. November 2014; Payton A. Flynn, Sr. Cybersecurity: Utilizing Fusion Centers to Protect State, Local, Tribal, and Territorial Entities Against Cyber Threats. Naval Postgraduate School. September 2016; Erick Mandt. On integrating cyber intelligence analysis and active cyber defense operations. Uitica College. 2015; Kenneth A. Minihan. Intelligence and Information Systems Security: Partners in Defensive Information Warfare. Defense Intelligence Journal. Vol 5 No 1. 1996; Why Them? Extracting Intelligence about Target Selection from Zeus Financial Malware. Workshop on the Economics of Information Security. June 2014; Alex Pinto. Data-Driven Threat Intelligence: Metrics on Indicator Dissemination and Sharing. SANS Cyber Threat Intelligence Summit. Alexandria, Virginia. 3-4 February 2016.

34. Chesney, Robert — "Military-Intelligence Convergence and the Law of the Title 10/ Title 50 Debate." *Journal of National Security Law and Policy*, http://jnslp.com/wp-content/uploads/2012/01/Military-Intelligence-Convergence-and-the-Law-of-the-Title-10Title-50-Debate.pdf; Harry M Friberg. U.S. Cyber Command Support To Geographic Combatant Commands. U.S. Army War College. February 2011; Joseph E. Sisson. Fleet Cyber Command/Tenth Fleet: Enabling Cyber Unity of Effort. March 2010; Richard Mesic, Myron Hura, Martin C. Libicki, Anthony M. Packard, Lynn M. Scott. Air Force Cyber Command (Provisional) Decision Support. RAND. 2010.

35. S. W. Stone — "Agility in decision-making for cyberspace operations," *Military Communications Conference MILCOM*, http://ieeexplore.ieee.org/document/7795294/; S.W. Stone — "Factors related to agility in allocating decision-making rights for cyberspace operations" *Diss. Robert Morris University*, http://static1.squarespace.com/static/53bad224e4b013a11d687e40/t/54da5be5e4b0e9d26e577151/1423596517506/096.pdf; Daryl L. Caudle. Decision-Making Uncertainty and the Use of Force in Cyberspace: A Phenomenological Study of Military Officers. Office of the Chairman of the Joint Chiefs of Staff. Strategic Plans and Policy (J5). October 2010; David M. Franklin. U.S. Command Relationships in the Conduct of Cyber Warfare: Establishment, Exercise, and Institutionalization of Cyber Coordinating Authority. Naval War College. March 2010; Joseph H. Scherrer, William C Grund. A Cyberspace Command and Control Model. Air War College. August 2009; Bradley L. Pybmu. Application of US Special Operations Command Model to Department of Defense Cyberspace Force. United States Marine Corps Command and Staff College, Marine Corps University. 2009; Norman R. Howes, Michael Mezzino, John Sarkesain. On Cyber Warfare Command and Control Systems. Department of Defense, Command and Control Research Program. Command and Control Research and Technology Symposium. 2004; Norman R. Howes, et. al. Cyber Warfare Command and Control System Users Manual. Institute for Defense Analysis. July 2003; Russell J. Caldwell. Information Operations (IO) Organizational Design and Procedures. Naval Postgraduate School. November 2004.

36. Long, Austin - "A Cyber SIOP? Operational Considerations for Strategic Offensive Cyber Planning." *Journal of Cybersecurity*, www.researchgate.net/publication/313933131_A_cyber_SIOP_Operational_considerations_for_strategic_offensive_cyber_planning; Fredrick Okello, Richard Ayres, Patrice Bullock, Brahim Erhili, Bruce Harding, Allan Perdigao. Information Warfare: Planning the Campaign. Air Command & Staff College. April 1996; Steven J. Smart. Joint Targeting in Cyberspace. Air and Space Power Journal. Winter 2011.

37. iSIGHT Partners. Potential 'Dead Hand' C&C Architecture Suggested by Adversary Adaptation Following Failed Botnet Takedown Attempt. February 2010; JD Work. Autonomy & Conflict Management In Offensive & Defensive Cyber Engagement. InfowarCON. Nashville, Tennessee. 5-7 April 2016.

38. Peter R. Stephenson. Towards Improving Attribution Confidence in Cyber Attacks. Journal of Cyber Conflict Studies. Vol 1 Issue 1. September 2006; Rudolph "Reb" Butler, Dick Deckro, Jeff Weir. Using Decision Analysis to Increase Commanders Confidence for Employment of Computer Network Operations. IOSphere. Fall 2005; Lou Anne DeMattei. Developing A Strategic Warning Capability For Information Defense. Defense Intelligence Journal. Vol 7 No 2. 1998; D.M. Rock. Cyber Attack: The Department of Defense's Inability to Provide Cyber Indications and Warning. Marine Corps Command and Staff Coll. 7 February 2006; Marco Roscini. Evidentiary Issues in International Disputes Related to State Responsibility for Cyber Operations. Texas International Law Journal Volume 50, Symposium Issue 2.

39. Frans Mulschlegel, Jim Christy. Corporate Vigilantism: Striking Back. InfoWarCon. 1999; Winn Schwartau, W. Hutchinson. Corporate Vigilantism and the Hostile Perimeter. InfoWarCon. 1999; Stewart Baker, Orin Kerr, and Eugene Volokh. "The Hackback Debate". Steptoe & Johnson LLP. 22 November 2012.

[40] Christopher Paul, Isaac R. Porche III, Elliot Axelband. The Other Quiet Professionals. Lessons for Future Cyber Forces from the Evolution of Special Forces. RAND. 2014; Timothy Franz. The Cyber Warfare Professional. Air & Space Power Journal. Summer 2011; Zhang Jun-qi,KE Hong-fa,ZHU Ji-luck. Discussion on Core Competencies and Construction Elements of Cyberwarfare Forces. Journal of Ordnance Equipment Engineering. July 2015; Joel Hill. Transforming Intelligence Education to Support Information Operations. Defense Intelligence Journal. Vol 12 No 1. 2003; Lynn M. Scott, Raymond E. Conley, Richard Mesic, Edward OíConnell, Darren D. Medlin. Human Capital Management for the USAF Cyber Force. RAND. 2010.

The Cyber Conflict Studies Association (CCSA) promotes and leads international intellectual development efforts to advance the field of cyber conflict research. These activities include workshops that bring together professionals from industry, academia and government to discuss strategic issues surrounding cyber conflict and the publication of insightful research articles and position papers and books. CCSA also plays an important role in our national cyber-readiness strategy, serving as a resource for national security decision-makers and helping to frame and promote national cyber conflict policy. CCSA brings together the best and the brightest individuals in the field of cyber conflict study to further the goals of national security and the field of cyber.

This work was supported in part by the Minerva Research Initiative. The Minerva Research Initiative, administered jointly by the Office of Basic Research and the Office of Policy at the U.S. Department of Defense, supports social science research aimed at improving our basic understanding of security.

Funded in part by the Office Naval Research under the OSD Minerva program, grant number: N00014-17-1-2423.

COLUMBIA | SIPA
School of International and Public Affairs

CCSA
Cyber Conflict Studies Association

International Security and the Strategic Dynamics of Cyber Conflict

AUTHORS: Melissa K. Griffith and Adam Segal
SERIES EDITOR: Justin Key Canfil EXECUTIVE EDITOR: Jason Healey

Introduction

What is the state of the field of cyber conflict within the fields of international relations and international security? Which questions have been answered and which remain unexplored? Where should those hoping to push forward discussion around and promote an understanding of the strategic dynamics of cyberspace focus their intellectual energy and efforts? What are the questions scholars need to be asking now?

These were the driving concerns of the International Security and Strategic Dynamics panel at the 2017 State of the Field of Cyber Conflict Conference,[1] where researchers, policy analysts, and practitioners came together to discuss the emerging scholarship on cybersecurity. The goal of this particular panel, as summarized in this White Paper, was to capture the evolution of the field since the 2016 State of the Field Conference, to provide an overview of existing scholarship focusing on the strategic dynamics of cyberspace, and to identify where more rigorous research could expand the frontiers of the field.

About the State of the Field Series

This article is part of the 2017 Cyber Conflict State of the Field (SOTF) paper series, under the auspices of the Cyber Conflict Studies Association and Columbia University's School of International and Public Affairs.

The conference, held annually since 2016, brings together experts from various academic disciplines, including political science, law, economics, and policy research, to define key questions and map the research frontier in the emerging field of cyber conflict studies. The conference is cumulative: each year builds upon past discussions. As a result, discussions have necessarily matured at different rates as new topics are added.

The papers in this series are meant to capture the findings of the 2017 conference. Together, the papers represent the conference attendees' understanding of the present state of the field in the academic study of cyber conflict.

The 2017 panel drew heavily on work completed by Ryan C. Maness and Adam Segal for the inaugural 2016 State of the Field Conference.[2] To assess the state of the field in 2016, Maness and Segal compiled traditional international security literatures and then looked for corresponding research emerging on cyberspace within those categories. The categories of interest were drawn from the foundational sub-literatures within international security, including deterrence, the offense-defense balance, security dilemmas, foreign policy doctrines, arms races, and norms or taboos. By mapping existing work onto these traditional security sub-literatures, Maness and Segal were able to identify emerging consensus and persisting disagreements and gaps. The questions asked in 2016 then became "what work has been done around deterrence," "what are the limitations of this approach," and "what gaps remain in the study of deterrence," rather than the far broader "what are the strategic dynamics of cyberspace?"

The 2017 panel first sought to update the research listed under these various categories to include work that had emerged over the preceding year and, second, to identify persisting gaps and potential limitations to this approach. In addition to newly published work, we also considered presentations at the 2017 International Studies Association (ISA) Conference as markers of unpublished work in progress by scholars.

Notably absent from this White Paper's analysis, however, are the bodies of work addressing intelligence, economics, operational and tactical dynamics, the history of cyber conflict, and the legal and ethical issues embedded within cyber conflict. Each of these topics, given its importance to the field of cyber conflict and security research, comprised its own panel and subsequent White Paper. Similarly, work focusing on cybercrime, although a subset of the cyber security debate, remains outside the scope of this project. Given the conceptual distinctions drawn between White Papers and between cyber conflict and other forms of cyber incident, this White Paper specifically focuses on strategic dynamics by mapping out the relationship between the strategic study of cyber conflict and the foundational sub-literatures within international security. It leaves the analysis and review of these related bodies of work to others.

Many of our findings are consistent with the trends identified in 2016. Articles continue to be heavily focused on larger states (e.g. the U.S., the U.K., China, and Russia). In terms of theoretical frameworks, the categories of deterrence and offense-defense balance represent the largest bodies of work. Certain sub-fields that receive significant attention in other security domains, such as international cooperation, remain systematically understudied. Finally, an overarching refrain during the plenary discussion was a continued questioning of whether international security and its sub-literatures provide useful analytical leverage for studying the strategic dynamics of cyber security, whether important dynamics remain uncaptured, and whether more progress could be made through greater reliance on analytical tools from other disciplines. This debate was reflective of the fundamental question of whether cyberspace has created a revolution for security politics or whether some aspects of previous security politics remain relevant for the study of cyber security.

Given these findings, we conclude that progress in the field remains slow and evolutionary. Within the field of international security, several key gaps remain. Moreover, as in other political science fields, there is a sharp divide between practitioners/policy analysts and academics on the utility and applicability of units of analysis, theories, and methodologies drawn from international security studies.

This White Paper will proceed in three parts. First, we will discuss the major takeaways from the 2017 panel discussion. This section will focus on the overall structure of the field and emerging oversaturation and gaps within it. Second, we will outline the current state of the field. This section will break existing work into sub-categories, or sub-literatures, and tie those works directly to questions currently being asked while highlighting unanswered questions that have persisted within and between categories. While this section of the paper does not act as a comprehensive literature review of work on the strategic dynamics of cybersecurity, we hope it will be a useful foundational reference for those entering the field. Throughout, we will strive to identify new work rather than simply recapping the work already covered in the 2016 State of the Field Report. Third, we will provide a short summary of key observations from our preparation and the subsequent panel discussion, identify limitations of the approach used to develop the panel and this paper, and provide a few recommendations for moving the field forward.

Major Takeaways from 2017

During our preparation for this panel and in the subsequent panel and plenary discussion, four key takeaways emerged: tensions remain between academics and policy practitioners regarding the utility of international security studies to cyber security research; the field continues to over-study some topics and ignore others; there is a high degree of overlap between the panels on tactical/operational dynamics and those on strategic dynamics; and there is a need to better link the destructive or disruptive effects of cyberattacks with the strategic goals states pursue in cyberspace.

First, the utility and applicability of international security studies and international relations more broadly remain contentious, especially between practitioners/policy analysts and academics. The main concerns here point to diverging intellectual interests. These fields bring with them assumptions about which variables are most influential and which dynamics are most important in the study of conflict. For example, international security and international relations focus heavily on states. Yet the state is only one of many actors in cyber conflict, and the private sector in particular plays a large role. This discussion left us once again with a question voiced last year: in what ways might traditional security discussions and theoretical frameworks limit discussion of cyberspace and its strategic dynamics?

Second, the same areas continue to be studied. Gaps identified in 2016 largely persist. Empirical work and case studies focus on a select few countries.

Given these dominant topics and persisting gaps, why has the field developed in the manner in which it has? Why, for example, do we see a heavy, persistent focus on deterrence and the offense-defense balance and, simultaneously, a hesitance to address cybersecurity using other sub-literatures within international security?

Is this merely a question of timing? After exploring preliminary questions through the dominant sub-literatures, will scholars move on to the additional sub-literatures? This would almost seem to imply that there is a natural progression for any topic located within international security. Certain dynamics may be studied before others because the concepts addressed are building blocks for later sub-literatures, because the outcomes inherent to some sub-literatures occur historically later than outcomes studied in others, or because the prioritized sub-literatures are dominant in the broader field of international security.

Perhaps it is a question of applicability, especially for policymakers and practitioners. However, the applicability of deterrence and offense-defense balance to this new threat space is highly contested. Much of the debate after our panel focused on moving away from deterrence models because they were not applicable.

The core observation remains: the evolution of the field is largely consistent. The question as to why remains unanswered and underexplored.

Third, clearly delineating between tactical, operational, and strategic levels of war in cyber conflict is easier in theory than in practice. While conceptually the State of the Field Conference separated tactics and operations from strategic dynamics, in the subsequent discussions for both sessions there was significant overlap.

This overlap may have been due in part to an open forum format conducive to the blurring of boundaries or the way some topics—such as work focusing on how standard operating procedures shape strategic cyber planning—naturally bridge buckets. But the more interesting explanation is the particular challenge cyber conflict poses to the levels of war categorization. The issue here is not the more general point that the dynamics in one level affect dynamics in another but rather that in cyberspace, dynamics at the tactical level increasingly reverberate at the strategic level. In cyber conflict, including cyber-facilitated information operations, "analysts face the challenge of the strategic corporal in a more dramatic fashion: tactical behaviors can rapidly have strategic effects."[3] Given these dynamics, questions of interest reside between levels of war rather than in discrete buckets more often here than in the study of conflict on air, land, and sea.

Why, then, should scholars keep a levels of war distinction? There are several advantages to utilizing this conceptual distinction for research in international security and the strategic dynamics of cyberspace. First, there are different dynamics at each of these levels worthy of study. Take the dynamic of speed as an example. At the tactical level, things occur at the speed of light with very little reaction time, but at the strategic level, campaigns are more prolonged. Second, the levels of war provide a useful frame for understanding particular outcomes. Take the cyberattacks on Estonia in 2007 as an example. Tactically, Estonia lost, with widespread outages in the face of massive DDOS attacks. Strategically, Estonia still moved the statue and has, over the last decade, used the attacks to establish itself as a leader in cybersecurity and international norms.

This third takeaway from the 2017 Strategic Dynamics panel points to the need for additional attention to be paid to the ways in which cyber conflict blurs traditional conceptual boundaries utilized in other security domains. When do levels of war remain distinct and productive categories? When do these categories limit inquiry into the dynamics of cybersecurity?

Fourth, this panel merged two previously distinct panels: international security and strategic dynamics. In addition to increasing the range of relevant literature and potential research questions, these overlapping fields do have two distinct framings. In some contexts, we are focusing on the destructive effects of conflict. In others, the outcome of interest is not destruction but rather the strategic effects. These are not one in the same. To grapple with the strategic dynamics of this space, we need to frame discussions around the strategic outcomes motivating and/or driving conflict. These focuses lead to two very different sets of questions: (1) what are the dynamics or core characteristics of cyber conflict, and (2) how do specific actors pursue strategic outcomes using cyber means?

2017 State of the Field: International Security and Strategic Dynamics of Cybersecurity

There are many different ways to conceptualize a "state of the field." When we ask what work is missing, what questions have been asked, and what questions need to be asked, the intended target audience should be at the forefront of any discussions. Different audiences require different information and often hold different research goals and desire different deliverables. These diverging preferences were on full display at the 2017 State of the Field Conference, which brought scholars, policy analysts, and practitioners together to grapple with a research agenda for cybersecurity.

What, then, do we mean when we identify gaps in the literature and posit from these gaps what the next wave of research questions should be? Is this research directed at policymakers tasked with developing cybersecurity policy, PhD students pursuing an academic career, or scholars contributing to their chosen field? While there is some overlap thematically between these three audiences, the exact research questions vary.

This panel, like that of 2016, focused on the second and third categories and compiled resources to that end. We recognize, however, that the state of the field—work completed and remaining gaps—would be organized differently for policymakers or policy analysts. Moving forward, research on the areas of overlap between the academic and policy communities may be the most fruitful.

Our 2017 review of the state of the field is organized around eight subtopics of cyber conflict and security within international security. These subtopics are not mutually exclusive, and many overlap. Indeed, some subtopics include broadly grouped work, a byproduct of the limited research available in these areas. As more research is pursued, these broader categories can and should be broken apart into their constituent parts. The subtopics are as follows:

1. The Structure of the Security Environment: Defining the Degree of Change
2. Deterrence, Dissuasion, and Attribution
3. Offense versus Defense, the Security Dilemma, and Escalation
4. Power and Influence
5. Foreign Policy and Doctrine
6. The Relationship between State and Non-State Actors
7. Norms and Norm Diffusion
8. International Institutions and Cooperation

We take each of these eight categories in turn, providing a short summary followed by relevant works and key questions. The conclusion of the section captures the questions that emerged in the subsequent panel and plenary discussions. Where applicable, subsequent panel discussion questions have been folded into their relevant bucket. After reviewing the subtopics, we draw attention to a range of concerns that do not fall as neatly into a single category.

It is worth noting that although we organize this White Paper thematically, there is an alternative chronological framing. Scholarship in the 1990s and early 2000s focused heavily on the Revolution in Military Affairs (RMA) and how information technology altered the ways in which wars were fought and conflict was conducted.[4] This scholarship largely gave way to research focusing on the broader implications of technological shifts for conflict and war in the international system.

Research from the early to mid 2000s focused heavily on the structural implications of cyberspace: whether potential cyber conflict represented an evolution or a revolution in the nature of warfare. By the late 2000s and early 2010s, academics had begun to move beyond structural implications and into research on the specific strategic dynamics of conflict, such as deterrence, compellence, offense-defense balance, cooperation, and norms development. This uptick in interest within international relations and international security will likely lead to additional phases of research that move into categories of inquiry so far overlooked.

The Structure of the Security Environment: Defining the Degree of Change

A debate from the 2016 conference that carried forward into 2017 is whether cyber conflict represents a revolution or an evolution of conflict.[5] If it is the latter, previous theoretical models more readily describe the phenomenon. If it is the former, then many, if not most, of our theoretical models contain assumptions and dynamics that are inappropriate to the study of cyber conflict and cyber security.

While it is easy to simplify this discussion into an "everything has changed" and "nothing has changed" caricature, it is more useful to ask: Which aspects have changed? By how much? What has remained the same? How similar? Given these changes to the structure of the security environment, what are the strengths and limitations of our current models?

Early scholarship examined the future of the Internet as a domain of conflict, worst- and best-case scenarios in cyberspace, and whether the structure of the "5th domain of warfare" radically departed from previous security domains.

In the 2016 report, this discussion was housed, in part, under "Opening Concerns with IR Categories" and "How Does Cyber Fit into IR." Surveyed work included Arquilla and Ronfeldt (1993), Clarke and Knake (2011), Choucri (2012), Junio (2013), Kello (2013), Lawson (2013), Rid (2013), and Lindsay and Kello (2014).[6] Additional works published within the last year include Kello (2017), Lindsay (2017), and Perkovich and Levite (2017).[7]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around the structure of the cyber security threat environment. We highlight two here:

1. How should we characterize the structure in relation to other security environments? What models best capture cyber security and conflict?

1a. How does the structure compare to other security environments?

The purpose here is to identify similarities and differences between a range of security environments and then to determine which of these similarities and differences most influence the structure and outcomes in this space. How does the structure compare across other warfighting domains (air, land, and sea)? Is it more or less escalatory? Does the central role of the private sector make cyberspace fundamentally different from other domains? Does the range of actors and their relative power differ? Is conflict less discrete here than in previous threat spaces?

1b. Which models can we draw on beyond those used for other warfighting domains? What are the limitations of state-centric kinetic force models?

There was strong consensus in the room over the clear limitations of international security models. In the panel discussion, several audience members raised concerns over the kinetic force assumptions built into much of international security studies. They pushed instead for a focus on other types of scholarship geared around non-security systems. Suggestions for bodies of work that might better capture the structure of this threat space included human security, public health,[8] and economics.

In addition, it would be useful to move beyond state-to-state behavior and toward systems behaviors. We know that non-state actors (individuals, corporations or firms, terrorist organizations, etc.) play a significant role in cyber conflict and security. There are international relations sub-literatures that focus on a wider range of actors. We should more readily draw on sub-literatures including terrorism, intrastate conflict, and organized crime in our work on cyber conflict. Furthermore, we could utilize system-based or network analysis-based approaches to mapping out the underlying structure of cyber conflict.

2. Examining structure is useful for developing research on cyber conflict, especially as

it relates to international security. But what is the nature of the relationship between the structure and the ways we are organizing within this domain?

One thread of the panel discussion focused on exploring the structure of cyberspace and cyber-facilitated conflict: the underlying dynamics that all actors in this space face. Two sub-questions animated this discussion. First, what organizations are emerging from the structure? Second, how do those organizations affect the structure itself? These two questions point to a desire to move away from unidirectional cause and effect thinking in this space. Rather, structure shapes organizations and organizations can in turn shape the underlying structure. This process is iterative. We need to move beyond conversations on structure alone and assumptions that it is independent of efforts to grapple with this new threat space.

Deterrence, Dissuasion, and Attribution

Consistent with the 2016 report, deterrence remains an area with comparatively significant coverage. However, given the perceived limitations of the deterrence model, it has been suggested that discussions should be rooted in the fundamental characteristics of cyberattacks and conflict; this is a non-kinetic space defined by continuous rather than discrete conflict, where attribution is difficult and non-state actors are increasingly important players. Addressing these characteristics will likely require us to look beyond deterrence and toward models of dissuasion, compellence, bargaining, and restraint.

In the 2016 report, surveyed work included Kugler (2009), Libicki (2009), Goodman (2010), the National Research Council (2010), Nye (2011), Cooper (2012), Valeriano and Maness (2014), and Gartzke and Lindsay (2015).[9] Other notable works published prior to 2017 include Harknett (1996), Clark and Knake (2010), Liff (2012), Tsagourias (2012), Gartzke (2013), Iasiello (2013), Schmitt and Vihul (2014), Lindsay (2015), Healey (2016), and Lupovici (2016).[10] Additional work published in this category within the last year includes Borghard and Lonergan (2017), Chen (2017), Davis et al. (2017), Edwards et al. (2017), Fischerkeller (2017), Harknett and Fischerheller (2017), Harknett and Nye (2017), Mandel (2017), Nye (2017), and Sharp (2017).[11] Papers and posters on this topic at ISA 2017 include Cunningham, Lupovici, and Wilner.[12]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around deterrence, dissuasion, coercion, and attribution. We highlight four here:

1. How can you deter a cyberattack?

This question was highlighted in the 2016 State of the Field Conference, persisted in the 2017 panel, and continues to animate policymakers and academia.

There is widespread agreement that cyber conflict raises specific challenges for classical models of deterrence. Classical deterrence relies on (1) a credible threat of the imposition of costs in retaliation (deterrence by punishment), and/or (2) the ability to deny strategic benefit (deterrence by denial) if an attack does occur. These mechanisms for deterrence face unique challenges in cyberspace for four broadly discussed reasons (see Libicki (2009), Iasiello (2013), and Rid and Buchanan (2015)).

First, attribution presents a unique challenge in cyberspace. Complexities include the time it may take to technically or politically attribute an attack to a specific actor; difficulties raised by false flags, plausible deniability, and proxy actors; and reliance in some instances on private actors for forensic attribution. Attribution can be more or less challenging depending on the type of cyberattack in question and the resources a state can bring to bear.

Second, reliance on cyberspace is asymmetric. Some states and non-state actors have smaller relative attack surfaces than others, limiting the potential scope and scale of retaliation in kind. In contrast with nuclear weaponry, to which all states are vulnerable, potential adversaries may not be equally vulnerable to cyberattacks.

Third, the difficulty of signaling cyber cost-imposing capabilities further complicates matters. Cyber capabilities are less visible than their kinetic counterparts and have limited life spans (i.e., once attacked, the target is made aware of a vulnerability and has an opportunity and incentive to address it).

Fourth, proportionality or retaliation requires proper categorization of an incident and tailoring of a proportional response. In the cyber realm, however, the purpose and scale of an attack is often ambiguous. An observable outcome could be a failed effort at a more major network breach, a warning shot, espionage, or an operational preparation of the environment (OPE) for future activity. On the other side of the coin, the effects from any given attack can be unpredictable and can far exceed the root cause. Taken together, these four challenges undermine the ability of states to credibly threaten the imposition of costs in retaliation and to signal their capability of denying the benefits gained from cyberattack.

While there is consensus over these major challenges to deterrence, there is little agreement in the literature about whether or not these challenges can be overcome. Several scholars argue that despite these limitations, deterrence is still possible in cyberspace. Gartzke and Lindsay (2015), for example, focus on cross-domain deterrence and argue that ways of imposing costs beyond punishment in kind can overcome concerns around asymmetry and signaling. Goodman (2010) contests that deterrence is harder in theory than in practice, while Nye (2017) draws attention to the array of deterrence mechanisms beyond punishment and denial: entanglement and norms/taboos. Others have proposed entirely new models of managing potential cyber aggression that move away from deterrence entirely. For example, Harknett and Fischerkeller (2017) argue that, given its limitations, deterrence is not a credible strategy for cyberspace, and that we should turn instead to a strategy of cyber persistence.

Ultimately, much of the work focusing on how states can deter cyberattacks or best manage those attacks that do occur remains largely theoretical. It has yet to delve deeply and systematically into empirical analysis focusing on instances of and strategies for deterrence in this space.

2. How critical is attribution to deterrence?

Early scholarship and opinion pieces on deterrence in cyber conflict focused on the central role of attribution and the ways in which it is slow or imprecise in this realm. Rid and Buchanan (2015) and Libicki's (2016) work rests here. Research questions in this vein begin with the role that attribution plays in deterrence models, then spiral out into the types of attribution that are possible within the context of cyber conflict. From there the question becomes, given these particular dynamics of attribution, is it possible to utilize traditional deterrence models? What are the limitations of deterrence in this context? Nye (2017), for example, argues that deterrence and dissuasion in cyberspace are comprised of four mechanisms and that only the first—threats of punishment—requires attribution.

Moreover, attribution is not merely a technical act. Edwards et al. (2017) highlight the differences between the strategic and technical components of attribution, as well as reasons why states might choose not to undertake it. In a similar vein, Healey's (2016) attribution scale as well as the work of Rid and Buchanan (2014) and Davis et al. (2017) provide foundations for attempts to move away from technical forensics around attribution and toward more political determinations.

3. How relevant is the nuclear deterrence model to cyber deterrence?

There was widespread agreement during this panel that the nuclear model does not provide useful leverage for deterrence in the context of cyberattacks. This past year, for example, Chen (2017) argued that we need to move away from the nuclear model of punishment and denial and toward a model focusing on engagement and surprise. The limitations of the nuclear model range from a focus on state actors to attribution to credible punishment. Many of these same limitations apply to deterrence models more generally. For example, Fischerkeller (2017) argued for an offensive component to cyber deterrence, and U.S. General James Cartwright referenced the importance of demonstrated offensive capabilities to deterring adversaries.[13]

In addition, Clark and Knake (2010) highlighted the limitations of the nuclear model for deterrence, given that it rests on large scale retaliation in kind and assumes total prevention of nuclear attacks. In cyber conflict, retaliation may not be in kind and both the attack and response may fall below thresholds of war. Moreover, in a domain of constant contact, total prevention is not the goal of any defense strategy. Rather, states should focus on preventing escalation beyond low-level activity and maintaining society-wide cyber resilience.

4. What are the persisting limitations of deterrence models? Given those limitations, what alternative models should we utilize?

These types of questions represent the most vibrant paths forward for deterrence research. Deterrence, or at least the punishment model, may be a dead end for cyber conflict studies, but broader questions regarding conflict mediation and prevention remain central and productive areas of inquiry.

Should we use kinetic literatures to analyze the strategic dynamics of a non-kinetic space? How applicable are theories of deterrence to a threat space defined by constant, rather than discrete, conflict? Given these defining characteristics, what alternative models might apply to conflict prevention in cyberspace?

Four potential alternative approaches were suggested during the panel discussion: dissuasion, compellence, bargaining, and restraint. Notably, Gartzke and Lindsay (2015), Valeriano and Maness (2015), Harknett and Fischerkeller (2017), and Nye (2017) all provide alternative models to deterrence in their work. During the panel discussion, it was also suggested that we should look to scholarship like Fearon's work on bargaining and cooperation for inspiration.[14]

In conclusion, it would be productive to move away from literatures focusing on conflict as either on or off, and into conflict management and prevention strategies resting on understandings of conflict as continuous.

Offense Versus Defense, the Security Dilemma, and Escalation

Consistent with the 2016 report, offense versus defense and the security dilemma continue to garner significant coverage. The core assumption of work cataloging the offensive and defensive characteristics of cyberspace is that these characteristics drive the likelihood, intensity, or cost of conflict. Existing work has been motivated by two broad questions: (1) is cyberspace offense or defense dominant, and (2) what factors determine whether cyberspace is offense or defense dominant? Newer work has begun to challenge the technological determinism endemic to the study of cyber conflict and apply instead reframe the existing theory these lessons to a broader security dilemma discussion and to the dangers of escalation. This emerging literature could be expanded further still, to include more empirical work and a greater focus on the differences between the strategic dynamics present in offensive and defensive efforts.

In the 2016 report, surveyed work included Andres (2012), Liff (2012), Fielder (2013), Gartzke (2013), Peterson (2013), Saltzman (2013), Rid and Buchanan (2014), Valeriano and Maness (2014), and Gartzke and Lindsay (2015).[15] Other notable works published prior to 2017 include Libicki (2007), Lin (2010), Lin (2012), Malone (2012), Fielder (2013), Kello (2013), Lieber (2013), Lindsay (2013), Rid (2013) and Harknett and Goldman (2016).[16] Additional works published within the last year include Buchanan (2017), Slayton (2017), and Smeets (2017).[17] Papers and posters on this topic at ISA 2017 include Buchanan, Loleski, Monte, and Slayton.[18]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around offense versus defense and the security dilemma. We highlight three here:

1. Is cyberspace offense or defense dominant?

What factors determine whether cyberspace is offense or defense dominant? Does a theory built around the relative ease of holding or taking physical geographic territory make sense when applied to cyberspace?

There is widespread support in academic and policy circles for viewing cyberspace as offense dominant (e.g. Libicki (2007), Nye (2010), Liff (2012), Kello (2013), Lieber (2013)). However, a vocal minority (e.g. Lindsay (2013) and Rid (2013)) is pushing back against the claim that offense has the upper hand.

Embedded within the question of whether offense or defense has the upper hand is a series of arguments around how to determine or measure this balance. One subgrouping of scholars (e.g. Malone (2012) and Saltzman (2013)) assesses the relative costs of taking versus holding cyber territory to determine the offense-defense balance, while others (e.g. Buchanan (2017)) examine whether offensive cyber operations have a first-mover advantage. Additionally, some scholars tie their arguments around the primacy of the offense directly to the challenges facing deterrence in cyberspace (see previous section of this White Paper). In their efforts to determine the balance, however, Gartzke and Lindsay (2015) distinguish between ease of deception and ease of attack in cyberspace. This distinction proves important, as they argue that, while these two dynamics are often conflated, the former's impact on the offense-defense balance remains largely uncertain.

Scholarship on the cyber offense-defense balance has not been limited to the cyber domain. Gartzke (2013), for example, posits that offense dominance in cyberspace corresponds with defense dominance in terms of kinetic conflict and physical territory by illustrating an inverse relationship between the two domains. Goodman (2010), Andres (2012), Gartzke (2013), Lindsay (2013), and Valeriano and Maness (2014) also highlight that while offense may dominate in cyberspace, this balance likely does not affect the broader balance of power in the international system.

More recently, Slayton's (2017) work seeks to move away from the largely shared assumption that cyberspace is offense dominant. Slayton argues that the costs of cyber operations are determined by organizational skills and capacity rather than some quality intrinsic to the technology itself. As such, we should move away from the dominant approach of conceptualizing this balance in terms of technology and toward skilled practice.

As technology moves forward, however, interest in using offense and defense categories for classification and analysis persists. For example, Frank L. Smith III, a senior lecturer at the University of Sydney, has sought to apply the offense-defense balance to his work on quantum computing.[19]

2. How should we characterize the security dilemma in the context of cyberspace?

Compared to the debate around offense-defense dominance in cyberspace, this area remains largely understudied. At ISA 2017, two presentations focused on questions around the security dilemma: Loleski and Buchanan. Buchanan (2017) also published a book this past year examining the cybersecurity dilemma, with a particular focus on its mitigation. Earlier work in this area focused on select aspects of these dynamics, such as deception (Gartzke and Lindsay (2015)), cyber posturing (Saltzman (2013)), offensive cyber weapon acquisition and deployment (Smeets (2017)), and escalation (Lin (2012) and Fielder (2013)).

3. Do the strategic dynamics of cyberspace differ between offensive and defensive efforts? If so, how?

We need to decouple offensive and defensive conversations around cyber conflict, since the dynamics present in one may very well differ from the dynamics present in the other. In the panel discussion, two dueling narratives emerged. On the one hand, powerful states dedicate billion-dollar budgets and make constant calls for more resources and money to develop their cyber capabilities. On the other hand is the single actor, such as a teenager successfully hacking NASA. North Korea is becoming an active and sometimes very effective actor in cyber conflict. The first example seems to support the narrative of cyberspace favoring traditionally strong actors. The second and third examples support the narrative of cyberspace challenging traditional definitions of power. Does cyber conflict favor the strong and powerful? Or does it empower a wide variety of actors? In other words, is cyber power centralized or diffuse? And given that the latter examples are largely offensive in nature, does the answer differ when talking about offensive versus defensive aspects of this threat space?

As for other sub-categories discussed in this White Paper, work on the offensive and defensive characteristics of cyber conflict and resulting security dilemmas is highly theoretical and would benefit from more empirical study and focus on cases outside of the U.S., Russia, and China.

Power and Influence

What is cyber power, and how can we measure it? What does "net assessment" mean in cyberspace? What qualifies as influence in cyberspace, and how can we measure it? Is cyber power soft or hard power, or both? These are the types of questions that animate the emerging literature on power and influence in cyberspace.

In the 2016 report, surveyed work included Rattray (2001), Libicki (2007), Kramer (2009), Nye (2010), Betz and Stevens (2012), Rid and McBurney (2012), Sheldon (2012), Healey (2013), Lindsay (2013), and Segal (2016).[20] Other notable works published prior to 2017 include Kramer et al. (2010), Klimburg (2011), Betz (2012), Ebert and Maurer (2013), and Sheldon (2014).[21] Additional works published within the last year include

Borghard and Lonergan (2017), Kello (2017), Maurer (2017), and Sharp (2017).[22] Papers and posters on this topic at ISA 2017 include Langø and Maness.[23]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around power and influence. We highlight three here:

1. How much of a role does soft power play in cyber power?

Discussion during the panel focused on how to categorize cyber power within the soft power-hard power divide. Stuxnet, often regarded as the world's first cyber weapon, destroyed Iranian centrifuges used to enrich uranium. Information warfare deployed during the 2016 U.S. election targeted public opinion in key demographics. These are both examples of cyber conflict, yet the former falls more neatly into traditional conceptualizations of hard power, while the latter is an example of psychological and information warfare and the use of soft power. Given the plethora of work from the last year on cyber coercion (Borghard and Lonergan (2017), Sharp (2017), and both Langø and Maness at ISA 2017), does it make sense to think of cyber coercion as largely hard or soft? Empirically, what have been the most prevalent forms of coercion in the past? Is this likely to change, and under what conditions?

2. How do we measure cyber power?

One potential definition of "cyber power" is the ability to inflict damage on adversaries and defend against outside attack. Though this categorization of the definition could be taken to apply to any form of power. But the relationship between capabilities and vulnerabilities is complex in cyberspace. The U.S., for example, is simultaneously one of the most capable and one of the most vulnerable countries in cyberspace. North Korea is both capable and largely disconnected, leaving it with few vulnerabilities. As these examples demonstrate, power in the cyber realm must be more than a ratio between capability and vulnerability.

Within existing scholarship, efforts have been made to categorize different elements of state cyber power. Take, for example, the following two notable frameworks. In their book, Betz and Stevens (2012) identify four types of cyber power: (1) coercion to modify behaviors of another actor, i.e. "compulsory cyber-power"; (2) control over a cyberspace actor through institutions, i.e. "institutional cyber-power," (3) maintenance of the overall structures, or network society, in which all actors are embedded, i.e. "structural cyber power"; and (4) the production and dissemination of discourse through cyberspace, i.e. "productive cyber power." In contrast, and looking specifically at the state, Klimburg (2011) breaks cyber power down into three different components: (1) operation and policy coordination within the state, i.e. "integrated government capability"; (2) coherent policy within international institutions and agreements, i.e. "integrated systems capability"; and (3) cooperation with non-state actors, i.e. "integrated national capability."

As part of the discussion about defining cyber power, scholars are also grappling with the theoretical and empirical task of measuring it. As explicitly discussed in the 2016 State of the Field report, "if existing measures usually count tangible things, how can this be adjusted for cyber, especially since many aspects of cyber power are confidential, deceptive, or intangible?"

3. In what ways does cyber power alter the broader distribution of power?

We often speak of cyber power in isolation from other forms of power, but how do they combine across domains? Is it merely additive? Are particular forms of power in other domains undermined by a lack of cyber power? Can they directly determine levels of cyber power? Take, for example, arguments that correlate economic and private domestic industry strength to components of cyber power. The U.S. and Israel are commonly cited in this type of analysis, but in the context of security policy, highly networked militaries like that of the U.S. also face significant cyber vulnerabilities. In this instance, a certain level of cyber power is required to ensure the continued utility of other forms of power. Given these complex interrelations, how, from an empirical and a theoretical standpoint alike, can we measure the relationship between various forms of power?

We care about power because it is the central currency of international politics. Power, influence, and coercion all allow actors to pursue outcomes they desire. The focus then becomes the ways in which cyber power and the distribution of power alter the broader ability of states and other actors to pursue or achieve their strategic goals. Cyber power and politics are not independent of the broader geo-strategic positions of the states themselves. Consistent with this view, Ebert and Maurer (2013) highlight how, with the economic rise of the BRICS and increasing divergence between U.S. preferences and those of Brazil, Russia, India, China, and South Africa, cyberspace has become increasingly contested.

Can cyber power alter broader balances of power in international politics? Betz (2012) argues that the effect of cyber power on the international balance of power is relatively small, which is consistent with the views of scholars arguing that the offensive dominance of cyber space is unlikely to significantly alter the balance of power (see previous section). Nye (2010) looks at the diffusion of power in cyberspace but similarly concludes that this diffusion should not be mistaken for equal distribution.

In his 2017 book, Kello identifies three types of shocks, or "revolutions," occurring in international relations due to the introduction of cyber conflict. First, it empowers new actors and challenges the supremacy of states as the fundamental building blocks of the Westphalian system. Second, it empowers a new set of "revolutionary states," and through a shift to the balance of power alters the order of international politics. Third, it presents entirely new dynamics of engagement between actors and/or states that, in turn, alter our understanding of dynamics such as deterrence and the offense-defense balance. Kello's breakdown of the three levels of shocks now impacting international relations raises the question, which actors are empowered by this process and which are disadvantaged? How has it altered the types of tools these actors need to deploy in international politics to maintain power or influence? How might these shocks to pre-established distributions alter the overall balance of power?

Persisting gaps in the broader literature include the ways cyber power interacts across domains and the role it plays in broader global distributions of power for state and non-state actors. In both queries, the question of how to measure power and changes in power remains a central and critical challenge.

Foreign Policy and Doctrine

How do cyber operations impact foreign policy? What shapes the foreign policy doctrines of specific actors? In 2016, this was one of the most dynamic discussions during the panel. In 2017, our focus on this topic moved toward the states that remain outside the emerging literature and analysis. As was reported in the 2016 report, few in-depth studies of cyber operations impact on decision makers exist. Existing studies focus on the U.S, Russia, China, Israel, and the U.K., leaving gaps for theoretical and empirical work within this sub-category of research.

Work surveyed in the 2016 report includes Cavelty (2007), Gvosdev (2012), Reveron (2012), Guitton (2013), Inkster (2013, 2016), Junio (2013), Segal (2013), Axelrod and Iliev (2014), Gompert and Libicki (2014), Lindsay (2014), Geers (2015), Jaitner and Mattson (2015), Lindsay, Cheung, and Reveron (2015), Kaplan (2016) and Maness and Valeriano (2016).[24] Additional works published within the last year include Inkster (2016).[25] Papers and posters on this topic at ISA 2017 include Barrinha and Renard.[26]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around foreign policy and doctrine. We highlight four here:

1. How do leaders make decisions around cyber conflict? Which heuristics are useful?

How have different states approached cyber foreign policy and how are they developing cyber doctrine or incorporating cyber capabilities and operations into existing doctrine? Will the crafting of cyber foreign policy follow the pattern established by previous new technologies and domains, such as space?

2. Which countries' foreign policies or doctrines will shape the deployment of these technologies?

Which actors have the greatest ability to shape cyberspace? Historically, the U.S. has been the dominant actor. Will the character of cyberspace and cyber diplomacy change if, as Inkster (2016) posits, China replaces the U.S.? How can we measure these changes?

There is another line of inquiry that rests under the question of doctrine diffusion and the role of institutions and alliances. In what ways have alliances such as NATO, ANZUS, or U.S.-Japan been conduits for disseminating U.S. doctrine abroad? What role have these institutions, and others like them, played in shaping doctrine?

Estonia played a central role in Ukraine's response to Russian cyber intrusion. To what extent do neighboring states or particular sets of states serve as information sharing hubs for other states?

3. What can comparative studies add to the discussion?

We are in dire need of comparative studies outside of the dominant countries that occupy much of the scholarship. Ideally, our case studies would cover countries of varying geographic sizes and locations, levels of economic strength, military capabilities, industry characteristics, etc. In addition to work on the U.S. (e.g. Cavelty (2007), Segal (2013), and Gompert and Libicki (2014)), China (Inkster (2013), Segal (2013), Gompert and Libicki (2014), Lindsay et al. (2015), and Inkster (2016)), and Russia (Gvosdev (2012), Geers (2015), and Jaitner and Mattson (2015)), we hope to see a greater diversity of scholarship emerging on Australia, Brazil, France, Germany, Israel, Iran, Japan, New Zealand, North Korea, Singapore, South Africa, South Korea, and the U.K.

4. How do international organizations go about setting foreign policy and doctrine in this space? Or are they avoiding these types of decisions altogether?

We need more scholarship focusing on supranational and intergovernmental actors such as the United Nations, NATO, and the EU.

The Relationship Between State and Non-State Actors

Given the importance of non-state actors in cyberspace, there is a need for more work from within traditional international security as well as other fields. While the unit of interest in a great deal of security studies is the nation-state, there is robust literature on terrorism, organized crime, and other non-state actors.

In the 2016 report, surveyed work included two studies of non-state actors: Benson (2014) and Weinmann (2015).[27] Other notable works published prior to 2017 include Deibert (2003), Cavelty and Suter (2009), Applegate (2011), Healey (2011), Rattray and Healey (2011), Segal (2012), Fielder (2013), Bussolati (2015), Golden (2015), Tropina and Callanan (2015), and Borghard and Lonergan (2016).[28] Additional works published in this category in 2017 include Gross, Canetti, and Vashdi (2017), King et al. (2017), and Maurer (2017).[29] Papers and posters on this topic at ISA 2017 include Christensen.[30]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around state and non-state actors. We highlight three here:

1. What international security literatures that focus on non-state actors could be applied to this realm?

The many international security literatures that do not take the state as the dominant unit of analysis should be applied to cyber conflict. These include works on terrorism, intrastate conflict, organized crime, piracy, and public-private partnerships. There is also an extensive literature in international relations focused on corporations, non-governmental organizations, and regional and international institutions. This type of scholarship is present but sparse in the study of cyber conflict. For example, Benson (2014), Gabriel (2015), and Gross, Canetti, and Vashdi (2017) have published work examining cyberterrorism. Applegate (2011) and Segal (2012) focused on the emergence of cyber militias. Tropina and Callanan (2015) focused on the Internet industry's role in cybersecurity and crime, and Golden (2015) focused on the creation of new private-public partnerships. Most recently, King (2017) focused on social media and information campaigns. Future work drawing on these literatures and delineating how cyber conflict diverges from existing models would help to close the non-state scholarship gap in the emerging field.

2. What is the importance of proxy actors? What are the characteristics of relationships between states and non-state actors?

Our discussion referenced two types of relationships: non-state actors as proxies for states and public-private partnerships. Scholarship is emerging on both: works by Borghard and Lonergan (2016) and Maurer (2017) examine proxy actors and mercenaries in cyberspace, while works by Cavelty and Suter (2009), Golden (2015), and Tropina and Callanan (2015) examine public-private partnerships. Can we create a more general typology of relationships between state and non-state actors? What types of relationships exist beyond public-private partnerships, proxies, militias, and mercenaries? What types of case studies would be useful for illuminating these relationships and their strategic consequences?

3. Can non-state actors can be both victims and perpetrators in cyber conflict.

Non-state actors such as proxies, militias, terrorist groups, and individuals can be perpetrators of cyber-attacks. Yet non-state actors such as infrastructure or utility providers, private companies, and individuals can also be the victims of these attacks. Other non-state actors provide tactical support and are neither perpetrator nor victim but rather quasi-first responders. Take Computer Emergency Readiness Teams (CERTs) as an example. What types of relationships exist between states and non-state actors with different connections to cyber conflict?

In conclusion, in order to fully grapple with the role non-state actors play in cyber conflict, we must first classify those actors by type, the characteristics of their relationships with each other and with states, and the role they play in conflicts.

Norms and Norm Diffusion

The discussion of norms and norm diffusion was relatively developed compared to many of the other subtopics presented in this White Paper. While the Copenhagen School has taken the lead on researching the securitization of cybersecurity, there remain many topics of interest to scholars and practitioners. These questions include: How do states promote norms in cyberspace, and when do they accept them? Do existing norms apply, or will new ones be developed? What norms are emerging? Does just war theory apply? How are non-state actors engaged in norm entrepreneurship?

In the 2016 report, surveyed work included Nissenbaum (2005), Hansen and Nissenbaum (2009), Maurer (2011), Cavelty (2015), Grigsby (2015), and Mazanec (2015).[31] Other notable works published prior to 2017 include Finnemore (2011), Tikk (2011), Yannakogeorgos (2011), Stevens (2012), Hurwitz (2014), Farrell (2015), Erskine and Carr (2016), and Osula and Rõigas (2016).[32] Additional works published within the last year include Farrell and Glaser (2017).[33] Papers and posters on this topic at ISA 2017 include Barrett and Shamai and Mazanaec.[34]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around norms and norm diffusion. We highlight four here:

1. Can we begin to determine norms in cyberspace?

Do we have examples of norms and norm formation around cyber conflict? Are there issues that certain actors are championing as norms but that have yet to rise to the standard of a norm? Have any taboos been established? In what areas, such as state-sponsored espionage, have international norms simply not been significantly pursued or developed yet? n Across all of these questions one question stands out, how can we prove the existence of a norm? Are apparent examples of self-restraint around certain norms enough?

Identifying norms in cyberspace is particularly challenging. Erskine and Carr (2016) point to three complicating factors. First, cyberspace is a domain in which practices and the technology they stem from are both new and rapidly changing. Second, there are persisting, competing value systems that create additional tension for and differing perceptions of cybersecurity. Third, while norms set guidelines and expectations for behavior by specific actors, cyberspace involves a wide range of potential actors, and it can be unclear to which actor (or moral agent) a norm might apply.

With these limitations in mind, much of the discussion here remains prescriptive (what norms we would like to see developed) or theoretical (how cyber norms might emerge and why). Tikk (2011), for example, argues that four sources of law—soft, organizational, national and international agreements, and customary law—are necessary for cyber security. In contrast, Finnemore (2011) applies lessons from the emergence of international norms in the past to predict patterns for diffusion, contestation, modification, and acceptance of potential cyber norms. In a departure from largely theoretical and/or prescriptive work, Yannakogeorgos (2011) conducted an empirical examination of an existing cyber norm by tracing the development of ICANN Internet governance over the past decade.

Significant attention has also been paid to the application of legal norms and agreements around conflict and warfare to the cyber domain. This topic is covered in more detail in the Law White Paper in this series.

2. What factors are driving norm creation?

If specific norms are in fact emerging, what factors are driving them? Is it concerns around destabilization or escalation? Are they then emerging from a general consensus that cyber conflict without norms would be dominated by undesirable characteristics? Or perhaps these norms are driven by states with limited resources and significant external threats. Perhaps, as Farrell and Glaser (2017) argue, norms in U.S. doctrine are created in response to effects rather than means. Cyber weapons, then, only require a new set of norms if their effects diverge from those for which norms already exist.

In this process, are norms (or potential norms) from different domains or communities competing? Do norms from business communities crash into norms from departments of defense? Are some states advocating for particular norms but encountering resistance from others? Where, by contrast with the points of convergence discussed above, are the points of tension?

3. What role are non-state actors playing in establishing norms?

Given that non-state actors play a much more central role in this security space than in others before it, how are they shaping norm formation? Which types of actors are taking the lead? What are their interests? Take Microsoft's efforts to establish a digital Geneva Convention as just one possible example.

International Institutions and Cooperation

This category is a broad catch-all for many different forms of cooperation, such as military alliances, governance, regime creation, and international institutions. Given the lack of depth and breadth in studies emerging so far, somewhat disparate works found themselves grouped into this single bucket. During the panel discussion, this category was pointed to as a rich area for future research. Interestingly, a large number of the papers and posters presented at ISA 2017 were on these topics.

While a broad international institutions and cooperation category was not part of the 2016 report, highlighted work focusing on a subset of cooperation —"Cyber Arms Control Institutions and Regimes"—included Dipert (2010), Geers (2010), Knake (2010), Lin (2012), Schmitt (2013), and Valeriano and Maness (2014).[35] Other notable works published prior to 2017 in the broader cooperation and international institutions literature include Axelrod (2010), Hathaway (2010), Hunker (2010), Tikk (2010), Healey and Bochoven (2011), Hurowitz (2012), Forsyth (2013), Goldsmith (2013), Clark et al. (2014), DeNardis (2014), and Shackelford and Craig (2014).[36] Additional work published in 2017 includes Lindsay (2017) and Rovner and Moore (2017).[37] Papers and posters on this topic at ISA 2017 included An, Brandao and Camisao, Coleman, Diersch, Griffith, and Yoo.[38]

Several central questions emerged from our preparatory materials and the subsequent panel discussion around international institutions and cooperation. We highlight one here:

1. What about international security literatures that focus on cooperation?

Within international security, numerous literatures focus on cooperation in the face of security threats. Pulling these literatures into discussions around cyberspace would fill a clear and significant gap in existing scholarship. Early efforts in this regard remain sporadic and nascent.

Of note is emerging scholarship testing mechanisms identified as driving security cooperation more broadly within the context of cyberspace. For example, Rovner and Moore (2017) investigate whether hegemonic stability theory applies to cyberspace, specifically whether hegemonic leadership on the part of the U.S. ameliorates collective action problems in cyberspace. Forsyth (2013) also looked at the role of great powers. Ultimately, he argued, in contrast to Hurwitz's (2012) view, that great powers have provided existing structures for cyberspace. In other words, "cyberspace is what great powers make of it."

In addition, the literatures on alliance formation, evolution, and termination can be applied to questions on the utility of certain types of alliances for cyber conflict, whether existing alliances will be utilized, and how alliances will evolve given this new threat. Griffith's 2017 ISA paper on alliance theory sought to draw attention to the utility of these theories by examining U.S. alliances. By specifically addressing the utility of alliance theory, this work builds on previous research regarding the role of military alliances in cyber-defense, which had largely taken two forms: (1) empirical, as in Healey and Bochoven's 2011 work, examining how existing alliances have responded to potential cyber conflicts, and (2) prescriptive, as in works by Hathaway (2010), Hunker (2010), and Tikk (2010), asking how existing alliances should respond to the changing threat environment.

Similarly, arms control literature can be drawn upon to examine cyber arms control and its limitations. Geers (2010) and Lin (2012) provide a useful starting place here.

Two broader fields of study or literatures that extend beyond security cooperation are also worth mentioning here. First, literatures focusing on international law can be and have been applied to cyber conflict. The *Tallinn Manuals* (see Schmitt (2013)) represent a significant effort in this regard. Law and legal cooperation is discussed in greater detail in the White Paper on the legal and ethical issues embedded within the study of cyber conflict. Second, questions around the governance of cyberspace likewise span multiple White Papers, touching upon issues ranging from legal to economic concerns and from operational to strategic dynamics. Out of all the cooperative efforts examined so far, governance has received a significant amount of attention. Take, for example, Axelrod (2010), Clark et al. (2014), and Shackelford and Craif (2014), as well as the ongoing debate between multilateral (state and public agency led) and multi-stakeholder (public-private partnership) governance models.

In conclusion, international security cooperation remained one of the core gaps identified in our preparation for the panel and in the subsequent panel discussion. One possible first step for future work is to examine the applicability of existing theories of security cooperation to cyberspace. As that work is completed, we hope to see this category break apart into several distinct sub-literatures.

Panel and Plenary Discussions

During the panel discussion directly following our presentation of the state of the field on this topic, audience members raised several questions that did not easily fit into the buckets above. The following is an overview of four central questions for your reference.

1. In what ways does cyberspace affect other security interests?

We should also be looking at how cyberspace affects other outcomes of interest, such as nuclear stability. In addition, we should ask how cyber tools affect the broader outcome of interest in each literature, such as balance of power, cooperation, stability, escalation, etc. How do cyber tools alter, for example, the overall offense-defense balance at a strategic level?

2. Where do psychological and information warfare fit?

This question was raised in the 2016 panel and again in 2017 with more vigor, given the recent Russian information operations during the U.S. and French elections. Psychological warfare requires more study both as an independent form of warfare and as a domain where cyber tools play a role.

3. How should we approach potential drivers of change? Should we expect all or some of the dynamics that define cyber conflict today to persist in the future?

As a field, we need to be conducting future-oriented research rather than assuming that the dynamics we observe today will persist. How might the strategic dynamics of conflict be different in five years? In ten years? In twenty years? What new technologies (e.g., artificial intelligence) might drive that change?

4. How can we integrate economic concerns into cyber conflict and cyber security?

Our adversaries are attempting to use economic means to hurt us and help themselves. This plays into the broader strategic concerns of any given country. What are the boundaries between economic and security concerns? Is there agreement on these boundaries?

Summary and Recommendations

The purpose of this White Paper was threefold: (1) to capture the evolution of the field since the 2016 State of the Field Conference, (2) to provide an overview of the existing scholarship focusing on the strategic dynamics of cyberspace, and (3) to identify where more rigorous research could expand the frontiers of the field.

The paper also highlighted four takeaways. First, there is a persisting tension between academics and policy practitioners regarding the utility of international security to the study of cyber conflict. Second, the topics the field has chosen to study remained largely consistent between the 2016 conference and the 2017 conference. Deterrence remains over-researched, while domestic politics remains severely underdeveloped. International cooperation saw the most attention in the form of ISA papers. Empirical work and case studies are focused on a handful of countries. Third, the conceptual boundaries between tactical, operational, and strategic dynamics in cyberspace are difficult to maintain, but still lead to useful research. Fourth, as a field, we need to pay attention to both the destructive outcomes of cyber conflict and the motivating strategic goals and outcomes for which cyber tools are being mobilized. These are two separate intellectual inquiries and should be clearly distinguished from one another.

It is also important to note that there remain some key limitations to the design of this panel and subsequent White Paper. First, some important new works may be missing from this overview. We hope to continue to expand the scope of the work represented and encourage readers to point us to work, new and old, that we may have missed. Second, no single discipline will be able to comprehensively address the realities of cyber conflict. International security, as highlighted in this White Paper, maintains broad utility around certain sets of questions, while other fields of inquiry will provide additional insight into others. Moreover, while it is important to make reference to the international security sub-literatures explored here, it will also be productive to show the ways in which some of these theories and models hold limited or incorrect explanatory power in this new threat space. That said, many of these models continue to hold significant utility and should not simply be discarded on the assumption that "everything has changed." Third, this overview of the state of the field is centered around an academic political science audience. The questions policy practitioners are asking likely have some overlap with those discussed but also branch out into different interests, which are not captured here.

With these limitations in mind, we hope that this White Paper proves useful as both a starting place for new entrants into this field and a reference for those already enmeshed in these debates. To echo Maness and Segal's 2016 report, scholars need to continue to be "theoretically innovative and empirically grounded" to move the field forward. During the 2018 State of the Field Conference, we hope to see many of the questions outlined above answered, as well as new efforts to fill those gaps that persist.

About the Authors

Melissa K. Griffith is a Ph.D. Candidate in Political Science at the University of California, Berkeley.

Adam Segal is the Ira A. Lipman Chair in Emerging Technologies and National Security and Director of the Digital and Cyberspace Policy Program at the Council on Foreign Relations.

End Notes

1. The 2017 Conference was the second State of the Field of Cyber Conflict meeting. The event was co-hosted by the Cyber Conflict Studies Association and the Saltzman Institute of War and Peace Studies at Columbia University's School of International and Public Affairs (SIPA). For an overview of the conference, refer to www.cyberconflict.org/blog/2017/6/23/state-of-the-field-of-cyber-conflict-workshop-june-2017.html

2. To review Maness and Segal's 2016 State of the Field assessment, refer to SIPA's "Cyber Conflict State of the Field Workshop Report": http://static1.1.sqspcdn.com/static/f/956646/27604316/1498262610577/SOTF_Review_Copy.pdf?token=OROvr5%2FKs5iuP2UX8tX6Cv-J%2FI0U%3D

3. Griffith, Melissa K. and Trey Herr. 2017. "Is the Strategic Corporal on Your Twitter Feed?" in *Net Politics* and *Digital and Cyberspace Policy Program* from the Council on Foreign Relations: July 12. www.cfr.org/blog/strategic-corporal-your-twitter-feed

4. For examples of work on the RMA, which falls largely outside the scope of this specific review of literature, refer to Metz, Steven and James Kievit. 1994. "The Revolution in Military Affairs and Conflict Short of War." *United States Army War College Strategic Studies Institute*; Blank, Stephen J. 1996. "Preparing for the Next War." *Strategic Review* 24(2): 17–25; Biddle, Stephen. 1996. "Victory Misunderstood: What the Gulf War Tells Us About the Future of Conflict." *International Security* 21(2): 139–179; Cohen, Eliot A. 1996. "A Revolution in Warfare." *Foreign Affairs*; Davis, Norman C. 1996. "An Information-Based Revolution in Military Affairs." *Strategic Review* 24(1): 43–53; Friedman, George and Meredith Friedman. 1998. *The Future of War: Power, Technology and American World Dominance in the Twenty-first Century* (St. Martin's Griffin Publishers); Gongora, Thierry and Harald Von Riekhoff. 2000. *Toward a Revolution in Military Affairs? Defense and Security at the Dawn of the Twenty-First Century* (Greenwood Press); Andréani, Gilles, Christoph Bertram, and Charles Grant. 2001. *Europe's Military Revolution* (Center for European Reform); and Sloan, Elinor C. 2002. *The Revolution in Military Affairs*. 1st 3dition (McGill-Queen's University Press).

5. Lindsay, Jon R. and Lucas Kello. 2014. "Correspondence: A Cyber Disagreement." *International Security* 39(2): 181–192.

6. Arquilla, John and David Ronfeldt. 1993. "Cyberwar Is Coming!" Comparative Strategy 12(2): 141–65. doi:10.1080/01495939308402915; Clarke, Richard A. and Robert Knake. 2011. *Cyber War: The Next Threat to National Security and What to Do About It.* Reprint edition (Ecco); Choucri, Nazli. 2012. *Cyberpolitics in International Relations* (MIT Press); Junio, Timothy. 2013. "A Theory of Information Warfare." University of Pennsylvania dissertation; Kello, Lucas. 2013. "The Meaning of the Cyber Revolution: Perils to Theory and Statecraft." *International Security* 38(2): 7–40; Lawson, Sean. 2013. "Beyond Cyber-Doom: Assessing the Limits of Hypothetical Scenarios in the Framing of Cyber-Threats." *Journal of Information Technology & Politics* 10(1): 86–103. do i:10.1080/19331681.2012.759059; Rid, Thomas. 2013. *Cyber War Will Not Take Place* (Hurst & Company); and Lindsay, Jon R. and Lucas Kello. 2014. "Correspondence: A Cyber Disagreement." *International Security* 39(2): 181–92. doi:10.1162/ISEC_c_00169.

7. Kello, Lucas. 2017. *The Virtual Weapon and International Order*. Kindle edition (Yale University Press); Lindsay, Jon Randal. 2017. "Restrained by Design: The Political Economy of Cybersecurity." *Digital Policy, Regulation and Governance* 19(6); Perkovich, George and Ariel E. Levite, eds. 2017. *Understanding Cyber Conflict: Fourteen Analogies* (Georgetown University Press).

8. For one such example, refer to Mulligan, Deirdre K. and Fred B. Schneider. 2011. "Doctrine for Cybersecurity." *Daedalus* 140(4): 70–92, who apply the provision of public health as a public good to cybersecurity provisions.

[9] Kugler, Richard L. 2009. "'Deterrence of Cyber Attacks.'" In *Cyberpower and National Security*, Franklin Kramer, Stuart H. Starr, and Larry K. Wentz, eds. 1st edition (Potomac Books): 309–42; Libicki, Martin C. 2009. *Cyberdeterrence and Cyberwar* (RAND). www.books24x7.com/marc.asp?bookid=54204; Goodman, Will. 2010. "Cyber Deterrence: Tougher in Theory than in Practice?" *U.S. Senate Washington, DC Committee on Armed Services* 4(3). http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA528033; National Research Council, *Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options for U.S. Policy* (The National Academies Press); Nye, Joseph S. 2011. "Nuclear Lessons for Cyber Security?" *Strategic Studies Quarterly* 5(4): 18–38; Cooper, Jeffrey. 2012. "A New Framework for Cyber Deterrence." In *Cyberspace and National Security: Threats, Opportunities, and Power in a Virtual World*, Derek S. Reveron, ed. (Georgetown University Press): 105–120. www.jstor.org/stable/j.ctt2tt6rz; Valeriano, Brandon and Ryan C. Maness. 2014. "The Dynamics of Cyber Conflict between Rival Antagonists, 2001-2011." *Journal of Peace Research* 51(3): 347–60. doi:10.1177/0022343313518940; and Gartzke, Erik and Jon R. Lindsay. 2015. "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace." *Security Studies* 24(2): 316–48. doi:10.1080/09636412.2015.1038188.

[10] Harknett, Richard J. 1996. "Information Warfare and Deterrence." *Parameters*: 93–107; Clark, Richard A. and Robert K. Knake. 2010. *Cyber War: The Next Threat to National Security and What to Do About It* (HarperCollins); Liff, Adam P. 2012. "Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War." *Journal of Strategic Studies* 35(3): 401–428; Tsagourias, Nicholas. 2012. "Cyber Attacks, Self-Defence and the Problem of Attribution." *Journal of Conflict and Security Law* 17(2): 229–44; Gartzke, Erik. 2013. "The Myth of Cyberwar: Bringing War in Cyberspace Back Down to Earth." *International Security* 38(2): 41–73; Iasiello, Emilio. 2013. "Is Cyber Deterrence an Illusory Course of Action?" *Journal of Strategic Security* 7(1): 54–67; Schmitt, Michael N. and Liis Vihul. 2014. "Proxy Wars in Cyberspace: The Evolving International Law of Attribution." *Fletcher Security Review* 1(2): 54–73; Lindsay, Jon R. 2015. "Tipping the Scales: The Attribution Problem and the Feasibility of Deterrence against Cyberattack." *Journal of Cybersecurity* 1(1): 53–67; Healey, Jason. 2016. "Beyond Attribution: Seeking National Responsibility in Cyberspace." *Atlantic Council*; and Lupovici, Amir. 2016. "The 'Attribution Problem' and the Social Construction of 'Violence': Taking Cyber Deterrence Literature a Step Forward." *International Studies Perspectives*.

[11] Borghard, Erica D and Shawn W. Lonergan. 2017. "The Logic of Coercion in Cyberspace." *Security Studies* 26(3): 452–481; Chen, Jim. 2017. "Deterrence and its Implementation in Cyber Warfare." In *ICCWS 2017-Proceedings of the 12th International Conference on Cyber Warfare and Security* (ACPIL); Davis II, Jon S., Benjamin Adam Boudreaux, Jonathan William Welburn, Jair Aguirre, Cordaye Ogletree, Geoffrey McGovern, and Michael S. Chase. 2017. *Stateless Attribution: Toward International Accountability in Cyberspace* (RAND); Edwards, Benjamin, Alexander Furnas, Stephanie Forrest and Robert Axelrod. 2017. "Strategic Aspects of Cyberattack, Attribution, and Blame." *Proceedings of the National Academy of Sciences of the United States of America* 114(11): 2825–2830; Fischerkeller, Michael. 2017. "Incorporating Offensive Cyber Operations into Conventional Deterrence Strategies." *Survival: Global Politics and Strategy* 59(1): 103–134; Harknett, Richard J. and Michael P. Fischerkeller. 2017. "Deterrence is Not a Credible Strategy for Cyberspace." *Orbis* 61(3); Harknett, Richard R. and Joseph Nye Jr. 2017. "Is Deterrence Possible in Cyberspace?" *International Security* 42(2): 196–199; Mandel, Robert. 2017. *Optimizing Cyber Deterrence: A Comprehensive Strategy for Preventing Foreign Cyberattacks* (Georgetown University Press); Nye, Joseph. 2017. "Deterrence and Dissuasion in Cyberspace." *International Security* 41(3): 44–71; Sharp, Travis. 2017. "Theorizing Cyber Coercion: The 2014 North Korean Operation Against Sony." *Journal of Strategic Studies* 40(7): 898–926.

[12] ISA 2017: Cunningham, Fiona (MIT). "Seizing the Initiative or Controlling Escalation? China's Changing Approach to Cyber Deterrence"; Lupovici, Amir (Tel Aviv University). "Israel and the (Social) Construction of Cyber Deterrence"; and Wilner, Alex (Carleton University). "State and Non-State Cyber Deterrence: Bridging IR by Crossing Disciplines."

[13] Healey, Jason. 2016. "The Cartwright Conjecture: The Deterrent Value and Escalatory Risk of Fearsome Cyber Capabilities." Submitted to *The Journal of Cybersecurity*: www.americanbar.org/content/dam/aba/administrative/law_national_security/Jason%20Healey%20The%20Cartwright%20Conjecture.authcheckdam.pdf

[14] See, e.g., Fearon, James. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52 (Spring): 269–305.

[15] Andres, Richard. 2012. "The Emerging Structure of Strategic Cyber Offense, Cyber Defense, and Cyber Deterrence." In *Cyberspace and National Security: Threats, Opportunities, and Power in a Virtual World*, Derek S. Reveron, ed. (Georgetown University Press). www.jstor.org/stable/j.ctt2tt6rz.; Liff, Adam P. 2012. "Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War." *Journal of Strategic Studies* 35(3): 401–28. doi:10.1080/01402390.2012.663252; Fielder, James D. 2013. "Bandwidth Cascades: Escalation and Pathogen Models for Cyber Conflict Diffusion." *Small Wars Journal* 9(6). http://smallwarsjournal.com/jrnl/art/bandwidth-cascades-escalation-and-pathogen-models-for-cyber-conflict-diffusion; Gartzke, Erik. 2013. "The Myth of Cyberwar: Bringing War in Cyberspace Back Down to Earth." *International Security* 38(2): 41–73. doi:10.1162/ISEC_a_00136; Peterson, Dale. 2013. "Offensive Cyber Weapons: Construction, Development, and Employment." *Journal of Strategic Studies* 36(1): 120–24. doi:10.1080/01402390.2012.742014; Saltzman, Ilai. 2013. "Cyber Posturing and the Offense-Defense Balance." *Contemporary Security Policy* 34(1): 40–63. doi:10.1080/13523260.2013.771031; Rid, Thomas and Ben Buchanan. 2015. "Attributing Cyber Attacks." *Journal of Strategic Studies* 38(1–2): 4–37. doi:10.1080/01402390.2014.977382; Valeriano, Brandon and

Ryan C. Maness. 2014. “The Dynamics of Cyber Conflict between Rival Antagonists, 2001-2011.” *Journal of Peace Research* 51(3): 347–60. doi:10.1177/0022343313518940; and Gartzke, Erik and Jon R. Lindsay. 2015. “Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace.” *Security Studies* 24(2): 316–48. doi:10.1080/09636412.2015.1038188.

[16] Libicki, Martin C. 2007. *Conquest in Cyberspace: National Security and Information Warfare* (Cambridge University Press); Lin, Herbert. 2010. “Offensive Cyber Operations and the Use of Force.” *Journal of National Security Law & Policy* 4(63): 63–86; Lin, Herbert. 2012. “Escalation Dynamics and Conflict Termination in Cyberspace.” *Strategic Studies Quarterly* 6(3): 46–70; Malone, Patrick J. 2012. “Offense-Defense Balance in Cyberspace: A Proposed Model.” *Naval Postgraduate School*; Fielder, James D. 2013. “Bandwidth Cascades: Escalation and Pathogen Models for Cyber Conflict Diffusion.” *Small Wars Journal* 9(6); Kello, Lucas. 2013. “The Meaning of the Cyber Revolution: Perils to Theory and Statecraft.” *International Security* 38(2): 7–40; Lieber, Keir. 2013. “The Offense-Defense Balance and Cyber Warfare.” In Emily O. Goldman and John Arquilla, eds., *Cyber Analogies* (Naval Postgraduate School); Lindsay, Jon R. 2013. “Stuxnet and the Limits of Cyber Warfare.” *Security Studies* 22(3): 365–404; Rid, Thomas. 2013. *Cyber War Will Not Take Place* (Oxford University Press); and Harknett, Richard and Emily Goldman. 2016. “The Search for Cyber Fundamentals.” *Journal of Information Warfare* 15(2): 81–88.

[17] Buchanan, Ben. 2017. *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations* (Oxford University Press); Slayton, Rebecca. 2017. “What is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment.” *International Security* 41(3): 72–109; and Smeets, Max. 2017. “A Matter of Time: On the Transitory Nature of Cyberweapons.” *Journal of Strategic Studies* 41(1–2).

[18] ISA 2017: Buchanan, Ben (Harvard University). “The Cybersecurity Dilemma”; Loleski, Steven (University of Toronto). “Exploit It All: Threats, Vulnerabilities, and the Cyber Security Dilemma”; Monte, Matthew (Self). “Offense-Defense Asymmetries: A View from Inside Cyber Operations”; Slayton, Rebecca M. (Cornell University). “Reframing the Cyber Offense-Defense Balance: From Technology to Skilled Practice.”

[19] Smith III, Frank L. 2017. “Quantum Technologies and International Security.” CLTC Visiting Scholar Paper Workshop, UC, Berkeley.

[20] Rattray, Gregory J. 2001. *Strategic Warfare in Cyberspace* (MIT Press); Libicki, Martin C. 2007. *Conquest in Cyberspace: National Security and Information Warfare*. 1st edition (Cambridge University Press); Kramer, Franklin, ed. 2009. *Cyberpower and National Security*. 1st edition (Potomac Books); Nye, Joseph S. 2010. “Cyber Power.” Essay from the Belfer Center for Science and International Affairs, Harvard Kennedy School: http://belfercenter.ksg.harvard.edu/publication/20162/cyber_power.html; Betz, David J. and Tim Stevens. 2012. *Cyberspace and the State: Towards a Strategy for Cyber-Power*. 1st edition (Routledge); Rid, Thomas and Peter McBurney. 2012. “Cyber-Weapons.” *The RUSI Journal* 157(1): 6–13. doi:10.1080/03071847.2012.664354; Sheldon, John B. 2012. “Toward a Theory of Cyber Power: Strategic Purpose in Peace and War.” In *Cyberspace and National Security: Threats, Opportunities, and Power in a Virtual World*, Derek S. Reveron, ed. (Georgetown University Press): 207–224; Healey, Jason, ed. 2013. *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012* (Cyber Conflict Studies Association); Lindsay, Jon R. 2013. “Stuxnet and the Limits of Cyber Warfare.” *Security Studies* 22(3): 365–404. doi:10.1080/09636412.2013.816122; and Segal, Adam. 2016. *The Hacked World Order: How Nations Fight, Trade, Maneuver, and Manipulate in the Digital Age*. 1st edition (Public Affairs).

[21] Kramer, Franklin D., Stuart H. Starr, and Larry K. Wentz, eds. 2010. *Cyberpower and National Security* (National Defense University Press and Potomac Books); Klimburg, Alexander. 2011. “Mobilizing Cyber Power.” *Survival* 53(1): 41–60; Betz, David. 2012. “Cyberpower in Strategic Affairs: Neither Unthinkable nor Blessed.” *Journal of Strategic Studies* 35(5): 689–711; Ebert, Hannes, and Tim Maurer. 2013. “Contested Cyberspace and Rising Powers.” *Third World Quarterly* 34(6): 1054–1074; Sheldon, John B. 2014. “Geopolitics and Cyber Power: Why Geography Still Matters.” *The Journal of the National Committee on American Foreign Policy* 36(5): 286–293.

[22] Borghard, Erica D. and Shawn Lonergan. 2017. “The Logic of Coercion in Cyberspace.” *Security Studies* 26(3); Kello, Lucas. 2017. *The Virtual Weapon and International Order*. Kindle edition (Yale University Press); Maurer, Tim. 2017. *Cyber Mercenaries: The State, Hackers, and Power* (Cambridge University Press); and Sharp, Travis. 2017. “Theorizing Cyber Coercion: The 2014 North Korean Operation Against Sony.” *Journal of Strategic Studies*.

[23] ISA 2017: Langø, Hans-Inge (University of Texas at Austin). “Mutually Assured Vulnerability: An Ecological Approach to the Study of Coercion and Power in Cyberspace”; and Maness, Ryan (Northeastern University). “Cyber Compellence: Applying Coercion in the Information.”

[24] Cavelty, Myriam Dunn. 2007. *Cyber-Security and Threat Politics: U.S. Efforts to Secure the Information Age* (Routledge). www.tandfebooks.com/isbn/9780203937419; Gvosdev, Nikolas K. 2012. “The Bear Goes Digital: Russia and its Cyber Capabilities.” In *Cyberspace and National Security: Threats, Opportunities, and Power in a Virtual World*, Derek Reveron, ed. (Georgetown University Press). www.jstor.org/stable/j.ctt2tt6rz; Reveron, Derek S., ed. 2012. *Cyberspace and National Security: Threats, Opportunities, and Power in a Virtual World* (Georgetown University Press); Guitton, Clement. 2013. “Cyber Insecurity as a National Threat: Overreaction from Germany, France and the UK?” *European Security* 22(1): 21–35. doi:10.1080/09662839.2012.749864; Inkster, Nigel. 2013. “Chinese Intelligence in the Cyber Age.” *Survival: Global Politics and Strategy* 55(1): 45–66. doi:10.1080/00396338.2013.767405; Inkster, Nigel. 2016. *China’s Cyber Power* (The International Institute for Strategic Studies); Junio, Timothy. 2013. “A Theory of Information Warfare.” University of Pennsylvania dissertation; Segal, Adam. 2013. “The code not taken: China, the United States, and the future of cyber espionage.” *Bulletin of the Atomic Scientists*, 69(5), 38–45; Axelrod, R. and R. Iliev. 2014. “Timing of Cyber Conflict.” Proceedings of the National Academy of Sciences 111(4): 1298–1303.

doi:10.1073/pnas.1322638111; Gompert, David C. and Martin Libicki. 2014. “Cyber Warfare and Sino-American Crisis Instability.” *Survival* 56(4): 7–22. doi:10.1080/00396338.2014.941543; Lindsay, John. R. 2014. “The Impact of China on Cybersecurity: Fiction and Friction.” *International Security* 39(3): 7–47; Geers, Kenneth. 2015. *Cyber War in Perspective: Russian Aggression against Ukraine* (CCDCOE). http://scholar.google.com/scholar?cluster=9137378561972954249&hl=en&oi=scholarr; Jaitner, Margarita and Peter A. Mattsson. 2015. “Russian Information Warfare of 2014.” In *Cyber Conflict: 2015 Conference on Architectures in Cyberspace* (IEEE): 39–52. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7158467; Lindsay, Jon R., Tai Ming Cheung, and Derek S. Reveron, eds. 2015. *China and Cybersecurity: Espionage, Strategy, and Politics in the Digital Domain.* 1st edition (Oxford University Press); Kaplan, Fred. 2016. *Dark Territory: The Secret History of Cyber War* (Simon & Schuster); and Maness, Ryan C. and Brandon Valeriano. 2016. “The Impact of Cyber Conflict on International Interactions.” *Armed Forces & Society* 42(2): 301–23. doi:10.1177/0095327X15572997.

25. Inkster, Nigel. 2016. *China’s Cyber Power* (IISS).

26. ISA 2017: Barrinha, Andre Filipe (Canterbury Christ Church University and Centre for Social Studies) and Thomas Renard (Egmont, Brussels). “Cyber-diplomacy and Change in World Politics.”

27. Benson, David C. 2014. “Why the Internet is not Increasing Terrorism.” *Security Studies* 23(2): 293–328. doi: 10.1080/09636412.2014.905353; and Weimann, Gabriel. 2015. *Terrorism in Cyberspace: The Next Generation* (Columbia University Press).

28. Deibert, Ronald J. 2003. “Black Code: Censorship, Surveillance, and the Militarisation of Cyberspace.” *Millennium: Journal of International Studies* 32(3); Cavelty, Myriam Dunn and Manuel Suter. 2009. “Public-Private Partnerships are No Silver Bullet: An Expanded Governance Model for Critical Infrastructure Protection.” *International Journal of Critical Infrastructure Protection* 2(4): 179–187; Applegate, Scott D. 2011. “Cybermilitias and Political Hackers—Use of Irregular Forces in Cyberwarfare.” *IEEE Security and Privacy Magazine* 9(5): 16–22; Healey, Jason. 2011. “The Spectrum of National Responsibility for Cyberattacks.” *Brown Journal of World Affairs* 18(1): 57–69; Rattray, Gregory and Jason Healey. 2011. “Chapter 5: Non-State Actors and Cyber Conflict.” In *America’s Cyber Future: Security and Prosperity in the Information Age*, Kristin M. Lord and Travis Sharp, eds. (CNAS): 67–83; Segal, Adam. 2012. “The Rise of Asia’s Cyber Militias.” *The Atlantic*: www.theatlantic.com/international/archive/2012/02/the-rise-of-asias-cyber-militias/253487/; Fielder, James D. 2013. “Bandwidth Cascades: Escalation and Pathogen Models for Cyber Conflict Diffusion.” *Small Wars Journal* 9(6); Bussolati, Nicolò. 2015. “The Rise of Non-State Actors in Cyberwarfare.” In *Cyber War: Law and Ethics for Virtual Conflicts*, Jens David Ohlin, Kevin Govern, and Claire Finkelstein, eds. (Oxford Scholarship Online); Golden, Chris. 2015. “Creating New Private-Public Partnerships in Cybersecurity.” *National Cybersecurity Institute Journal* 2(3); Tropina, Tatiana and Cormac Callanan. 2015. “Self-and Co-regulation in Cybercrime, Cybersecurity and National Security” (SpringerBriefs in Cybersecurity); and Borghard, Erica D. and Shawn W. Lonergan. 2016. “Can States Calculate the Risks of Using Cyber Proxies?” *Orbis* 60(3): 395–416.

29. Gross, Michael L., Daphna Canetti, and Dana R. Vashdi. 2017. “Cyberterrorism: Its Effects on Psychological Well-being, Public Confidence and Political Attitudes.” *Journal of Cybersecurity* 3(1); King, Gary, Jennifer Pan, and Margaret E. Roberts. 2017. “How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument.” *American Political Science Review* 111(3): 484–501; Maurer, Tim. 2017. *Cyber Mercenaries: The State, Hackers, and Power* (Cambridge University Press).

30. ISA 2017: Christensen, Kristoffer (University of Copenhagen). “‘It Is Not Even There’: Topologies of Cyber Security in the Practice of Private Companies.”

31. Nissenbaum, Helen. 2005. “Where Computer Security Meets National Security.” *Ethics and Information Technology* 7(2): 61–73. doi:10.1007/s10676-005-4582-3; Hansen, Lene and Helen Nissenbaum. 2009. “Digital Disaster, Cyber Security, and the Copenhagen School.” *International Studies Quarterly* 53(4): 1155–75. doi:10.1111/j.1468-2478.2009.00572.x; Maurer, Tim. 2011. “Cyber Norm Emergence at the United Nations—An Analysis of the UN’s Activities Regarding Cyber-Security.” Belfer Center for Science and International Affairs, Harvard Kennedy School, discussion paper. http://belfercenter.ksg.harvard.edu/publication/21445/cyber_norm_emergence_at_the_united_nationsan_analysis_of_the_uns_activities_regarding_cybersecurity.html; Cavelty, Myriam Dunn. 2007. “The Normalization of Cyber-International Relations.” In *Strategic Trends* 2015, Oliver Thränert and Martin Zapfe, eds. (Center for Security Studies): 81–98. www.researchgate.net/publication/274076687_The_Normalization_of_Cyber-International_Relations; Grigsby, Alex. 2015. “The UN GGE on Cybersecurity: What is the UN’s Role?” *Council on Foreign Relations—Net Politics* (April) http://blogs.cfr.org/cyber/2015/04/15/the-un-gge-on-cybersecurity-what-is-the-uns-role/; and Mazanec, Brian M. 2015. *The Evolution of Cyber War: International Norms for Emerging-Technology Weapons* (University of Nebraska Press).

32. Finnemore, Martha. 2011. “Cultivating International Cyber Norms.” In *America’s Cyber Future: Security and Prosperity in the Information Age*, Kristin M. Lord and Travis Sharp, eds. (CNAS); Tikk, Eneken. 2011. “Ten Rules for Cyber Security” *Survival* 53(3): 119–132; Yannakogeorgos, Panayotis. 2011. “Cyberspace, the New Frontier – and the Same Old Multilateralism.” In *Global Norms, American Sponsorship and the Emerging Patterns of World Politics*, S. Reich, ed. (Palgrave); Stevens, Tim. 2012. “A Cyberwar of Ideas? Deterrence and Norms in Cyberspace.” *Contemporary Security Policy* 33(1):148–70; Hurwitz, Robert. 2014. “The Play of States: Norms and Security in Cyberspace.” *American Foreign Policy Interests* 36(5); Farrell, Harry. 2015. “Promoting Norms for Cyberspace.” *Council on Foreign Relations*; Erskine, Toni and Madeline Carr. 2016. “Beyond ‘Quasi-Norms’: The Challenges and Potential of Engaging with Norms in Cyberspace.” In *International Cyber Norms: Legal, Policy & Industry Perspectives* (NATO Cooperative

Cyber Defence Centre of Excellence); and Osula, Anna-Maria and Henry Rõigas, eds. 2016. *International Cyber Norms: Legal, Policy & Industry Perspectives* (NATO Cooperative Cyber Defence Centre of Excellence).

33. Farrell, Henry and Charles L. Glaser. 2017. "The Role of Effects, Saliencies and Norms in U.S. Cyberwar Doctrine." *Journal of Cybersecurity* 3(1).

34. ISA 2017: Barrett, Edward T. (U.S. Naval Academy). "Reliable Old Wineskins: The Applicability of the Just War Tradition to Military Cyber Operations"; and Shamai, Patricia (University of Portsmouth) and Brian M. Mazanaec (George Mason University). "Stigmatizing Cyber War: Mission Impossible?"

35. Dipert, Randall R. 2010. "The Ethics of Cyberwarfare." *Journal of Military Ethics* 9(4): 384–410; Geers, Kenneth. 2010. "Cyber Weapons Convention." *Computer Law & Security Review* 26(5): 547–551; Knake, Robert K. 2010. *Internet Governance in an Age of Cyber Insecurity*. Council Special Report, no. 56 (Council on Foreign Relations); Lin, Herbert S. 2012. "Arms Control in Cyberspace: Challenges and Opportunities." *World Politics Review* (March); Schmitt, Michael N., ed. 2013. *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge University Press); and Valeriano, Brandon and Ryan C. Maness. 2014. "The Dynamics of Cyber Conflict between Rival Antagonists, 2001-2011." *Journal of Peace Research* 51(3): 347–60. doi:10.1177/0022343313518940.

36. Axelrod, Robert. 2010. "Beyond the Tragedy of the Commons: A Discussion of *Governing the Commons*." *Perspectives on Politics* 8(2): 580–82; Hathaway, Melissa E. 2010. "Toward a Closer Digital Alliance." *SAIS Review of International Affairs* 30(2); Hunker, Jeffrey. 2010. "Cyber War and Cyber Power: Issues for NATO Doctrine." *NATO Defense College, Rome*, 62; Tikk, Eneken. 2010. "Global Cybersecurity–Thinking About the Niche for NATO." *SAIS Review of International Affairs* 30(2); Healey, Jason and Leendert van Bochoven. 2011. "NATO's Cyber Capabilities: Yesterday, Today, and Tomorrow." Issue Brief for *The Atlantic Council*; Hurwitz, Roger. 2012. "Depleted Trust in the Cyber Commons." *Strategic Studies Quarterly* 6(3); Forsyth, James W. 2013. "What Great Powers Make of It: International Order and the Logic of Cooperation in Cyberspace." *Strategic Studies Quarterly* 7(1); Goldsmith, Jack. 2013."Cybersecurity Treaties: A Skeptical View." Hoover Institution; Clark, David, Thomas Berson, and Herbert S. Lin, eds. 2014. *At the Nexus of Cybersecurity and Public Policy: Some Basic Concepts and Issues* (National Academies Press); DeNardis, Laura. 2014. *The Global War for Internet Governance* (Yale University Press); and Shackelford, Scott and Amanda Craig. 2014. "Beyond the New 'Digital Divide': Analyzing the Evolving Role of National Governments in Internet Governance and Enhancing Cybersecurity." *Ω*(5)0.

37. Lindsay, Jon R. 2017. "Restrained by Design: The Political Economy of Cybersecurity." Digital Policy, Regulation, and Governance 19(6); and Rovner, Joshua and Tyler Moore. 2017. "Does the Internet Need a Hegemon?" Journal of Global Security Studies 2(3): 184–203.

38. ISA 2017: An, Jungbae (Yonsei University). "A Handful of Technocrats or Supranational Policy Network: Technical Community in Global Internet Governance"; Brandao, Ana Paula (CICP, University of Minho) and Isabel Camisao (CICP, University of Coimbra). "Framing the Cybersecurity Agenda: An Analysis of the Commission's Entrepreneurship"; Coleman, Liv (University of Tampa). "Pressure on Defense: Cybersecurity and the U.S.-Japan Alliance"; Diersch, Verena (University of Cologne). "Cyberspace as a Realm for Intelligence Cooperation – Is Technology the Driving Factor?"; Griffith, Melissa K. (UC, Berkeley). "The Durability of an Alliance: NATO and Cyber-Defense"; and Yoo, In Tae (Yonsei University). "New Wine into Old Wineskins? Regime Diffusion in Cyberspace through International Trade."

The Cyber Conflict Studies Association (CCSA) promotes and leads international intellectual development efforts to advance the field of cyber conflict research. These activities include workshops that bring together professionals from industry, academia and government to discuss strategic issues surrounding cyber conflict and the publication of insightful research articles and position papers and books. CCSA also plays an important role in our national cyber-readiness strategy, serving as a resource for national security decision-makers and helping to frame and promote national cyber conflict policy. CCSA brings together the best and the brightest individuals in the field of cyber conflict study to further the goals of national security and the field of cyber.

This work was supported in part by the Minerva Research Initiative. The Minerva Research Initiative, administered jointly by the Office of Basic Research and the Office of Policy at the U.S. Department of Defense, supports social science research aimed at improving our basic understanding of security.

Funded in part by the Office Naval Research under the OSD Minerva program, grant number: N00014-17-1-2423.

COLUMBIA | SIPA
School of International and Public Affairs

Exploring the Dimensions of Intelligence in Cyber Conflict

AUTHORS: Dr. Michael Warner and Steven Loleski
SERIES EDITOR: Justin Key Canfil

Introduction

Since its introduction last year, the State of the Field conference has featured a number of analytically distinct but overlapping panel topics addressing issues of concern to the cyber domain. In particular, the panel on Intelligence and Adversaries has addressed the centrality of intelligence to the cyber realm. This year our panel focused on the reciprocal relationship between cyber and intelligence in order to unpack the dynamics between them before more closely zooming in on strategic, operational, and tactical issues relevant to intelligence processes. Our discussion touched on issues covered by other panels namely those covering Strategic, Tactical and Operational dynamics, and Cyber Conflict History.

The inaugural conference last year helpfully laid out some important questions and gaps in the literature in each respective topic area. With respect to the Intelligence field, discussion focused on conceptual matters surrounding intelligence itself along with attribution of cyber attacks. This year there was an effort to build upon this foundation and move beyond a discussion of intelligence processes to explore the broader relationship between cyber and intelligence. In other words, how has the cyber domain impacted intelligence and in turn, how intelligence has affected the cyber realm. While both intelligence in cyber and cyber in intelligence are at work, there were no clear answers to these questions and much work remains to be done specifying these dynamics. Participants raised comparisons to intelligence history in order to better assess the continuities or evolution from the past.

About the State of the Field Series

This article is part of the 2017 Cyber Conflict State of the Field (SOTF) paper series, under the auspices of the Cyber Conflict Studies Association and Columbia University's School of International and Public Affairs.

The conference, held annually since 2016, brings together experts from various academic disciplines, including political science, law, economics, and policy research, to define key questions and map the research frontier in the emerging field of cyber conflict studies. The conference is cumulative: each year builds upon past discussions. As a result, discussions have necessarily matured at different rates as new topics are added.

The papers in this series are meant to capture the findings of the 2017 conference. Together, the papers represent the conference attendees' understanding of the present state of the field in the academic study of cyber conflict.

Big Takeaways from 2017

This year's Intelligence and Adversaries panel continued to develop discussions from last year but endeavoured to frame the topic in a broader context. There were a few overarching themes that were touched upon at various points throughout our panel discussion.

1. ***History matters:*** while popular conceptions of cyber intelligence may dwell on the novelty of this domain, history can inform our understanding of evolutionary changes and continuities. In particular, concerns about the intelligence process and production at the operational and tactical level have clear parallels to Cold War operations. Moreover, exploring the path dependence of the emergence of cyber security largely under the auspices of intelligence organizations has affected how we approach this domain.
2. ***Escalation or restraint?*** There were a number of open questions about the nature of intelligence operations in cyberspace and whether these are (or are perceived) as offensive or defensive in nature. However, there was a consensus that we need to move beyond viewing cyberspace as inherently offense or defense dominant and instead look at changing strategic circumstances that make offense or defensive operations more likely. Toward this end, many participants were aware that intelligence operations can be both defensive and offensive and were concerned with how to credibly signal intentions especially across different nation-states. There was some overlap with other panels that explored strategic, operational, and tactical dynamics exclusively.
3. ***Structuring the discussion:*** there is a tendency to look at the familiar intelligence lifecycle to show how cyber has or has not affected traditional intelligence processes. This year our panel attempted to situate the discussion in broader terms and to be aware of the reciprocal influences between cyber and intelligence along strategic, operational, and tactical levels. However, these levels were not always easy to analytically separate in discussion.

2017 State of the Field: Intelligence and Adversaries

As mentioned, we had structured our discussion first to highlight the nature of cyber intelligence by encouraging participants to think about the mutually entangling dynamics between its constituent parts.

Intelligence in Cyber versus Cyber in Intelligence: It's both![1]

- To what extent does cyber emerge from the history of intelligence?[2]
- How analytically useful is the intelligence life cycle process in conceptualizing the cyber domain? What are some key differences?[3]
- What are the conceptual and organizational boundaries of intelligence and reconnaissance?
- To what extent are intelligence operations defensive or offensive in nature?[4]
- Who are the actors or communities producing intelligence? How do private or non-state actors figure into this discussion?
- What does it mean to collect intelligence through cyber?[5]
- How does intelligence affect cyber writ large or particular targets at the operational or tactical level?
- How do different nation-states approach the cyber domain and how can we establish credible measures of signalling intentions?[6]

Strategic-level considerations

How does cyberspace offer new opportunities for strategic intelligence forecasting? This section considers how cyberspace has affected decision-making by policymakers and in turn how it can be leveraged to provide strategic insight.

- How do decision-makers know what to expect in cyberspace?[7]
- How is cyberspace affecting the practice, organization, and legitimacy of intelligence?[8] How is that affecting citizens and enterprises?

- How is the cyber domain leveraged and used alongside conventional policy domains to achieve strategic objectives?
- How does cyberspace represent an opportunity for intelligence agencies to face strategic surprise a new way?[9]
- How can we leverage private firm reporting in documenting operational policy success or failure? What are the methodological challenges with doing so?
- To what extent will emerging technologies (Artificial Intelligence, machine learning, and quantum computing for example) destabilize or stabilize the cyber domain?

Operational-level considerations

This section covers the role of intelligence in operations and was largely discussed in the U.S. context. Further comparative national study of cyber operational and doctrinal mandates would enrich discussion.[10]

- How can we build and sustain capabilities that meet requisite strategic needs?
- How are cyber capabilities or cyber intelligence assets measured? What metrics can we employ to assess relative power in the cyber domain?
- How does organizational culture matter in the procurement and sustainment of cyber capabilities and intelligence assets?
- How are allies and adversaries planning and operating in cyberspace?[11]
- What is the relationship and role of intelligence in and during cyber operations? How has the intelligence community affected approaching cyber operations?
- To what extent does U.S. law and legal authorizations help or hinder U.S. cyber operations or intelligence gathering?[12]
- How can we think about the cost of cyber operations in achieving policy objectives?[13]

Tactical-level considerations

This section zoomed in to consider some tactical-level considerations on intelligence in cyberspace. Specifically, attribution and vulnerability disclosures seem to be drawing the most attention.

- Attribution! Who does what and to whom?[14]
- How does disclosing vulnerabilities affect intelligence collection?[15]
- Doctrine
  - What does "cyber" do to/for other intelligence disciplines?[16]
  - How are traditional intelligence methods being used in cyber?
  - Who will do all this, and how, and where? Will they know how?

Summary and Recommendations

The opening panel questions provoked a number of interesting comments and further questions pushing the discussion into other areas. One general observation was that there was not much in the way of sustained discussion or consensus about the distinctiveness of the cyber domain on intelligence. A participant in passing noted that the scale, speed, and risk of cyber operations have changed the dynamic but this remains to be explored in some depth. There may even be reasons to challenge these factors given that known cyber espionage campaigns have tended to be long-term operations and it is not altogether clear why or how cyber espionage has been more provocative than past espionage.

Second, a frequent tendency among participants was to conflate strategic, operational, and tactical issues in conversation or it may reflect some ambiguity with how these terms are precisely used with respect to the cyber domain. It may be helpful here to enter into conversation with other panels on strategic, operational, and tactical issues to develop common standards about how these terms are used in the cyber realm. Or this may be a broader issue related to the complexity of cyberspace itself as a domain with emergent properties.[17] For example, the Snowden disclosures were mentioned as an example where a single individual had the capacity to change the strategic conversation on why and how the United States collects foreign intelligence.[18] It is not altogether clear what value added there is analytically by compartmentalizing and reifying the discussion into strategic, operational, and tactical dimensions unless warranted first and foremost by the research question under investigation. With these observations in mind, there are a few potential suggestions that may be worthwhile to consider for the future:

1. ***Encourage puzzle-driven research:*** the hallmark of most established scholarly disciplines is to enter into a puzzle- or problem-solving stage where interesting questions are addressed and middle-range theory develops. Instead of falling back to the comfort of comparing and contrasting cyber to the familiar intelligence process models, it would be useful to encourage a focus on different puzzles or testing middle-range theories. As cyber conflict history continues to be documented, this type of research has the potential to offer complementary case studies and also discuss the methodological challenges with cyber conflict research such as omitted or confounding variables. Also, this will encourage viewing cyber operations alongside more conventional covert operations that may have the benefit of specifying differences or the value added of cyber operations to outcomes. Another research gap identified above is the lack of attention to non-state group given the proliferation of capabilities open to small groups and individuals to conduct cyber espionage or targeted attacks.[19]

2. ***Developing research design and methods:*** social science disciplines in recent years have been moving to increased transparency surround research design and methods. As an emerging field, cyber conflict studies could benefit from a sustained focus on not only research agendas but also research designs and methods best suited to address the problems both unique to this domain and common with other conventional areas. Currently, most discussion of cyber conflict or intelligence remains focused on logical possibilities without sustained testing of those propositions. What needs to be done, following others,[20] is to develop datasets or cases to test and substantiate these claims. This has the potential to move discussion beyond whether cyberspace was inherently offense or defense dominant, which many participants agreed, was becoming exhausted. Having said that, it should be noted that single-case studies combined supplemented with other methods offers great potential value in generating new theoretical insights but also offer rigorous testing of causal mechanisms through within case observations.[21] Case studies may be all the more important in cyber conflict studies given the apparent but not insurmountable methodological problems of developing reliable datasets. Citizen Lab's seminal *Tracking GhostNet* investigation stands out as an example of multi-method work on a single case that yielded considerable insights on the nature of a particular cyber espionage operation.[22]

3. ***Encourage cross-pollination of panels:*** while the stand-alone panels have produced rich discussions in their own right, some observers noticed that certain topics like offense/defense balance were talked about simultaneously with the risk of certain groups talking past each other instead of with each other. It would be useful to encourage related panels to get together where certain issues that seem at odds could potentially be reconciled in a wider cyber operations frame of reference. For example, one observation during our panel was that it was difficult to separate where intelligence ends and an operation begins. As others have noted it quickly becomes clear that discussion bleeds into other panel topics where cyber operations leads to talk about norms and law for instance. Going back to the first suggestion, something to consider more further into the future as the State of the Field develops around common themes and literatures is to encourage paper submissions on cyber conflict writ large and for conference organizers to develop panels based on submissions from the field.

About the Authors

Dr. Michael Warner is the command historian of US Cyber Command and an adjunct faculty member at American University and Johns Hopkins University.

Steven Loleski is a Ph.D. Candidate in the Department of Political Science at the University of Toronto.

End Notes

[1] Michael Warner, “Intelligence in Cyber—and Cyber in Intelligence,” in *Understanding Cyber Conflict: Fourteen Analogies*, ed. George Perkovich and Ariel E. Levite (Georgetown University Press, 2017).

[2] Gordon Corera, *Cyberspies: The Secret History of Surveillance, Hacking, and Digital Espionage*, 1 edition (New York: Pegasus Books, 2016); Fred Kaplan, *Dark Territory: The Secret History of Cyber War* (New York: Simon & Schuster, 2016); Craig Wiener, “Penetrate, Exploit, Disrupt, Destroy: The Rise of Computer Network Operations as a Major Military Innovation” (George Mason University, 2016), http://search.proquest.com/docview/1864633371/.

[3] Jon R. Lindsay, “Cyber Espionage,” in *Oxford Handbook of Cyber Security*, ed. Paul Cornish (Oxford University Press, forthcoming 2018).

[4] Erik Gartzke and Jon R. Lindsay, “Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace,” *Security Studies* 24, no. 2 (April 3, 2015): 316–48; Gary Brown, “Economic Espionage: Spying and Fighting in Cyberspace: What Is Which?,” *J. Nat’l Security L. & Pol’y* 8 (2016): 621–621; Aaron F. Brantly, “Aesop’s Wolves: The Deceptive Appearance of Espionage and Attacks in Cyberspace,” *Intelligence and National Security* 31, no. 5 (July 28, 2016): 674–85, https://doi.org/10.1080/02684527.2015.1077620; Rebecca Slayton, “What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment,” *International Security* 41, no. 3 (January 1, 2017): 72–109.

[5] Intelligence and National Security Alliance (INSA), “Cyber Intelligence: Setting the Landscape For An Emerging Discipline” (Intelligence and National Security Alliance (INSA), 2011), https://www.insaonline.org/wp-content/uploads/2017/04/INSA_CyberIntel_WP.pdf.

[6] Jon R. Lindsay, “Introduction—China and Cybersecurity: Controversy and Context,” in *China and Cybersecurity: Espionage, Strategy, and Politics in the Digital Domain*, ed. Tai Ming. Cheung, Derek S. Reveron, and Jon R. Lindsay ([S.l.]: Oxford University Press, 2015), 1–28; Mark Galeotti, *Putin’s Hydra: Inside Russia’s Intelligence Services* (European Council on Foreign Relations, 2016), http://www.ecfr.eu/page/-/ECFR_169_-_INSIDE_RUSSIAS_INTELLIGENCE_SERVICES_(WEB_AND_PRINT)_2.pdf.

[7] Aaron Franklin Brantly, *The Decision to Attack: Military and Intelligence Cyber Decision-Making*, First Edition first Printing edition (Athens, GA: University of Georgia Press, 2016).

[8] Michael V. Hayden, *Playing to the Edge: American Intelligence in the Age of Terror* (New York: Penguin Press, 2016); Jennifer Stisa Granick, *American Spies: Modern Surveillance, Why You Should Care, and What to Do about It* (Cambridge, United Kingdom: Cambridge University Press, 2017).

[9] Richard K. Betts, *Enemies of Intelligence : Knowledge and Power in American National Security* (New York: Columbia University Press, 2007); Daniel Byman, “Strategic Surprise and the September 11 Attacks,” *Annual Review of Political Science* 8 (2005): 145–70.

[10] “Cybersecurity and Cyberwarfare: National Doctrine and Organization,” accessed March 27, 2018, http://stefanomele.it/news/dettaglio.asp?id=275.

[11] Austin Long, “A Cyber SIOP? Operational Considerations for Strategic Offensive Cyber Planning,” *Journal of Cybersecurity* 3, no. 1 (March 1, 2017): 19–28.

[12] Andru E. Wall, “Demystifying the Title 10-Title 50 Debate: Distinguishing Military Operations, Intelligence Activities & Covert Action,” *Harv. Nat’l Sec. J.* 3 (2011): 85.

[13] Slayton, “What Is the Cyber Offense-Defense Balance?”

[14] Thomas Rid and Ben Buchanan, “Attributing Cyber Attacks,” *Journal of Strategic Studies* 38, no. 1–2 (January 2, 2015): 4–37.

[15] Jason Healey, “The U.S. Government and Zero-Day Vulnerabilities: From Pre-Heartbleed to Shadow Brokers,” JIA SIPA, November 1, 2016, https://jia.sipa.columbia.edu/online-articles/healey_vulnerability_equities_process.

[16] Michael Warner, “Reflections on Technology and Intelligence Systems,” *Intelligence and National Security* 27, no. 1 (February 2012): 133–53; Mark M. Lowenthal and Robert M. Clark, eds., *The Five Disciplines of Intelligence Collection*, 1 edition (CQ Press, 2015).

[17] Robert Jervis, *System Effects: Complexity in Political and Social Life* (Princeton, NJ: Princeton University Press, 1997).

[18] “Presidential Policy Directive -- Signals Intelligence Activities,” whitehouse.gov, January 17, 2014, https://obamawhitehouse.archives.gov/the-press-office/2014/01/17/presidential-policy-directive-signals-intelligence-activities.

[19] Cortney Weinbaum, Steven Berner, and Bruce McClintock, “SIGINT for Anyone,” Product Page, 2017, https://www.rand.org/pubs/perspectives/PE273.html; Marcos Degaut, “Spies and Policymakers: Intelligence in the Information Age,” *Intelligence and National Security* 31, no. 4 (June 6, 2016): 509–31; Ronald J. Deibert et al., “Tracking Ghostnet: Investigating a Cyber Espionage Network,” 2009.

[20] Jason Healey, ed., *A Fierce Domain: Conflict in Cyberspace*, 1986 to 2012 (Vienna, VA: Cyber Conflict Studies Association, 2013); Jon R. Lindsay, “Stuxnet and the Limits of Cyber Warfare,” *Security Studies* 22, no. 3 (July 2013): 365–404.

[21] Alexander L. George and Andrew Bennett, *Case Studies and Theory Development in the Social Sciences* (MIT Press, 2005); Wiener, “Penetrate, Exploit, Disrupt, Destroy.”

[22] Deibert et al., “Tracking Ghostnet.”

Bibliography

Betts, Richard K. *Enemies of Intelligence : Knowledge and Power in American National Security*. New York: Columbia University Press, 2007. http://www.loc.gov/catdir/toc/ecip0710/2007003937.html.

Brantly, Aaron F. “Aesop’s Wolves: The Deceptive Appearance of Espionage and Attacks in Cyberspace.” *Intelligence and National Security* 31, no. 5 (July 28, 2016): 674–85. https://doi.org/10.1080/02684527.2015.1077620.

Brantly, Aaron Franklin. *The Decision to Attack: Military and Intelligence Cyber Decision-Making*. First Edition first Printing edition. Athens, GA: University of Georgia Press, 2016.

Brown, Gary. “Economic Espionage: Spying and Fighting in Cyberspace: What Is Which?” *J. Nat’l Security L. & Pol’y* 8 (2016): 621–621.

Byman, Daniel. “Strategic Surprise and the September 11 Attacks.” *Annual Review of Political Science* 8 (2005): 145–70. https://doi.org/10.1146/annurev.polisci.8.082103.104927.

Corera, Gordon. *Cyberspies: The Secret History of Surveillance, Hacking, and Digital Espionage*. 1 edition. New York: Pegasus Books, 2016.

“Cybersecurity and Cyberwarfare: National Doctrine and Organization.” Accessed March 27, 2018. http://stefanomele.it/news/dettaglio.asp?id=275.

Degaut, Marcos. “Spies and Policymakers: Intelligence in the Information Age.” *Intelligence and National Security* 31, no. 4 (June 6, 2016): 509–31. https://doi.org/10.1080/02684527.2015.1017931.

Deibert, Ronald J., Rafal Rohozinski, A. Manchanda, Nart Villeneuve, and G. M. F. Walton. “Tracking Ghostnet: Investigating a Cyber Espionage Network,” 2009.

Galeotti, Mark. *Putin’s Hydra: Inside Russia’s Intelligence Services*. European Council on Foreign Relations, 2016. http://www.ecfr.eu/page/-/ECFR_169_-_INSIDE_RUSSIAS_INTELLIGENCE_SERVICES_(WEB_AND_PRINT)_2.pdf.

Gartzke, Erik, and Jon R. Lindsay. “Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace.” *Security Studies* 24, no. 2 (April 3, 2015): 316–48. https://doi.org/10.1080/09636412.2015.1038188.

George, Alexander L., and Andrew Bennett. *Case Studies and Theory Development in the Social Sciences*. MIT Press, 2005.

Granick, Jennifer Stisa. *American Spies: Modern Surveillance, Why You Should Care, and What to Do about It*. Cambridge, United Kingdom: Cambridge University Press, 2017.

Hayden, Michael V. *Playing to the Edge: American Intelligence in the Age of Terror*. New York: Penguin Press, 2016.

Healey, Jason, ed. *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012*. Vienna, VA: Cyber Conflict Studies Association, 2013.

———. “The U.S. Government and Zero-Day Vulnerabilities: From Pre-Heartbleed to Shadow Brokers.” JIA SIPA, November 1, 2016. https://jia.sipa.columbia.edu/online-articles/healey_vulnerability_equities_process.

Intelligence and National Security Alliance (INSA). “Cyber Intelligence: Setting the Landscape For An Emerging Discipline.” Intelligence and National Security Alliance (INSA), 2011. https://www.insaonline.org/wp-content/uploads/2017/04/INSA_CyberIntel_WP.pdf.

Jervis, Robert. *System Effects: Complexity in Political and Social Life*. Princeton, NJ: Princeton University Press, 1997.

Kaplan, Fred. *Dark Territory: The Secret History of Cyber War*. New York: Simon & Schuster, 2016.

Lindsay, Jon R. “Cyber Espionage.” In *Oxford Handbook of Cyber Security*, edited by Paul Cornish. Oxford University Press, 2018.

———. “Introduction—China and Cybersecurity: Controversy and Context.” In *China and Cybersecurity: Espionage, Strategy, and Politics in the Digital Domain*, edited by Tai Ming. Cheung, Derek S. Reveron, and Jon R. Lindsay, 1–28. [S.l.]: Oxford University Press, 2015. http://myaccess.library.utoronto.ca/login?url=http://books.scholarsportal.info/viewdoc.html?id=/ebooks/ebooks3/oso/2015-04-27/1/9780190201265-Lindsay.

———. “Stuxnet and the Limits of Cyber Warfare.” *Security Studies* 22, no. 3 (July 2013): 365–404. https://doi.org/10.1080/09636412.2013.816122.

Long, Austin. “A Cyber SIOP? Operational Considerations for Strategic Offensive Cyber Planning.” *Journal of Cybersecurity* 3, no. 1 (March 1, 2017): 19–28. https://doi.org/10.1093/cybsec/tyw016.

Lowenthal, Mark M., and Robert M. Clark, eds. *The Five Disciplines of Intelligence Collection*. 1 edition. CQ Press, 2015.

“Presidential Policy Directive—Signals Intelligence Activities.” whitehouse.gov, January 17, 2014. https://obamawhitehouse.archives.gov/the-press-office/2014/01/17/presidential-policy-directive-signals-intelligence-activities.

Rid, Thomas, and Ben Buchanan. “Attributing Cyber Attacks.” *Journal of Strategic Studies* 38, no. 1–2 (January 2, 2015): 4–37. https://doi.org/10.1080/01402390.2014.977382.

Slayton, Rebecca. “What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment.” *International Security* 41, no. 3 (January 1, 2017): 72–109. https://doi.org/10.1162/ISEC_a_00267.

Wall, Andru E. “Demystifying the Title 10-Title 50 Debate: Distinguishing Military Operations, Intelligence Activities & Covert Action.” *Harv. Nat’l Sec. J.* 3 (2011): 85.

Warner, Michael. “Intelligence in Cyber—and Cyber in Intelligence.” In *Understanding Cyber Conflict: Fourteen Analogies*, edited by George Perkovich and Ariel E. Levite. Georgetown University Press, 2017.

———. “Reflections on Technology and Intelligence Systems.” *Intelligence and National Security* 27, no. 1 (February 2012): 133–53. https://doi.org/10.1080/02684527.2012.621604.

Weinbaum, Cortney, Steven Berner, and Bruce McClintock. “SIGINT for Anyone.” Product Page, 2017. https://www.rand.org/pubs/perspectives/PE273.html.

Wiener, Craig. “Penetrate, Exploit, Disrupt, Destroy: The Rise of Computer Network Operations as a Major Military Innovation.” George Mason University, 2016. http://search.proquest.com/docview/1864633371/.

The Cyber Conflict Studies Association (CCSA) promotes and leads international intellectual development efforts to advance the field of cyber conflict research. These activities include workshops that bring together professionals from industry, academia and government to discuss strategic issues surrounding cyber conflict and the publication of insightful research articles and position papers and books. CCSA also plays an important role in our national cyber-readiness strategy, serving as a resource for national security decision-makers and helping to frame and promote national cyber conflict policy. CCSA brings together the best and the brightest individuals in the field of cyber conflict study to further the goals of national security and the field of cyber.

This work was supported in part by the Minerva Research Initiative. The Minerva Research Initiative, administered jointly by the Office of Basic Research and the Office of Policy at the U.S. Department of Defense, supports social science research aimed at improving our basic understanding of security.

Funded in part by the Office Naval Research under the OSD Minerva program, grant number: N00014-17-1-2423.

COLUMBIA | SIPA
School of International and Public Affairs

Economic Dimensions of Cyber Conflict

AUTHORS: Dr. Chris Demchak and Benjamin Schechter
SERIES EDITOR: Justin Key Canfil

Introduction

As the relevance and academic interest in cyber conflict studies deepen and the methods and theory applied to the subject matter multiply, it is evitable that the relevance of cybersecurity studies to other academic areas should also expand. The increasing linkages between cyber conflict studies and other established disciplines mirror cyber's deepening integration in diverse areas ranging from communications and finance to healthcare. The rise of highly insecure and highly interconnected globe-spanning digital networks has not only affected existing systems like finance but also radically redefine them. The question is raised whether cyber's impacts force modifications into existing theory and thought across other academic areas, or perhaps even radical reassessments. More to the point of this report, while cyber-driven interconnectedness spreads, so do the threats that accompany the use of cyber. These threats present serious, unanswered quandaries for some academic disciplines, such as economics.

This year was the first State of the Field to feature the Economic Impact of Cyber Threats panel. The panel explored how cyber insecurity impacts economies at the firm and state level, and how that may have implications for existing economic theory. The panel discussions investigated how existing economic theory does, or does not, account for increasingly severe cyber threats and systemic cyber vulnerability. This panel was intended to mark the beginning of an ongoing discussion, laying a groundwork for more substantive work.

About the State of the Field Series

This article is part of the 2017 Cyber Conflict State of the Field (SOTF) paper series, under the auspices of the Cyber Conflict Studies Association and Columbia University's School of International and Public Affairs.

The conference, held annually since 2016, brings together experts from various academic disciplines, including political science, law, economics, and policy research, to define key questions and map the research frontier in the emerging field of cyber conflict studies. The conference is cumulative: each year builds upon past discussions. As a result, discussions have necessarily matured at different rates as new topics are added.

The papers in this series are meant to capture the findings of the 2017 conference. Together, the papers represent the conference attendees' understanding of the present state of the field in the academic study of cyber conflict.

Discussions touched on a host of topics, related to the effects of a highly integrated world, deeply reliant on cyberinfrastructure. There was a consensus among discussants that the impacts of cyber threats on national and international economics were consequential and had security implications. Furthermore, there was a general consensus that this problem set had received insufficient scrutiny and study by established scholars. The discussion had a number of relevant takeaways. As was expected, there is still pushback on whether the degree to which this research is necessary or if the problem is being framed is correct. Is it a fundamental issue at the theoretical level or if it is evidence of insufficient application of existing theory to the problem. Issues of how to effectively conduct the research and how to make this line of inquiry relevant and viable for scholars were also central themes. Over the course of the panel, three key knowledge gaps were highlighted: the absence of effective theory linking cybersecurity to economic realities, the lack of relevant, reliable, or exploitable largescale data or data collection to effectively enable this kind of research, and the missing participation of economics scholars in cyber conflict debates.

Based on discussions there are no clear canonical works in this area; rather, there are works that could inform this type of research and the development of an authoritative body of academic literature. This review presents the three threads with associated questions raised in the discussion.

Creation of the Economic Impact of Cyber Threats Panel

This year was the first to feature a dedicated panel to discuss the challenges of cyber insecurity to economics. However, the panel was created in response to the scarcity of cohesive thought on the challenges of cyber insecurity and vulnerability to economies and societies. The objective of this panel was established to explore the potential challenges to existing economic theory and how to correct any theory shortcomings.

Takeaways from 2017

The inaugural Economic Impact of Cyber Threats discussed far-ranging topics and concepts, leveraging the diverse participants of the Cyber Conflict Studies Association. The breadth of topics discussed indicated a willingness to conceptualize of the panel's topic through a variety of methodological and theoretical lenses. However, this also meant discourse remained largely at higher, conceptual levels. Nonetheless, three major clusters of interest emerged through the panel discussion.

1. The first discussion area was oriented around the validity of the panel's core question regarding economic theory. These concerns fell into two broad categories. The first is that the problem is a fleeting or outside the scope of cyber conflict studies. Specifically, current challenges will either be resolved as cybersecurity becomes normalized or as status quo challengers. Countries like China will become more established and less inclined to tolerate or engage in malicious cyber activity. The second was that the topic was inherently within the realm of economics. Some argued that cyber threats to economies would be managed and treated like any other form of risk and managed accordingly and that this was not an issue of theory but time, that equilibrium would be restored. Others argued that these cyber challenges to specifically neoclassical economic models may be sufficiently critical and profound that they require a critical reassessment of existing foundational economic theory.

2. The second discussion area was how to effectively investigate the issues of cyber insecurity's effect on economies and what would be necessary to develop new economic theory, if necessary. These concerns followed one of three threads: definitions, data, and /or methodology. There was agreement that there needed to be a clearer, more precise use of terms when engaging in multidisciplinary research, as this would entail. Even during the discussion, there was debate over terms of reference, such as those associated with economic or accounting loss, among others. This highlighted the ongoing and well-known challenges of adopting a common terminology. Data was the largest issue, acknowledging that any research would require relevant and trusted data, which is extremely difficult to obtain given the covertness of cybersecurity operations generally and institutional reluctance to reveal information on breaches or other losses. Furthermore, establishing a collection regime, even with cooperative and trusted states and firms, is challenging. There is no consensus on what data is needed, the best methods for collecting it, and how to know if some critical data has been overlooked. There were questions on how to effectively and predictably quantify things like intellectual property theft. Finally, there were concerns raised about the methodology, establishing causality in areas as complex as national economies and cybersecurity, and most critically, what economic theory might be suspect and in what ways.

3. The final area of discussion was how cyber conflict studies could advance economic theory and role for future scholars, both in what types of

research could be done and how to make this line of research viable for academics, both established and emerging. A part of this was how to draw support and interest from economists to engage with the challenges of cyber conflict scholars.

2017 State of the Field: Economic Impact of Cyber Threats

Lack of Theorizing

While the economic impacts of cyber insecurity are becoming increasingly clear, there has been scant progress in developing theories to explain what the consequences this cyber insecurity is to national economies and systems. Cyberspace has facilitated global interconnectedness and revolutionized a host of fields, as well as giving rise to cyber conflict studies. However, there have been insufficient efforts to develop theories that span multiple disciplines in the same way that the impacts of cyberspace have changed the world they study. Economic theories, international relations theories, and theories of warfare need to be better integrated into the cybersecurity debates and vice versa. Emerging powers, such as China, and non-state actors utilize cyberspace in unintended ways that directly affect areas of interest to a range of academic disciplines. Despite the changing status quo, the guiding frameworks across these disciplines of studies continue to be generally fragmented in presumptions, logic, explanations, and policy conclusions.

Key Questions

- To what extent can—or should—cyber studies or cyber conflict studies be isolated from other fields?
- How do the economic definitions need to be adjusted or redefined for a cybered world and what then happens to related theories' applicability to the emerging world?
- What do the standard economic terms mean in the context of cyber conflict, such as—but not limited to—cost, value, rationality, harm, information, trust, tradeoffs, concepts of utility (marginal, optimal, etc.), and comparative/absolute advantage?
- What new theories or theoretical adjustments are implied if the fundamental (and highly westernized) rule of law—taken for granted in modern economic models—no longer applies or is not uniformly applied/enforced in a deeply cybered world dominated by nonwestern and more authoritarian states?
- What gaps in current models need to be filled, or new theories developed to deal with observed and rising volume of theoretically excluded or unaccommodated—and largely unrestricted—behaviors in cyberspace and cyber conflict? These behaviors include, but are not limited to: zero marginal cost industries thriving and able to charge non-zero prices, economic development advances through theft of intellectual property, the rise of cyber national champions, economic coercion of large and small enterprises by hostile cyber actors, the rationality of risk calculations by criminal non-state actors (both groups and individuals) in cyberspace, large corporations experiencing massive breaches and loss of IP capital , and states not exercising effective governance over the IT capital goods sector.
- How does cyberspace as an unprecedented substrate underlying both democratic and authoritarian societies challenge existing economic and conflict models? To what extent does it break, bend, or make irrelevant current theories? In what way and through what mechanisms do these challenges express themselves, from its scale globally, its speed in complex system interactions, its opaqueness in basic structures, to its enhancement of global system sensitivity to large coherent and aggressive actors seeking dominance over a few critical sectors like telecommunications?
- How can theories of future states be crafted from current economic models? How can those models help inform national leaders and policymakers who are struggling to keep pace with rapid technological advancements? How can theory help us prepare for destabilizing, disruptive, or destructive cyber-accelerated systemic surprises in the future?

Lack of Data

There are no norms, standards, professional sharing practices, or regulations that provide reliable and large enough scale macro- and micro- economic and cybersecurity datasets suitable for economic analysis. These challenges apply to other data scientists and even quantitative social science researchers as well. Mandatory reporting could help, but as yet, it is not clear what variables and data, and at what granularity, is needed to study the intersection of cyber conflict and economics properly. Questions that are not easily quantified—such as economic resilience to cyber threats or the effects of cyber on societal resilience will not necessarily be attractive to economists or viable for statistical analysis.

Key Questions

- What data forms, collection, variables, and reliability are required and achievable?
- How does one establish the wider systemic implications of cyber insecurity and economics without data? Are there credible and verifiable surrogates—reliably available over time—that may be used if the preferred data is unobtainable?
- What are the usual and unusual sources of useable and reliable data across nations, cultures, and use cases?
- How can measures such as losses in GDP terms be made meaningful—and therefore theoretically informing—with additional context and verified methods of collection and analysis?
- How can concrete cases of significant economic gains from economic distortions, such as the Huawei IP theft leading to the bankruptcy of Nortel—or the WannaCry/Petya ransomware—be more rigorously investigated?
- How can cyber incidents of varying magnitudes, types, and targets be integrated to understand the larger effects on economies? For example, given the Chinese role in cyber extractions and its meteoric national rise in global economic influence, to what extent is a new model of economic growth necessary, one suitable for a deeply cybered world?
- What data will help disentangle the benefits of cyber from other economic benefits and behaviors across regions, cultures, and demonstrated governance preferences?
- What data—and from where—is necessary to explore if new models, language, rules, and trend indices of international trade are necessary to capture changing global economies and markets accurately?
- What emerging computational or statistical methods might be modified to help address the data challenges, such as gathering non-traditional digital signatures information or massive dataset analysis?

Lack of Neoclassical and Political Economists in Cyber Conflict Studies

Cyberspace has taken root globally and become a common substrate to developed and developing economies and societies. However, neoclassical economists and quantitative political-economy scholars are absent in discussions of rising interstate conflict, massive illicit wealth transfers, changes in trade dynamics, and other systemic evolutions and surprises. Scholars, especially junior scholars, have begun to question the standing theoretical assumptions and conclusions in fields ranging from international relations, security studies, and psychology to systems engineering, and even the information technology disciplines. However, economists have shown less interest or initiative in questioning their theoretical foundations. The few who are challenging the current models—largely built in the Cold War era—are those with Nobel Economics awards like Robert Shiller or a small group of critical economics theorists researching what they call "real world economics."[1] Even those challenging the models are not, however, integrating cyber in their critiques or new theories. Despite massively disruptive cyber incidents, effects on companies and markets, and other cybersecurity costs imposed on citizens, societies, and increasingly antiquated laws intended to keep markets transparent and fair; research remains scant.

Why aren't economists interested in the economic impacts of cyber threats? One argument has been the lack of data discourages the ease of using the quantitative tools now required in academically credible economic analysis. Another possibility is each subfield of economics viewing cyber's impacts as another field's problems. For example, massive intellectual property theft viewed as a firm-level problem only for micro-economists, or as a problem only for trade or macro-economists in that nations need better regulation and should use WTO mechanisms to arbitrate the costs and punishments. Leaving aside cyber, these three main fields in economics already are not integrated into approaches, and each is siloed conceptually, accepting for purposes of simplification the theoretical assumptions of the other two fields as given in the background of research. There is no incentive to include a systemic variable such as cyberspace which could be viewed as relevant to all three subfields.

Unfortunately, cyberspace does reach systemically across all varying levels of analysis in existing economics, micro, meso, macro, and trade. Economic models depend on simplification that may minimize distortions caused by systemic cyber vulnerabilities. The theories externalize systemic and background stabilization to governments and assume the norms of democratic civil society—from the assured value of currency to legal protection of contracts to policing of theft—are in full operation universally. The future also poses a problem in attracting economists to this field since they do not have tools to model the cyber in the future world. Cyberspace, as it has evolved, is creating even greater disparities between current models and theories and the emerging and conflictual surrounding reality.

Key Questions

- Why aren't existing economic assumptions being reassessed, as they are in many other social sciences?
- How can job opportunities exist for scholars looking at cyber economics at all, let alone the combination of cybersecurity, economics, and conflict?
- Does it hurt or help in attracting economic scholars to these questions if cyber economics is viewed as a distinct field, and how can the question of threats and conflict be included?
- What key terms—such as market failure and economic rationality—are particularly challenged in this space and how can the normal tools of neoclassical economists help them be sufficiently aware of these systemic changes?
- What is needed to have the political-economist become less focused on the political and more on the economic challenges and cyber in order to become the translating field of study in this area?
- What is the argument for and against having a separate subfield of cyberpolitical-economy in order to produce both integrating theories and large data collection necessary to fill the gaps in this space?

Summary and Recommendations

The key observation from this panel is that economics, while critical to the study of cybersecurity and cyber conflict, is woefully understudied by the scholars in economics. Making the case to that audience in particular about filling gaps in theories, data, and researchers is essential. There remains much to be done.

In the interim, recommendations for action include investigating potentially overlooked existing models whose data, methods, and theories could be adapted for use. Examples include information theories, socio-technical systems and surprise research, new forms of accounting (holistic), and works on normative synthesis. These and others could be drawn into use in this field to encourage openness to help acquire the missing data. The data problem will also require both an evaluation of the types of modeling done and what we consider viable data. As mentioned, there is a range of discipline that utilizes quantitative methods. Those fields also utilize diverse, and sometimes disparate, forms and types of data. Cyberspace is a digital space and is an abundant source of data, although not always in easily exploited forms or of obvious utility. However, using novel methods and models that abundance of data can be leveraged. Already we have seen the application of automation to collect, clean, and assemble massive datasets. Coupled with big data methods and analytics meaning can be derived to even apparently useless data. Assuming right question is being asked in the right way. This is just one example of an opportunity to derive important academic insights from novel methods and data sources.

Additionally, the literature on technological diffusion, corporate ventures, and cyber operations could be integrated to understand the role of cyber threats in changing inter-state relations. Making a case for how critical economics is to the conflict in cyberspace could be helped by investigating literature on economic coercion, arsenal democracies, gains from invasions physically and (now) digitally, and engineering lessons on creating systemic reliability even when combining unreliable systems. To help incorporate the work of other academic disciplines into cyber conflict studies will require both a willingness to accept existing lexicons, but also the willingness to adopt baseline terms and theoretical constructs. It is challenging to involve academics with diverse research backgrounds, such as economists, into the cyber field when their lexicon is inaccurately used and when the cyber lexicon is always in flux.

The bottom line is that cybersecurity scholars need to reach out and persuade economists to vigorously engage in the rising challenges of a deeply conflictual cybered world. That persuasion will take a while, and there are many urgent questions to be answered across this space. There are other fields that also need to be considered, and some may offer complementary data and explanations that lead to unexpected and supportive discoveries for all the disciplines involved.

About the Authors

Chris C. Demchak, Ph.D., is the RDML Grace M. Hopper Professor of Cyber Security and Director, Center for Cyber Conflict Studies (C3S), U.S. Naval War College.

Benjamin Schechter is a research associate at the Center for Cyber Conflict Studies(C3S), U.S. Naval War College.

End Note

[1] See for example the Real World Economics Review website, www.worldeconomicsassociation.org/journals/rwer/

Relevant Literature

While limited academic attention has been given to the issues raised in this panel, there is a relevant body of literature that can help inform future research into these issues. The literature presented here is not exhaustive but represents the first steps into exploring the systemic effects of cyber insecurity on economic theory.

Relevant contributions to Theorizing Gap

Holling, C. S. (2001). "Understanding the Complexity of Economic, Ecological, and Social Systems." *Ecosystems* 4(5): 390-405.

Parsons, T. and N. Smelser (1998). *Economy and society: A study in the integration of economic and social theory*, Routledge.

Baldwin, D. A. (1985). *Economic statecraft*, Princeton University Press.

Kahn, A. E. (1966). "The tyranny of small decisions: Market failures, imperfections, and the limits of economics." *Kyklos* 19(1): 23-47.

West, G. (2017). *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*. London, Orion Press.

Wang, Z. (2017). "The Economic Rise of China: Rule-Taker, Rule-Maker, or Rule-Breaker?" *Asian Survey* 57(4): 595-617.

Bloom, N., et al. (2013). "A Trapped-Factors Model of Innovation." *The American Economic Review* 103(3): 208-213.

Roe, E. (2012). *Taking complexity seriously: policy analysis, triangulation and sustainable development*, Springer Science & Business Media.

Relevant contributions to Data Identification and Acquisition Gap

Li, Z., et al. "Botnet economics: uncertainty matters." *Managing Information Risk and the Economics of Security*: 245-267.

Hannas, W. C., et al. (2013). *Chinese industrial espionage: Technology acquisition and military modernisation*, Routledge.

Relevant contributions to Missing Incentives for Economists Gap

Kakerlof, G. A. and R. J. Shiller (2015). *Phishing for phools: The economics of manipulation and deception*, Princeton University Press.

Keen, S. (2011). *Debunking Economics: the naked emperor dethroned?*, Zed Books Ltd.

Mishan, E. J. (2011 (1986)). *Economic Myths and the Mythology of Economics* (Routledge Revivals), Routledge.

Blanchard, O., et al. (2012). *In the wake of the crisis: Leading economists reassess economic policy*, MIT Press.
Akerlof, G. A., et al. (2014). *What Have We Learned?: Macroeconomic Policy After the Crisis*, MIT Press.

Romer, P. M. (2015). "Mathiness in the Theory of Economic Growth." *The American Economic Review* 105(5): 89.

The Cyber Conflict Studies Association (CCSA) promotes and leads international intellectual development efforts to advance the field of cyber conflict research. These activities include workshops that bring together professionals from industry, academia and government to discuss strategic issues surrounding cyber conflict and the publication of insightful research articles and position papers and books. CCSA also plays an important role in our national cyber-readiness strategy, serving as a resource for national security decision-makers and helping to frame and promote national cyber conflict policy. CCSA brings together the best and the brightest individuals in the field of cyber conflict study to further the goals of national security and the field of cyber.

This work was supported in part by the Minerva Research Initiative. The Minerva Research Initiative, administered jointly by the Office of Basic Research and the Office of Policy at the U.S. Department of Defense, supports social science research aimed at improving our basic understanding of security.

Funded in part by the Office Naval Research under the OSD Minerva program, grant number: N00014-17-1-2423.

COLUMBIA | SIPA
School of International and Public Affairs

CCSA
Cyber Conflict Studies Association

The International Law of Cyber Conflict

AUTHOR AND SERIES EDITOR: Justin Key Canfil
EXECUTIVE EDITOR: Jason Healey

Introduction

Is there international law to govern cyberspace and/or cyberspace operations? For years, this has been the million-dollar question. There is growing consensus that, far from the "wild west" it is often depicted as, cyberspace is indeed subject to extant legal and normative regimes. However, agreement on precisely which rules apply has proven elusive, and important issues remain unsettled. Participants in the 2017 State of the Field (SOTF) conference explored these issues through three main questions: (1) has progress been made in recent international legal discourse or diplomacy?; (2) does proposed theory reflect the reality of state practice?; and (3) what are the most important emerging issues?

The 2017 SOTF conference builds on findings from the 2016 SOTF conference, as well as several subsequent real-world developments. It first engages in a brief review of the findings of the 2016 panel. Next, it examines whether any theoretical progress has been made in the intervening period. It then turns to the question of whether theory accords with observed state practice. The years 2015 through 2017 saw numerous attempts by the international community to iron out consensus on the most pressing cyber law issues, but how closely do the fruits of these efforts mirror the claims of theorists and advocates? Finally, the report concludes by calling attention to several critical emerging issues and recommendations for future research areas.

About the State of the Field Series

This article is part of the 2017 Cyber Conflict State of the Field (SOTF) paper series, under the auspices of the Cyber Conflict Studies Association and Columbia University's School of International and Public Affairs.

The conference, held annually since 2016, brings together experts from various academic disciplines, including political science, law, economics, and policy research, to define key questions and map the research frontier in the emerging field of cyber conflict studies. The conference is cumulative: each year builds upon past discussions. As a result, discussions have necessarily matured at different rates as new topics are added.

The papers in this series are meant to capture the findings of the 2017 conference. Together, the papers represent the conference attendees' understanding of the present state of the field in the academic study of cyber conflict.

2016 Review

The 2016 SOTF Law and Ethics discussion broached three topics: *jus ad bellum*, the law governing the rights of states to resort to uses of force; *jus in bello*, the law that governs within armed conflict; and general issues involving sovereignty and neutrality. Leaning heavily on the 2013 *Tallinn Manual on the International Law Applicable to Cyber Warfare*, the 2016 working group's core premise was that, although legal ambiguity and challenges in enforcement remain, there has been a general trend towards consensus that international law does govern cyber operations, and on the application of some specific, fundamental principles of international law to the cyber domain. Lawyers and other interested stakeholders have undertaken the task of interpreting precedent [. . .]"in the hope that presenting a coherent legal framework will encourage states to embrace certain norms."[1,2] It is from that premise that this report proceeds, by examining more recent developments for evidence that international opinion is converging on a particular set of solutions.

The most important function of the 2016 report was to serve as a compilation of a body of canonical literature on the topic of cyber law. As the text notes, "rarely is a single author, work, or school of thought the starting point for a legal analysis in the manner that those might be in, for example, the field of international relations."[3] The usual canonical sources in international jurisprudence—treaty law, customary international law, and reference points such as landmark International Court of Justice (ICJ) precedent—largely arose before the advent of the cyber domain. Thus, given the proposition that extant legal regimes apply to cyberspace, panelists at the 2016 conference rightly recognized that importing such sources from across domains threatened to inject the discussion with an implicit tautology. Instead, "canonical" was taken to mean "works that have made a *direct* contribution to the relatively young field of the law governing cyber conflict. With a few notable exceptions, most are scholarly works," such as the *Tallinn Manual* and earlier writings.[4] The group saw the *Tallinn Manual*'s findings as *de rigeur*, especially on issues of *jus ad bellum*. However, several unanswered questions remained at the time of drafting:

- On cyber conflict below the threshold of *jus ad bellum*: how can retorsion and countermeasures be meaningfully distinguished? Which legal framework is ideal for sub-threshold incidents?

- On *jus in bello*: although international humanitarian law precepts apply, what can be done if states reject their application to cyber incidents? Note that this is a problem across operational domains, heightened by cyberspace characteristics. What types of persons or infrastructure qualify for protected status? How should cyber infrastructure be classified or distinguished?

- On sovereignty and neutrality: when does a third-party state forfeit neutrality? For the conduct of non-state actors, which test—effective versus overall control—is more sensible? What responsibility does the state have for the private sector?

On these and other questions, the 2016 SOTF workshop looked forward to further exploration in the relatively young body of scholarship dedicated to cyber law. However, participants cautioned that "those working on cyber conflict should approach the law with an appreciation for its inherent uncertainty. The emergent field of cyber conflict law is highly dependent on interpretation and implementation. Each new international cyber incident presents the potential for upending existing assumptions."[5] It was with that caveat in mind that the 2017 workshop participants sought to examine progress, developments in the relationship between theory and practice, and key emerging issues.

2017 Takeaways

The law and ethics panel arrived at several conclusions and broached several topics for which answers are not yet clear. This particular constellation of participants reminded the group of the utility of thinking about law in "two separate buckets": domestic and international. Thus far, cyber conflict discussions both in the literature and at SOTF Conferences have primarily concerned themselves with questions of public international law. Domestic law, especially the Title 10/ Title 50 debate concerning the blended role of military and intelligence operators in cyberspace under U.S. federal law, has often been overlooked by scholars.

Participants also discussed whether the "use of force" debate in the legal literature, which seeks to apply *jus ad bellum* concepts to cyber operations, is the most fruitful avenue for discourse. Several participants complained that, although the issue has been debated for over 20 years, all scholars have succeeded in doing is to clarify the questions. Furthermore, participants insisted that some operations will never fit into the use of force construct. Given this limitation, participants considered alternate paradigms for regulating operations below the "armed attack" threshold that triggers a state's inherent right to self-defense, including environmental law, public health law, and other *lex specialis* models that might make suitable analogs. They also considered whether sovereign noninterference might be a superior way to frame the issue.

Until states act, theories remain abstract. To solve many cyber problems, the world needs to build norms. However, as working group participants highlighted, many cyber characteristics—such as being clandestine and multi-stakeholder—heighten the challenges of norm development. Also at issue is the question of who gets a seat at the table. A greater number of seats would mean more buy-in from the international community, but at the cost of impeding consensus and diffusing control over outcomes. Beyond the number of seats is the question of who should populate those seats. Norms resultant from any working group directly reflect the interests of those shaping the norms. For example, if the great cyber power shape the norms, it is like those norms will best support the interests of the great powers at the costs or tertiary cyber actors.

The working group also tackled the important task of clarifying terms and definitions, specifically for the concepts of "norms" and "operations" in cyberspace. This exercise revealed a high degree of heterogeneity in the thinking of even a relatively homogenous (mostly American) group. Perhaps the most important outcome was the reiteration that norms are what states do, not what they expect. Not every offense is a "violation." The group then had a deeper debate over who it spoke for: Global citizens? International legal scholars? The United States? The working group alone? The group members as individuals? As this discussion made apparent, participants all wore different hats, which were difficult to remove, complicating their task of elucidating the state of the cyber conflict field.

Finally, the group tried to answer the question of whether cyber norms can only emerge through catastrophe. Although views were mixed, they were primarily optimistic: even if idealism does not carry the day, the threat of catastrophe can motivate just as well as catastrophe itself. States have managed to forecast and regulate (if not solve) many problems before they resulted in tragedy. Examples of successful precautionary regulation include treaties governing nuclear proliferation and space. It should be noted, however, that the 2017 SOTF conference was held not long before the United Nations Group of Governmental Experts (GGE) meeting, which ended the previously positive trend in cyber norm development by failing to reach consensus.[6] In any case, it cannot be known what the future holds.[7] Whether or not a catastrophe is necessary to catalyze norm formation, is one likely to occur?[8] The novelty of cyber means is commonly thought to raise the risk of inadvertent or accidental escalation, but we might also expect states' mutual interest in stability to be a dampening force. Furthermore, ambiguity in states' beliefs (and beliefs about their adversaries' beliefs) about cyber law may actually help in some cases by dampening enthusiasm for potentially provocative attacks.

Filling the Gaps: Recent Theoretical Progress

After defining terms for the purpose of the discussion, workshop participants addressed select "gaps" identified in the 2016 report: (1) the regulation of low-level conflict; (2) appropriate thresholds and standards for (a) the activation of the right to national self-defense, and (b) attribution of state responsibility for acts by non-state actors; (3) problems related to overlapping jurisdictional claims and enforcement; and (4) remaining controversies in international humanitarian law. Finally, they asked how definite markers for international norms and *opinio juris* (the beleive that an action is taken out of a sense of legal obligation) can be recognized in practice.

Definitions

The group recognized that while new definitions cannot entirely quell internal disagreement over policy preferences, cogent terminology helps prevent talking at cross purposes. Therefore, the participants' first task was to converge on definitions for several key terms, including "cyber operations" and "norms." The U.S. Army describes "cyber operations" as "the premeditated use of disruptive activities, or the threat thereof, against computers and/or networks, with the intention to cause harm or to further social, ideological, religious, political or similar objectives . . . or to intimidate any person in furtherance of such objectives."[9] Matthew Waxman puts it more succinctly: "Efforts to alter, disrupt, or destroy computer systems or networks or the information or programs on them."[10] Participants criticized these definitions as too focused on computers and computing systems, overlooking the personnel behind the keyboard as well as the Internet of Things (IoT)—assets beyond the digital. They also experienced difficulty in pinning down "evolving" technical terms, such as "network." Finally, conventional definitions of "operations" were seen as concentrated too much on attack and offense.

The definitions given to "Norms" in the social science literature range from "standards of appropriate behavior for actors of a given identity," to a special class of social institutions (or what Douglas North calls "rules of the game . . . [or] humanly devised constraints").[11] Participants considered whether "acceptable" should be substituted for "appropriate," reasoning that norms only become norms when they are diffused, which requires acceptance from other actors. One participant advised that "norms" be defined as both acceptable and *expected*, since many actions may be "acceptable" but never done and thus are not "normal" behavior. Finally, participants agreed that norms are what actors *do*, not what they *want*. For example, it is not a *violation* of any established norm to "dox" politicians, merely an affront. This distinction is important; if the positions held by cyber law advocates fail to reflect state practice in the international system, then their tenets are merely aspirations rather than decided norms.

Self-Defense Thresholds

Participants agreed that *jus ad bellum* customary law governs the permissibility of operations exceeding the threshold of "armed attack" articulated by the ICJ.[12] However, two problems remain. First, where is that threshold in practice? How would we know when a cyber operation had exceeded it? The answer is important both for preventing serious attacks (if a line is drawn, attackers might be deterred from crossing it) and for remedying serious attacks if and when they occur (dispute settlement fora need standards on which to draw, and the international community may need to be persuaded that the line has in fact been crossed). Yet, as participants noted, this question has already been debated for decades with no convincing answers. One participant described the obsession with this question as a "suffocating legal asymmetry," arguing that the United States is "paranoid" about committing an act of war, failing to recognize that, due to the extreme power imbalance between the United States and other states, very few (if any) U.S. actions in cyberspace, no matter how catastrophic or insidious, would be construed by its adversaries as *casus belli*. Whether or not this type of risk aversion is pervasive, it is counterintuitive to international relations theories about escalation in gray zones.

Not all cyber news is bad news. Since the 2016 SOTF report, theories on self-defense and state responsibility have advanced significantly, thanks in large part to the release of the *Tallinn Manual 2.0*. The *Tallinn Manual 1.0* outlined only basic thought on the circumstances in which retorsion or other countermeasures, often the only legal recourse a victim state has against a perpetrating state, might be permissible. The second *Manual* builds on the first to offer more specific advice, particularly with respect to cross-domain countermeasures. It also clarifies where the boundaries of knowledge lie: the *Manual*'s authors could not agree on whether collective security countermeasures were permissible, and no provision for countermeasures against non-state targets could be made (although the authors did agree that the "plea of necessity" for actions against non-state actors was reasonably analogous between cyber and conventional domains).

State Responsibility

For non-state actors to be held accountable, their actions must be connected to a state. As the UN General Assembly and UN GGE articulated in 2013, the principle of sovereignty, and thus the state veil protecting malicious non-state actors from international accountability, continues to hold in cyberspace. However, the UN Charter's exhortation for peace and security also applies, illustrating an inherent tension in public international law. Fortunately, the missing link—the law of state responsibility—has been slightly clarified since the 2016 SOTF report. The bad news is that this body of law does little to disincentivize state delegation to non-state actors, i.e., proxy wars, particularly in the cyber domain, where attribution to the state may be difficult to prove. The *Tallinn Manual 2.0* explains that extant state responsibility law applies, but not in a one-to-one mapping because virtual military assets, unlike physical ones, can be easily spoofed or commandeered. Given this difference, a more restrictive test of state "control" might be required. Similarly, although financial aid to malicious non-state groups is discouraged, state patrons are responsible only for the provision of aid itself, not for what non-state actors do with it. In the context of cyber conflict, the provision of aid—in the form of knowledge, funding, or code—may be all that is required for a non-state group to carry out complex and wide-ranging operations.

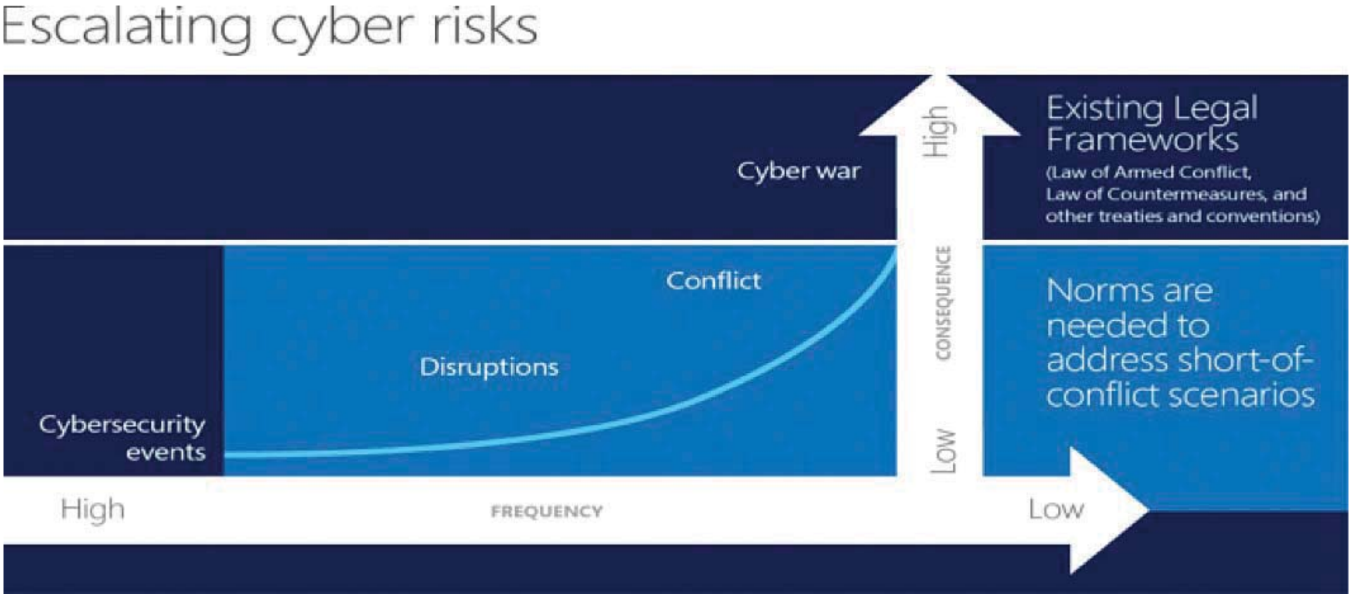
Other scholars have examined this issue. For example, a "Symposium on Cyber Proxies" was held at Columbia's School of International and Public Affairs in the spring of 2016.[13] Others have stressed the need to search outside of conventional sources of state responsibility law[14] for helpful general principles, including *sic utere tuo ut alienum non laedas* (use your own property so as not to harm another), which may also apply.[15] As Katharina Ziolkowski argues, the fact that states do not advertise their delegation to proxies gives weight to the idea that *opinio juris* may already exist in a liminal state.[16]

Even when state responsibility law is clear, operationalizing it requires technical attribution that is sometimes beyond the victims' capability.[17] While the international community may know victimization when it sees it, is case-by-case consideration enough, or must a clear legal standard for the burden of proof be developed? More efforts are needed to bridge applied legal standards with technical standards.[18] Participants agreed that another complication—and potential solution—lies in the increasing importance of private cyber intelligence firms, which are not subject to the same political or reputational constraints as governments or corporate data holders. The empirical question of whether the rise of private cyber intelligence firms has actually enhanced states' ability to hold perpetrators accountable remains unanswered. The proliferation of national disclosure laws and international Computer Emergency Readiness Team (CERT) cooperative arrangements may further support efforts to boost transparency.

Paradigms for Low-Level Operations (LLO)

Self-defense against armed attacks, a relatively easy debate, may not be the most necessary one. Much more difficult is the debate over how to regulate operations that definitely fall below the threshold of an armed attack or even a use of force. Angela McKay of Microsoft has stressed that whereas international law governs the region above the threshold, norms are needed to address everything else.[19] Figure 1, courtesy of Microsoft, illustrates this threshold.[20] Participants concurred that Figure 1 accurately represents the current framework and that the UN Charter Article 51 (referring to the right to self-defense) line could only be crossed by a cyberattack of extraordinary scale. However, participants (and, more broadly, scholars and states) have not reached consensus about the precise location of that line. Participants also noted a lack of evidence of any concrete norms in the sub-Article 51 region.

FIGURE 1: Use of Force Framework


Escalating cyber risks
Cyber war
High
Existing Legal Frameworks
(Law of Armed Conflict, Law of Countermeasures, and other treaties and conventions)
Conflict
CONSEQUENCE
Disruptions
Norms are needed to address short-of-conflict scenarios
Cybersecurity events
Low
High
FREQUENCY
Low

Rather than continuing to debate the location of the armed attack threshold, some panelists urged that the issue be reframed in terms of sovereignty. While LLOs will likely never be regulated by self-defense principles, they might be governed by the principle of sovereign non-interference.[21] Ironically, sovereignty—or rather, the normative prohibition against sovereign intrusions in international relations—is the way states such as China and Russia have traditionally framed the cyber debate. Although Western countries view the Internet fundamentally differently, as a domain that inherently transcends national borders, this is one area where a mutually agreeable bargaining solution may exist, potentially opening up space for new norms.

Of course, as some participants pointed out, the United States seems relatively accepting of inward-facing sovereignty standards (such as China's "Great Firewall") while less accepting of transborder activities (such as the 2015 attack on Github, for which China was blamed).[22] Likewise, the whole of the working group acknowledged practical limitations of applying sovereign non-interference to cyber activity. Establishing a shared definition of "sovereignty" in cyberspace, a necessary first step, would trigger fierce disagreement among states. While a well-established set of extant law on sovereignty could be ported, the political feasibility of states adhering to that law remains dubious.

Also at issue is the controversy over which existing legal paradigm might offer the greatest traction over LLOs. The concept of the "use of force" applies below the threshold of "armed attack" in theory, but its precise application is subject to considerable debate. Moreover, even if it were possible to establish an action as "unlawful" because it constituted a use of force in cyberspace, enforcement would present significant challenges in practice. The working group explored three alternative models. First, the criminal model espoused by one school of scholarly thought argues that direct culpability lies with individual perpetrators. When such individuals act of their own accord, frameworks like the Budapest Convention[23] and a system of overlapping mutual legal assistance treaties (MLATs) provide the means for victim states to seek justice.[24] But this paradigm falls short when LLOs are conducted on behalf of a state. Additionally, the system of MLATs and criminal cyber conventions is incomplete. The Budapest Convention, while a significant milestone, is not comprehensive and has not been ratified by many states. Further, many states lack bilateral MLATs with one another. Conflicting jurisdictional claims are another potential problem.

Participants next considered the environmental model.[25] Customary environmental law is based on doctrines such as the "no harm" and "good neighborliness" principles, which have been reaffirmed in several ICJ cases.[26] While transborder toxic spills, harmful emissions, and water contamination make colorful analogies, however, this body of law is largely undeveloped. There is a dearth of primary sources such as robust customary rules. Treaties drafted strictly for environmental issues would have to be renegotiated for cyber operations *per se*. Environmental law is also intensely fact-based, making it a poor analog for cyber-attack victims with limited forensic capabilities or disincentives to public disclosure.

Finally, participants considered the possibility that no extant model adequately fits. The scenario in which these LLOs are *non liquet*, as the earliest cyber legal theorists argued, presents both the greatest opportunity and greatest risk.[27] A new, well-designed, universal treaty would theoretically allow the international community to tailor legal standards to cyber operations. However, the divisiveness of the issue and political self-interest of states call into question the feasibility of *lex feranda*.[28] The rapporteur notes that this approach could also be a double-edged sword in that, until a cyber treaty is negotiated, treating cyber LLOs as *de novo* is an admission that the zone below armed attack is indeed the "wild west."

Participants also considered possible models not yet proposed in the literature, including public health, cross-sectoral retaliation (an economic concept), and the common/maritime law principle of hot pursuit, which allows for the pursuit of suspected belligerents across ordinary jurisdictional boundaries under exigent circumstances. Another proposal was to apply the "unwilling and unable" doctrine to cyberspace, although it is not clear how the doctrine, which is highly contested when applied to traditional military interventions, fits with LLOs, which are, by their nature, limited.[29] Counterintuitively, another possibility is to simply take a more permissive view of the law on countermeasures in the hopes that mutual risk will encourage host states to crack down on low-level activity.[30]

Jurisdiction & Enforcement

"Jurisdiction" refers to "the competence of States to regulate persons, objects, and conduct under their national law, within the limits imposed by international law."[31] In international law, jurisdiction is usually divided into three categories: prescriptive, adjudicative, and enforcement. Discussions around cyber law usually relate to the first category—the creation or articulation of rules in cyberspace. Scholarly attention has increasingly turned to the latter two; that is, the rights of states to respond when these rules are broken. Unfortunately, findings have only elucidated how limited these rights actually are.

"Jurisdiction to adjudicate," in this case, refers to the right of states to subject persons suspected of perpetrating cybercrimes to trial in courts with competency to hear such cases. As with physical claims, the strength of this type of jurisdiction turns on physical territoriality: extraterritorial jurisdiction over cybercrimes must be based on conventional principles.[32,33]

The legal concept of "jurisdiction to enforce" refers to the right to intervene against cybercrimes emanating across national borders. Extending this type of jurisdiction is an inherently political problem. The paucity of avenues for addressing cybercrime intensifies existing challenges. Enforcement against malicious actors who operate across borders hinges on the cooperation of host states. When host states are uncooperative, accusers generally have few legal avenues to resolve their grievances. Greater encouragement of MLATs or treaty frameworks like the Budapest Convention might help address this lack of recourse. Since there has been little scholarly work on MLATs since 2016, this may constitute a research opportunity. Others scholars have raised the idea of erecting an international cybercrime court to which states could submit their claims.[34] Short of that, coercive instruments, such as economic sanctions, may be the only means of resolution available to states that feel they have been wronged.

Controversies in International Humanitarian Law (IHL)

The debate over IHL's application to cyber warfare historically ranged between three schools. Idealists argued that, despite the fact that nowhere is cyber specifically enumerated in the law of armed conflict, non-derogable rules organically emerge by analogy or through public conscience.[35] Realists, conversely, postulated that applications exist but are more limited; namely, only beyond the kinetic divide.[36] Finally, skeptics held that IHL applies strictly to conventional domains, that cyber is *sui generis*, or that the factual challenges are insurmountable.[37]

As of 2017 and the publication of two *Tallinn Manuals*, this debate has (in theory) been settled. IHL applies in a restrictive sense, somewhere between the idealist and realist positions. The second *Tallinn Manual* further clarifies Geneva, consular, and human rights law, for instance by holding that the personally identifiable information of *hors de combat* and diplomatic staff are protected in times of war.[38] Still unanswered are several other important questions, such as clarifying what core IHL principles—including the prohibition on acts of perfidy and rules requiring fixation of distinctive emblems—would look like in cyber operations.

Theory Versus Reality: Developments in Practice

Linking theory and practice requires a quantitative analysis of state behavior and beliefs (*opinio juris*), the two necessary criteria for determining the emergence and formation of customary international law. While state practice on larger cyber issues, such as self-defense and IHL, has been limited, there is a growing body of data tracking unilateral state expressions. This data is important both for legal theory and for the state of the field. We must first know what ideals have been professed by the international community—what have states said?—before comparing nonverbal practice. As Michael Schmitt asserts, "states [now] need to roll into the game and start firming up the norms."[39]

The UN Office for Disarmament Affairs maintains a listing of the views of member states, which includes reports by the Secretary General as well as direct submissions by states parties.[40] In 2016, nineteen states made statements on the record, more than double than did in 2015. These included a number of influential players—Australia, Canada, India, Japan, and the United Kingdom—but also many from the developing world. The Geneva Internet Platform's Digital Watch Observatory, in partnership with the Internet Society, also maintains a collection of resources, including UN GGE reports and related General Assembly resolutions organized by year.[41]

Finally, the Carnegie Endowment for International Peace maintains a searchable "norms index," which the website describes as "track[ing] and compar[ing] the most important milestones in the negotiation and development of norms for state behavior in and through cyberspace."[42] Users can compare specific language from international declarations and other discourse on issues ranging from aspirational norms to threat perception. This project promises to be of significant value to policy researchers.

What does the record reveal about the state of the field? While more analysis is needed, it is possible to trace some broad themes. Since the Russian Federation first brought the issue of digital security to the UN General Assembly in 1998, the international debate has been divided into roughly two camps over the nature of cyber sovereignty.[43] Russia and China have cooperated closely on cyber norms since 2011, repeatedly reaffirming their shared vision in a series of joint statements that *The New York Times* branded a "nonaggression pact."[44,45]As Kristin Eichensehr succinctly writes, this camp "advocate[s] a sovereignty-based model of cyber governance that prioritizes state control," whereas the United States and its allies "argue that cyberspace should not be governed by states alone," but rather in conjunction with a multiplicity of stakeholders.[46]

Even as sovereignty has remained a sticking point, agreement has become increasingly possible on other types of norms. The UN GGE, a collection of representatives from 25 influential countries of "equitable geographic distribution," has met five times since 2004.[47] The GGE has progressively moved the debate forward—from discussions in 2005 to consensus reports in 2010 and 2013.[48] At the 2013 meeting, GGE members agreed that international law applies to cyberspace just as it does to other domains, although the precise implications of this admission were not discussed.[49] It also recognized that general IHL principles may in some cases apply, and that states retain territorial jurisdiction over cyber infrastructure.[50]

By 2015–2016, many onlookers became optimistic that a system of norms on several important issues was indeed coalescing.[51] The GGE and the Group of Twenty (G20) each produced consensus documents. The GGE report enumerated five "limiting norms" geared around the permissibility of conducting cyber warfare or knowingly allowing one's own cyber infrastructure to damage another state's, as well as six "positive duties" affirming the need for multilateral cooperation and information sharing in the event of an attack.[52] Likewise, the G20 communiqué suggested that industrial espionage was prohibited, in line with U.S. interests.[53] This enshrined a bilateral understanding reached between the U.S. and China earlier in 2015, in which Presidents Xi and Obama agreed to cooperate in four areas: (1) a joint commitment to norm-building, (2) anti-cybercrime dialogue, (3) abstaining from knowingly supporting cyber-theft of IP, and (4) providing timely responses to transnational investigations.[54] Finally, fora such as the NATO Cooperative Cyber Defense Center of Excellence's (CCDCOE) Cyber Conflict Conference welcomed input from a more diverse array of participants, including legal scholars from China.[55]

But what was *not* said in these discussions is as important as what *was* said.[56] While the 2015 GGE was hailed for its consensus over norms, its "progress on international law" has been described as "modest."[57] For example, agreement could not be reached at the 2015 GGE over language about Article 51, the exclusion of which, *Tallinn Manual* editor Michael Schmitt argues, is an "untenable notion as a matter of international law."[58] Then, as mentioned previously, in 2017, after expanding membership to 25 countries, the GGE suffered an embarrassing failure to reach consensus, ending (or at least stunting) its trajectory as the lead forum for multilateral cyber law discussions.

The 2017 GGE reportedly did make headway on a number of important issues, such as a definition for the "knowledge" requirement in the previous report's rule that "states should not knowingly allow their territory to be used for intentionally wrongful acts"; a proscription against "hackbacks" (offensive operations by private sector entities against suspected offenders); and a *de facto* categorization of the Domain Name System as critical infrastructure, off-limits to attack.[59] Agreement is said to have broken down over U.S. insistence that the Article 51 threshold should apply, at least in principle, to cyberattacks of a sufficient scale.[60] The U.S. government and Western observers were quick to blame Russia, China, and Cuba.[61] In her statements at the UN, U.S. delegate Michele Markoff attributed the impasse to "the reluctance of a few participants to seriously engage on the mandate on international legal issues."[62] However, the United States' insistence on the application and scope of key issues like Article 51 is by no means new, suggesting that it may have played a key role in the decision to end the discussions prematurely.

The failure of the 2017 UNGGE is not cause to abandon hope of normative convergence. Persuasion and adoption, when they occur, are gradual processes, and setbacks are inevitable. But as James Lewis has said, "the world's a long way from agreeing on basic principles of cyber sovereignty and those principles may not be written on U.S. terms."[63] Moreover, concerns have been raised that recent State Department shakeups could diminish the United States' say on cyber norm evolution moving forward.[64] As participants in the 2017 SOTF conference noted, to ensure favorable norms, the United States must both *stake out* and *set* precedents over time. Participants argued that what matters is not simply persuading the world, but rather maintaining that persuasive position.

Aside from (or in lieu of) *expressed* norms, there has been progress—as well as some retrogress—in *behavioral* norms. Following the United States' 2014 indictment of five Chinese People's Liberation Army (PLA) officers, the threat of economic sanctions, the 2015 Rose Garden agreement, and the G20 agreement, suspected Chinese hacking activity appeared to sharply decline, according to a much-reported FireEye analysis (although it is not clear to China-watchers whether more cooperative patterns are a result of, or epiphenomenal to, the agreements).[65] If U.S. strategies were indeed effective in dealing with China, similar indictments made against suspected Russian Federal Security Service (FSB) officers in March 2017 may also have a stabilizing effect.[66]

Given the political challenges of multilateral dialogue, private sector and other nongovernmental advocates may play a significant role in shaping cyber law discourse. A number of notable companies, including Google and Microsoft, have already displayed leadership. In early 2017, Brad Smith, Microsoft's President and Chief Legal Officer, called for a "digital Geneva Convention."[67] Shortly thereafter, Google proposed a framework that would obviate the need for MLATs by allowing governments to request evidence directly from Internet companies.[68] Norm entrepreneurs, particularly those from the private sector, will play an increasingly large role in the years to come.

Finally, although global norms are ideal for a globalized Internet, regional cooperation has gained more ground. The 2016 Organization for Security and Cooperation in Europe (OSCE), following up on its 2013 accord,[69] established a network of confidence-building measures and crisis hotlines spanning 57 countries. Similarly, in 2017, the European Council agreed on a "cyber diplomacy toolbox" that purports to streamline joint European diplomatic responses to cyber threats.[70] Whether regional cooperation will harden political blocs instead of helping to diffuse norms is unknown, but some cooperation is better than none for security and stability.

Modern Hague and Geneva law governing conduct in wartime arose in the wake of terrible historical experiences from the Battle of Solferino through the Crimean War, the American Civil War, World War I, and World War II. Is a catastrophe required for the formation of cyber norms? When this question was posed to SOTF participants, their answer was a unanimous "no." Although urgency turns discourse into action, anticipation of a terrible experience may be enough to catalyze change. If states universally recognize that particular types of behavior would lead to catastrophe, norms proscribing such activities should be easy to arrive at. This may be the case for rules like those against attacking CERTs and critical infrastructure in peacetime, for example, which have not been seen as controversial. Of course, this optimism hinges on a model of foreign policy decisionmakers as rational calculators, an assumption that may not always match reality. Nor is there any guarantee that norms established during peacetime would, as one reviewer put it, "survive first contact" in wartime.

Because norms are often established through leadership,[71] shaping norms in domains where actions are inherently secret or covert remains much more difficult. Further complicating matters, cyberspace has multiple stakeholders; most of the infrastructure and operators are nongovernmental. As a result, governments have only an indirect say over many behaviors, limiting their autonomy in this role.

Mistrust remains high. Despite China's apparent cooperativeness, one participant claimed that U.S. government insiders remain skeptical. There was also serious divergence throughout the room on the utility of the 2015 Rose Garden agreement,[72] with one participant arguing it was "100 percent a loss for the United States [and a] coup for China." Other participants were more optimistic, maintaining that, because the downturn in hacking trends appears to predate the agreement, indictments, or threat of sanctions, China's cooperation must be in its own interest. To realists and idealists alike, nothing is more stable than cooperation through mutual self-interest, at least until those interests change. The key may therefore be to codify law that aligns with the mutual self-interest of the most powerful states in the cyber domain, then shapes and constrains behavior even when those interests change.

Emerging Issues

Given the time constraints of the SOTF conference and the salience of particular topics, it was impossible to explore the full spectrum of legal and ethical issues that touch on cyberspace. However, some effort was devoted to brainstorming issues that may be of increasing importance in the years to come, for discussion at future SOTF workshops. First, although its use has evidently decreased in recent years, cyber network espionage (CNE)—especially industrial—will remain a hot-button issue. The *Tallinn Manual 2.0* discusses the issue at length, with no consensus about its permissibility at high levels. Widespread consensus holds that espionage, broadly, is not a violation of international law, but merely the domestic law of the target state (discussed more below). Cyber has made it increasingly possible for states to spy on entities outside the scope of traditional intelligence targets, such as foreign corporations. This has understandably given rise to heightened concern, particularly for high-tech, industrialized economies like the United States.

Unlike use of force issues involving computer network *attacks* (CNA), the debate over CNE norms has the potential to cut across traditional political divides. Even some Western countries do not share the U.S. view that the state and markets should be strictly compartmentalized and so, by logical extension, should their secrets.[73] However, the political consequences of being "for" industrial espionage are grave, creating a normative wedge whereby only the "against" side is vocal. In international law, silence can count as acquiescence.

Conventional espionage is normatively accepted between adversaries, and there is no international law proscribing it. As one participant put it, "espionage is more of an offense than a violation of anything." But in cyber operations, in the U.S. domestic law context, Title 10 (military) and Title 50 (intelligence and national security) roles are blurred.[74] And from the recipient's side, it is often hard to tell whether a system intrusion is an attempt to gather sensitive information or to plan an impending attack.[75] Although work has been done on the relationship between Title 10 and Title 50 in conventional spaces, there is a dearth of scholarly research on this dynamic in cyberspace.[76] The participants agreed that future SOTF workshops should consider U.S. national security law, broadly, without displacing the ongoing international law conversation.

Election influence or political inference is of obvious salience in the wake of the 2016 election. Cyber intelligence firm Crowdstrike released a report implicating Russian agents in a breach of Democratic National Committee (DNC) files shortly before the election.[77] About a month later, the U.S. Office of the Director of National Intelligence and the Department of Homeland Security stated that "the U.S. Intelligence Community (USIC) is confident that the Russian Government directed the recent compromises of emails from U.S. persons and institutions, including from U.S. political organizations."[78] In the time between then and the drafting of this report, Facebook surrendered information about 3,000 suspect advertisements to the U.S. Senate Intelligence Committee.[79] Similar concerns have arisen in other democracies: Emmanuel Macron, then a presidential candidate in France, had nine gigabytes of his emails leaked two days before the election.[80] Though repeatedly blamed, Russia is not the only source of such interference: online agitators from the United States were reportedly more culpable in interference with the 2017 German elections.[81] That the phenomenon of cyber interference in electoral politics has become routine not only provides a slew of case studies meriting deeper analysis, but also suggests the importance of this analysis for understanding issues like attribution, state responsibility, and sovereign non-interference.

Mass data breaches, ransomware, and other major cybercrimes are another topic of importance. These attacks have become more threatening in recent years, as was illustrated poignantly in the 2015 Office of Personnel Management hack and again in the 2017 Equifax breach, which included the personally identifiable information of 21.5 million and 143 million Americans, respectively.[82] Related to this is the issue of public-private sector cooperation and information sharing, particularly as it relates to the vulnerabilities equities process (VEP). During his administration, President Obama exerted considerable effort to bridge the trust gap widened by the Snowden leaks between private industry and U.S. intelligence collection agencies.[83] President Trump has made similar overtures.[84] Yet, as the 2017 WannaCry leaks revealed, the National Security Administration (NSA) exploited a vulnerability it discovered in the Windows operating system rather than notifying Microsoft to patch it.[85] After hackers stole the NSA tool used to exploit Windows and made it available on the open market, it was employed around the world. Systems in dozens of countries, including British hospitals and the Russian Interior Ministry, were infected. Despite the need for offensive capabilities, governments cannot treat cooperation as a one-way street and expect trust to be repaired. A 2017 U.S. Senate bill, the Protecting Our Ability to Counter Hacking Act of 2017 (PATCH Act), is one proposal to optimize this tradeoff by subjecting each new discovery to review.[86]

Given that artificial intelligence, autonomous systems, and the Internet of Things may all become more pervasive, some participants criticized *Tallinn 2.0* for focusing excessively on the human element. What ethical issues arise around self-executing, intelligent programs that remove human beings from the loop? How do they fit into the threshold debate portrayed in Figure 1?

Finally, because it may not be possible to say anything new about the "use of force" debate at the next SOTF conference, barring new and unexpected developments,[87] designating "sovereignty" as the umbrella topic might prove more fruitful. Beyond its public international law ramifications, sovereignty touches upon and could usher in subtopics that have thus far been overlooked in the SOTF forum: human rights law, domestic surveillance, comparative national security law, and public-private relations. These issues are all on a minable research frontier.

Summary & Recommendations

The past few years have been full of surprises. The world achieved a measure of cooperation on one major issue (economic espionage), lost it on another (the UNGGE), and witnessed the rise of new cyber threats (election interference). It is difficult to say where cyber norms are headed from here. Participants at the conference recalled "The Five Futures of Cyber Conflict and Cooperation," a 2011 article by Jason Healey. In it, the author imagines five potential futures—status quo, conflict domain, balkanization, paradise, and cybergeddon—assessing the nature, stability, intensity, and likelihood of each.[88] His conclusions are grim, but not dire: status quo or low-level conflict is likeliest, and cyber will continue to support a range of activities, both malevolent and benign. In 2018, is it reasonable to have the same expectations, or should we assign new probabilities to these potential futures? The participants were not able to reach consensus, but it was a useful thought exercise. Importantly, it highlighted that in each of the five potential futures, the law and ethics of cyber operations look very different.

The logical next question was to ask what the optimal U.S. grand strategy for cyberspace would be. In one participant's words, "which future should the United States be doubling down on?" Participants asserted that when the domain was in its nascent stages, the United States had an insurmountable lead in manpower, infrastructure, and companies; it built everything about the Internet. That may no longer be the case—many Internet companies operate abroad, foreign governments are increasingly competitive, and even in American universities, a great many computer science students are foreign nationals.

A minority of participants argued that the private sector is "still living in the Golden Days" of the past and fails to see the big picture. Balkanization is the future because it is the "min/max strategy" (that is, a strategy taken to minimize one's own maximum loss and maximize one's own minimum gain[89]) for states within which companies operate. A slim majority disagreed, arguing that because the private sector is globalized and profit-driven, it will align with whoever can make it money, putting bottom-up pressure on governments to keep online markets open. A small set of participants refuted this view by relating an anecdote about the "PLA 5" indictments, discussed in brief earlier. These participants claimed that Pittsburgh industrial players, angry that the federal government had not adequately defended them from intrusions, persuaded U.S. Attorney David Hickton to "go rogue" in a fait accompli that was not wholly coordinated with main Justice Department or State Department priorities. By allowing the Western District of Pennsylvania to name the PLA 5, these participants alleged, companies knowingly sacrificed business in China. Other participants were unable to verify this account.

What if the upward trend in cyber norms really has been broken? Do we even need law, or are politics and strategy enough to sustain an uneasy peace, as they were during the Cold War? Several participants agreed that legal and ethical ambiguity has some merit, as uncomfortable as this assertion makes those who look to the law expecting clarity through bright-line rules. If the Article 51 line was known with certainty, states might be more willing to walk directly up to it without crossing. Absent this knowledge, they may find it safer to act conservatively and avoid provocation.

Finally, the working group uncovered several key issues on which more research is needed. First, a listing or map of MLATs—as well as more social science research on their causes and effects—would be helpful, given their proliferation and increasing importance in solving jurisdictional disputes. Second, given the state of the Article 51 debate, a next step might be to plot actual cyber incidents on the schematic in Figure 1. In so doing, researchers may be able to reverse-engineer behavioral norms and thus infer where states believe the red lines lie based on how they have behaved. Third, more attention should be paid to issues of domestic cyber law, including the blurred lines between military and intelligence operations; the emergence of national legal frameworks in the United States, China, Europe, and elsewhere in recent years; and human rights principles, from which the debate over IHL has detracted attention. Research on these topics would be of tremendous value to both cyber theorists and cyber practitioners.

About the Author

Justin Key Canfil is a Ph.D. Candidate within the Columbia University Department of Political Science.

End Notes

[1] Schmitt, Michael N., ed. Tallinn Manual on the International Law Applicable to Cyber Warfare. Reprint edition. Cambridge University Press, 2013.

[2] Cantwell, Douglas. "Legal and Ethical Issues." Cyber Conflict State of the Field Workshop Report. Cyber Conflict Studies Association, 2016, pg. 102.

[3] Ibid., pg. 103.

[4] Ibid., pg. 104, emphasis mine. Other recent works include Osula, Anna-Maria, and Henry Roigas, eds. International Cyber Norms: Legal, Policy, and Industry Perspectives. NATO Cooperative Cyber Defence Centre of Excellence, 2016. https://ccdcoe.org/sites/default/files/multimedia/pdf/InternationalCyberNorms_full_book.pdf

[5] Cantwell, Ibid., pg. 105

[6] Väljataga, Ann. "Back to Square One? The Fifth UN GGE Fails to Submit a Conclusive Report at the UN General Assembly." NATO CCDCOE, September 1, 2017. https://ccdcoe.org/back-square-one-fifth-un-gge-fails-submit-conclusive-report-un-general-assembly.html

[7] See Maurer, Tim, and Kathryn Taylor. "Outlook on International Cyber Norms: Three Avenues for Future Progress." Just Security (blog), March 2, 2018. www.justsecurity.org/53329/outlook-international-cyber-norms-avenues-future-progress/

[8] On this, see relevant research in progress by the author: Canfil, Justin Key. "Defense Divinations: The Design of International Contracts Under Uncertainty About Military Technological Change," presented at the 2018 International Studies Association (ISA) conference. Working paper available upon request.

[9] "Cyber Operations and Cyber Terrorism." U.S. Army Training and Doctrine Command, Handbook No. 102. Fort Leavenworth, Kansas, August 15, 2005.

[10] Waxman, Matthew. "Cyber-Attacks and the Use of Force: Back to the Future of Article 2(4)." Yale Journal of International Law 36, no. 2 (January 1, 2011). http://digitalcommons.law.yale.edu/yjil/vol36/iss2/5

[11] Katzenstein, Peter J. "Introduction: Alternative Perspectives on National Security." In Peter J. Katzenstein (ed.), The Culture of National Security: Norms and Identity in World Politics. Columbia University Press, 1996, pg. 5; North, D. Institutions, Institutional Change and Economic Performance. Cambridge University Press, 1990.

[12] See, e.g., Case Concerning Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States of America); Merits, International Court of Justice, June 27, 1986.

[13] See also Healey, Jason. "The Spectrum of National Responsibility for Cyberattacks." Brown Journal of World Affairs 18, no. 1 (Fall/Winter 2011); Lin, Herbert. "Attribution of Malicious Cyber Incidents: From Soup to Nuts." Journal of International Affairs 70, no. 1 (Winter 2016): 106.

[14] Conventional sources are thought to include the 2001 International Law Commission's Draft Articles and ICJ caselaw; e.g. Corfu Channel, Iran Hostages, Tadic, Nuclear Weapons Advisory, Nicaragua, although primary sources per se are scant. See also notes from U.S. Cyber Command's "Cyberspace Operations in the Gray Zone" conference in February 2018: Adams, Michael J., and Megan Reiss. "International Law and Cyberspace: Evolving Views." Lawfare, March 4, 2018. www.lawfareblog.com/international-law-and-cyberspace-evolving-views

[15] Ziolkowski, Katharina. "Customary Rules of International Environmental Law—Can They Provide Guidance for Developing a Peacetime Regime in Cyberspace?" In Katharina Ziolkowski (ed.), Peacetime Regime for State Activities in Cyberspace. International Law, International Relations and Diplomacy. NATO CCD COE, 2013, pg. 135.

[16] Ibid.

[17] Refer to the chapter on attribution in this volume.

[18] For existing research in this vein, see Canfil, Justin Key. "Honing Cyber Attribution: A Framework for Assessing Foreign State Complicity." Journal of International Affairs 70, no. 1 (Winter 2016): 217–226.

[19] McKay, Angela. "International Cybersecurity Norms." Microsoft. Accessed 2017.

[20] Ibid.

[21] Although c.f. proposals like the "accumulation doctrine": Ruys, Tom. The Intangible "Armed Attack": Evolutions in Customary Practice Pertaining to the Right of States to Self-Defence and the Quest for a Definition of "Armed Attack" Under Article 51 UN Charter. Proefschrift, 2009; Gervais, Michael. "Cyber Attacks and the Laws of War." Berkeley Journal of International Law 30 (2012): 525.

[22] Rawlinson, Kevin. "Evidence Links China to GitHub Attack." BBC News, March 31, 2015, sec. Technology. www.bbc.com/news/technology-32138088

[23] Convention on Cybercrime, ETS No.185, Budapest, November 23, 2001. www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185

[24] Boer, Lianne J. M. "'Echoes of Times Past': On the Paradoxical Nature of Article 2(4)." Journal of Conflict and Security Law 20, no. 1 (April 1, 2015): 5–26. doi:10.1093/jcsl/kru012; Hathaway, Oona, and Rebecca Crootof. "The Law of Cyber-Attack." Faculty Scholarship Series, January 1, 2012. http://digitalcommons.law.yale.edu/fss_papers/3852; Todd, Graham H. "Armed Attack in Cyberspace: Deterring Asymmetric Warfare with an Asymmetric Definition Cyberlaw Edition." Air Force Law Review 64 (2009): 65–102.

[25] Healey, Jason, and Hannah Pitts. "Applying International Environmental Legal Norms to Cyber Statecraft." I/S: A Journal of Law and Policy for the Information Society 8, no. 2 (2012); Marauhn, Thilo. "Customary Rules of International Environmental Law—Can They Provide Guidance for Developing a Peacetime Regime in Cyberspace?" In Katharina Ziolkowski (ed.), Peacetime Regime for State Activities in Cyberspace. International Law, International Relations and Diplomacy. NATO CCD COE, 2013.

[26] See, e.g., Gabcikovo-Nagymaros Project (Hungary/Slovakia), I.C.J. 7 (1997).

[27] See Hollis, Duncan B. "Why States Need an International Law for Information Operations." Lewis & Clark Law Review 11 (2007).

[28] Segal, Adam, and Matthew Waxman. "Why a Cybersecurity Treaty Is a Pipe Dream." CNN.com, October 27, 2011. http://globalpublicsquare.blogs.cnn.com/2011/10/27/why-a-cybersecurity-treaty-is-a-pipe-dream/

[29] Deeks, Ashley. "'Unwilling or Unable': Toward a Normative Framework for Extra-Territorial Self-Defense." Virginia Journal of International Law 52, no. 3 (August 2012): 483. https://ssrn.com/abstract=1971326

[30] For a full exposition of countermeasures, see Schmitt, Michael N., ed. Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations. 2nd edition. Cambridge University Press, 2017.

[31] Lotus judgment, at 23, quoted in "Jurisdiction," Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations, Ch 3. Cambridge University Press, 2013.

[32] These jurisdictional bases include passive personality, universality, nationality, etc.

[33] Lotus judgment, at 23, quoted in "Jurisdiction," Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations, Ch 3. Cambridge University Press, 2013.

[34] Kraft, W., and Claudia Streit. "Ideas on the Establishment of an International Court for Cyber Crime." World Council for Law Firms and Justice, 2011; Choudhury, Rajarshi Rai, Somnath Basak, and Digbijay Guha. "Cyber Crimes-Challenges & Solutions. " International Journal of Computer Science and Information Technologies 4.5 (2013).

[35] Greenberg, Lawrence T., Seymour E. Goodman, and Kevin J. Soo Hoo. Information Warfare and International Law. National Defense University Press, 1998; Kelsey, Jeffrey T. G. "Hacking into International Humanitarian Law: The Principles of Distinction and Neutrality in the Age of Cyber Warfare." Michigan Law Review 106, no. 7 (May 2008): 1427–1451; Cook, James L. "Is there anything morally special about cyberwar?" In Ohlin, Jens David, Claire Oakes Finkelstein, and Kevin Govern (eds.), Cyberwar: Law and Ethics for Virtual Conflicts. Oxford University Press, 2015.

[36] Dinstein, Yoram. "The Principle of Distinction and Cyber War in International Armed Conflicts." Journal of Conflict and Security Law 17, no. 2 (July 1, 2012): 261–277; Schmitt, Michael N. "The Law of Cyber Warfare: Quo Vadis." Stanford Law & Policy Review 25 (2014): 269; Schmitt, Michael N. "Peacetime Cyber Responses and Wartime Cyber Operations Under International Law: An Analytical Vade Mecum. " Harvard National Security Journal 8 (2017): 239–426.

[37] Aldrich, Richard W. "The International Legal Implications of Information Warfare." No. INSS-OP-9. Air Force Academy, Colorado Springs, CO, 1996; Turns, David. "Cyber Warfare and the Notion of Direct Participation in Hostilities." Journal of Conflict and Security Law 17, no. 2 (2012): 279–297; Shackelford, Scott J. "From Nuclear War to Net War: Analogizing Cyber Attacks in International Law." Berkeley Journal of International Law 27 (2009): 192; Asslani, Jabbar. "Study on the Legal Dimensions of the Cyber Attacks from IHL Perspective Abstracts." International Studies Journal 10 (2013–2014): I–XIII; Arimatsu, Louise. "A Treaty for Governing Cyber-Weapons: Potential Benefits and Practical Limitations." In 4th International Conference on Cyber Conflict, pp. 1–19. IEEE, 2012; Droege, Cordula. "Get off my Cloud: Cyber Warfare, International Humanitarian Law, and the Protection of Civilians." International Review of the Red Cross (2012); Geiß, Robin, and Henning Lahmann. "Cyber Warfare: Applying the Principle of Distinction in an Interconnected Space." Israel Law Review 45, no. 3 (2012): 381–399; Dunlap, Charles. "Perspectives for Cyber Strategists on Law for Cyberwar." Strategic Studies Quarterly (January 2011): 81–99.

[38] *Hors de combat*, literally meaning "outside the fight" is a term of art referring to noncombatants, including prisoners of war, the sick and wounded, and unarmed civilians.

[39] Quoted by Ansley, Rachel. "Tallinn Manual 2.0: Defending Cyberspace." Atlantic Council. Accessed September 26, 2017. www.atlanticcouncil.org/blogs/new-atlanticist/tallinn-manual-2-0-defending-cyberspace

[40] "Developments in the Field of Information and Telecommunications in the Context of International Security." United Nations Office for Disarmament Affairs. www.un.org/disarmament/topics/informationsecurity/

[41] "UN GGE." Geneva Internet Platform Digital Watch Observatory. https://dig.watch/processes/ungge#Resorces

[42] "Cyber Norms Index." Carnegie Endowment for International Peace. http://carnegieendowment.org/publications/interactive/cybernorms

[43] United National General Assembly Resolution 53/70. January 4, 1999. www.un.org/ga/search/view_doc.asp?symbol=A/RES/53/70

[44] See Letter of September 12, 2011, https://ccdcoe.org/sites/default/files/documents/UN-110912-CodeOfConduct_0.pdf; Letter of January 9, 2015, https://ccdcoe.org/sites/default/files/documents/UN-150113-CodeOfConduct.pdf; Roth, Andrew. "Russia and China Sign Cooperation Pacts." The New York Times, May 8, 2015. However, c.f. Davidson, Lincoln. "Despite Cyber Agreements, Russia and China Are Not as Close as You Think." Council on Foreign Relations, June 30, 2016.

[45] "China, Russia Sign Joint Statement on Strengthening Global Strategic Stability." Xinhua News, June 2016. http://news.xinhuanet.com/english/2016-06/26/c_135466187.htm

[46] Eichensehr, Kristen E. "The Cyber-Law of Nations." Georgetown Law Journal 103 (2015). https://georgetownlawjournal.org/articles/63/cyber-law-of-nations

[47] "2015 UN GGE Report: Major Players Recommending Norms of Behaviour, Highlighting Aspects of International Law." CCDCOE, August 31, 2015. www.ccdcoe.org/2015-un-gge-report-major-players-recommending-norms-behaviour-highlighting-aspects-international-l-0

[48] Grisby, Alex. "The UN GGE on Cybersecurity: What is the UN's Role?" Council on Foreign Relations, April 15, 2015.

[49] Marks, Joseph. "UN Body Agrees to U.S. Norms in Cyberspace." Politico. Accessed September 26, 2017.

[50] Schmitt, Michael, and Liis Vihul. "International Cyber Law Politicized: The UN GGE's Failure to Advance Cyber Norms." Just Security, June 30, 2017.

[51] Healey, Jason, and Tim Maurer. "What It'll Take to Forge Peace in Cyberspace." Carnegie Endowment for International Peace, March 20, 2017. See also, Maurer, Tim. "Cyber Norm Emergence at the United Nations." Science, Technology, and Public Policy Program Explorations in Cyber International Relations Project. Harvard Belfer Center, September 2011; and Finnemore, Martha, and Duncan B. Hollis. "Constructing Norms for Global Cybersecurity." American Journal of International Law 110, no. 3 (July 2016).

[52] "2015 UN GGE Report: Major Players Recommending Norms of Behaviour, Highlighting Aspects of International Law." CCDCOE, August 31, 2015.

[53] "No country should conduct or support ICT-enabled theft of intellectual property, including trade secrets or other confidential business information, with the intent of providing competitive advantages to companies or commercial sectors." G20 Leaders' Communiqué, Antalya Summit, November 15-16, 2015, pg. 26. www.mofa.go.jp/files/000111117.pdf

[54] Nakashima, Ellen, and Steven Mufson. "The U.S. and China Agree Not to Conduct Economic Espionage in Cyberspace." Washington Post, September 25, 2015, sec. National Security. www.washingtonpost.com/world/national-security/the-us-and-china-agree-not-to-conduct-economic-espionage-in-cyberspace/2015/09/25/1c03f4b8-63a2-11e5-8e9e-dce8a2a2a679_story.html

[55] Professor Huang Zhi Xiong from Wuhan University was invited to provide comments on Tallinn Manual 1.0, according to Deeks, Ashley. "Tallinn 2.0 and a Chinese View on the Tallinn Process." Lawfare, May 31, 2015.

[56] See comments by James Lewis, quoted in Marks, Ibid.

[57] Korzak, Elaine. "UN GGE on Cybersecurity: The End of an Era?" The Diplomat, July 31, 2017; Korzak, Elaine. "International Law and the UN GGE Report on Information Security." Just Security, December 2, 2015. www.justsecurity.org/28062/international-law-gge-report-information-security/

[58] James Lewis, quoted in Marks, Ibid.; Schmitt and Vihul, Ibid.

[59] Sukumar, Arun Mohan. "The UN GGE Failed. Is International Law in Cyberspace Doomed As Well?" Lawfare, July 4, 2017.

[60] Ibid.

[61] Korzak, Elaine. "UN GGE on Cybersecurity: The End of an Era?" Ibid.

[62] Markoff, Michele. "Explanation of Position at the Conclusion of the 2016-2017 UN Group of Governmental Experts (GGE) on Developments in the Field of Information and Telecommunications in the Context of International Security." Remarks delivered to the United Nations, June 23, 2017. https://usun.state.gov/remarks/7880

[63] Marks, Ibid.

[64] Jason Healey, quoted in Starks, Tim. "Top State Cyber Official to Exit, Leaving Myriad Questions." Politico. Accessed September 26, 2017. http://politi.co/2vdSHtx

[65] "U.S. Charges Five Chinese Military Hackers for Cyber Espionage Against U.S. Corporations and a Labor Organization for Commercial Advantage." Department of Justice Office of Public Affairs Press Release, May 19, 2014; "Red Line Drawn: China Recalculates its Use of Cyber Operations." FireEye iSight Intelligence, June 2016.

[66] "U.S. Charges Russian FSB Officers and their Criminal Conspirators for Hacking Yahoo and Millions of Email Accounts." Department of Justice Office of Public Affairs Press Release, March 15, 2017. Although c.f. comments by James Lewis, who argues that Beijing may be uniquely susceptible to "naming and shaming" tactics: Groll, Elias. "DOJ Charges Russian Intelligence in Huge Yahoo Hack." Foreign Policy (blog), March 15, 2017. https://foreignpolicy.com/2017/03/15/doj-charges-russian-intelligence-in-huge-yahoo-hack/

[67] Smith, Brad. "The Need for a Digital Geneva Convention." Microsoft on the Issues, February 14, 2017. https://blogs.microsoft.com/on-the-issues/2017/02/14/need-digital-geneva-convention/

[68] Digital Security & Due Process: Modernizing Cross-Border Government Access Standards for the Cloud Era. Google, 2017.

[69] Annual Report. Organization for Security and Co-Operation in Europe, 2013. www.osce.org/secretariat/116947?download=true

[70] "Cyber Attacks: EU Ready to Respond with a Range of Measures, Including Sanctions." Council of the European Union Press Release. June 19, 2017. www.consilium.europa.eu/en/press/press-releases/2017/06/19-cyber-diplomacy-toolbox/

[71] For example, norms on nuclear weapons in space, customary international law on continental shelf maritime boundaries, and norms regarding satellite overflight.

[72] Davis, Julie Hirschfeld, and David E. Sanger. “Obama and Xi Jinping of China Agree to Steps on Cybertheft.” The New York Times, September 25, 2015, sec. Asia Pacific. www.nytimes.com/2015/09/26/world/asia/xi-jinping-white-house.html

[73] Canfil, Justin Key. “Cyber Security and the Law: Managing Cyber Risk.” 2014 Conference Report. The French-American Foundation, 2015.

[74] Wall, Andru E. “Demystifying the Title 10-Title 50 Debate: Distinguishing Military Operations, Intelligence Activities & Covert Action.” Harvard Law School, 2011.

[75] Title 10 of the U.S. Code refers to the military operational side, whereas Title 50 encompasses intelligence operations.

[76] e.g. Wall, Andru E. “Demystifying the Title 10-Title 50 Debate: Distinguishing Military Operations, Intelligence Activities & Covert Action.” Harvard Law School, 2011. For an important exception regarding cyberspace law, see Chesney, Robert. “Offensive Cyberspace Operations, the NDAA, and the Title 10-Title 50 Debate.” Lawfare, December 14, 2011. www.lawfareblog.com/offensive-cyberspace-operations-ndaa-and-title-10-title-50-debate

[77] Bump, Philip. “Here’s the Public Evidence that Supports the Idea that Russia Interfered in the 2016 Election.” Washington Post, July 6, 2017.

[78] Ibid.

[79] Campbell, Barbara. “Facebook to Turn Over 3,000 Ads to Congress in Russian Election Interference Probe.” NPR.org, September 21, 2017.

[80] Greenberg, Andy. “NSA Director Confirms that Russia Really Did Hack the French Election.” Wired, May 9, 2017.

[81] Hjelmgaard, Kim. “There is Meddling in Germany’s Election—Not by Russia, but by U.S. Right Wing.” USA Today, September 20, 2017.

[82] Nakashima, Ellen. “Hacks of OPM Databases Compromised 22.1 Million People, Federal Authorities Say.” Washington Post, July 9, 2015; Barrett, Devlin. “Chinese National Arrested for Allegedly Using Malware Linked to OPM Hack.” Washington Post, August 24, 2017; Gressin, Seena. “The Equifax Data Breach: What to Do.” Federal Trade Commission, September 8, 2017.

[83] Segal, Adam. “Rebuilding Trust Between Silicon Valley and Washington.” Council on Foreign Relations. Accessed September 26, 2017.

[84] Cherelus, Gina, and Dustin Volz. “Trump Meets Silicon Valley Elite after Mutual Mistrust in Campaign.” Reuters, December 15, 2016.

[85] Brandom, Russell. “The NSA’s Leaked Windows Hack Caused More Damage than Just WannaCry.” The Verge, May 17, 2017. www.theverge.com/2017/5/17/15655484/wannacry-variants-bitcoin-monero-adylkuzz-cryptocurrency-mining

[86] Protecting Our Ability to Counter Hacking Act of 2017, U.S. Senate, 115th Congress, 1st Session. www.schatz.senate.gov/imo/media/doc/BAG17434_FINAL%20PATCH.pdf

[87] C.f. Goodman, Ryan. “Cyber Operations and the U.S. Definition of ‘Armed Attack.’” Just Security (blog), March 8, 2018. www.justsecurity.org/53495/cyber-operations-u-s-definition-armed-attack/.

[88] Healey, Jason. “The Five Futures of Cyber Conflict and Cooperation.” Georgetown Journal of International Affairs (2011): 110–117.

[89] Definition from Hammoud, Naima, “Game Theory: Minimax, Maximin, and Iterated Removal,” lecture slides. University of Oxford, March 14, 2017.

The Cyber Conflict Studies Association (CCSA) promotes and leads international intellectual development efforts to advance the field of cyber conflict research. These activities include workshops that bring together professionals from industry, academia and government to discuss strategic issues surrounding cyber conflict and the publication of insightful research articles and position papers and books. CCSA also plays an important role in our national cyber-readiness strategy, serving as a resource for national security decision-makers and helping to frame and promote national cyber conflict policy. CCSA brings together the best and the brightest individuals in the field of cyber conflict study to further the goals of national security and the field of cyber.

This work was supported in part by the Minerva Research Initiative. The Minerva Research Initiative, administered jointly by the Office of Basic Research and the Office of Policy at the U.S. Department of Defense, supports social science research aimed at improving our basic understanding of security.

Funded in part by the Office Naval Research under the OSD Minerva program, grant number: N00014-17-1-2423.

Panel Discussion:

Cyber Conflict and Democratic Institutions

By Sean Kanuck

1. Introduction

This year's Global Digital Futures Policy Forum focuses on the tension between fragmentation of the Internet and globalization. While fragmentation, splintering, or "Balkanization" of the Internet has been a prominent topic of discussion for several years now, globalization has recently received a resurgence of attention in popular debate[i]. Globalization – long revered as a teleological objective of the Western liberal order – is increasingly being questioned by electorates in North America and Europe. Rising nationalist tendencies among certain political parties and candidates seek to re-assert domestic advantage and the self-interest of their constituents as their primary political goals. That trend, coupled with the legal debates about privacy and data localization in multiple jurisdictions, has reinvigorated interest in studying fragmented futures for the Internet.

This Panel will address cyber conflict as it pertains to the manipulation and/or compromise of democratic institutions – both directly and indirectly. Direct intervention in a democratic election could comprise either public efforts to personally obstruct voters or else clandestine alteration of actual vote tabulations; indirect intervention could consist of using proxy voices or inducing political, economic, or media events with secondary impacts on voter turnout and election results. Manipulative actions that do not directly alter the voting process or results are to be considered "influence operations", while actual changes to registered voters (including threats of violence or other means to physically deter eligible voters from attending the polls) or the ballots that are cast are typically deemed illegal "voter fraud", even when perpetrated by the state apparatus itself. (Figure 1 below reflects the fact that both direct intervention and indirect influence in democratic elections can be either overt or covert.)

Information communication technologies (ICT) present many new vectors for potentially interfering with democratic institutions. Foreign competitors, traditionally offset by geography, can now impose themselves on domestic political systems anywhere in the world. Social media platforms enable individuals or special interest groups to broadcast their policy positions at little or no cost and even to strategically misrepresent broader support for those positions. Internet-connected ICT networks are highly susceptible to unauthorized access, thereby rendering sensitive data vulnerable to theft and public release. In essence, the digital future – and liberal democratic processes that will rely upon it – is susceptible to interference and disruption. This Panel will consider ways to safeguard democracies and the international order from corruptive influences (or at least to minimize their impacts) in the future.

Figure 1: Examples of Methodologies for Manipulation of Democratic Elections

	DIRECT INTERVENTION	INDIRECT INFLUENCE
OVERT	Intimidating or deliberately misinforming voters in order to deter turn out. For example, unofficial "robocalls" used during the 2011 Canadian federal election to falsely claim changes to polling station locations.[ii]	Public campaign donations and/or speeches by non-candidates in support of specific ballot choices. For example, President Obama's 2016 speech in London opposing "Brexit" before that referendum.[iii]
COVERT	Secretly altering the election results in order to favor a specific candidate. For example, the historical allegations regarding Lucien Bonaparte's inflation of voting results in the French constitutional plebiscite of 1800.[iv]	Clandestine, third-party activity intended to increase or decrease support for specific candidates. For example, reputed Russian espionage and publicization of materials during the 2016 U.S. presidential campaign.[v]

2. Historical Precedent

When evaluating the impact of cyber modalities (i.e. ICT) on democratic institutions, one must first consider what is genuinely new in either the objectives or possible impacts. Regardless of which quadrant of Figure 1 is of concern, there is ample historical precedent from geo-politics. Thucydides recounted Athenian efforts to lobby the magistrates of Melos to capitulate without battle (i.e. indirect and overt influence). Similarly, Radio Free Europe and Voice of America were designed to provide the electorates of foreign polities with information that was otherwise unavailable and/or forbidden. Nor is history want for allegations of ballot-box stuffing (i.e. direct and covert intervention) or voter intimidation (i.e. direct and overt intervention). Digital manifestations of those forms of fraud are certainly illegal and deserving of policy attention, but they are not the focus of recent debate. What seems to capture the current imagination – and concern – is the heightened opportunity for indirect, covert influence through cyber means. Careful analysis is required, however, to properly assess the nature and foundation of that concern.

Framing Question 1: What is so new and inherently objectionable about digital influence campaigns?

If one reasonably acknowledges that foreign efforts to influence elections are as old as elections themselves, then one is left with either (i) a theoretical objection that is so counterfactual to historical practice that it is relegated to pure academic consideration, or (ii) a practical objection that employing a new technological means to an old political end is somehow unacceptable. It is worth recalling that public international law does not outlaw espionage – which is merely accepted as a feature of international relations. Nor is the publication and dissemination of political opinions generally deemed objectionable in liberal democracies. So what is really at issue here?

By way of example, several former U.S. intelligence officials have stated that they considered the theft of Office of Personnel Management records to be a "legitimate" foreign intelligence target.[vi] But even so, U.S. government officials have said that the scale and import of that espionage crossed a line that was unacceptable. So, it would seem that the objection stems from the quantitative scope of the activity in question (i.e. the sheer number of records compromised, the gross imbalance between the cost of conducting the activity versus its harm to the victim, the possible stand-off distance from which such an operation can be conducted without personal risk, etc.), rather than the qualitative nature of the activity itself (i.e. the theft of private information, the type of data targeted, etc.). Chivalric objections to the crossbow and guerilla warfare tactics should immediately come to mind, for new methods of conflict are often too efficacious for the establishment to accept at first outset.

Framing Question 2: When does a quantitative improvement in espionage constitute an unacceptable qualitative change? Do recent offensive cyber advances constitute a qualitative threat to democracy?

<u>Protected Infrastructure</u>

The U.S. Department of Homeland Security did not officially designate election systems as a critical infrastructure until January 2017.[vii] Yet, almost four years earlier in March 2013, the U.S. Director of National Intelligence (DNI) had identified an important incongruity related to how different nation states view online media and their political systems:

> *"Online information control is a key issue among the United States and other actors. However, some countries, including Russia, China, and Iran, focus on 'cyber influence' and the risk that Internet content might contribute to political instability and regime change. The United States focuses on cyber security and the risks to the reliability and integrity of our networks and systems. This is a fundamental difference in how we define cyber threats."*[viii]

That fundamental difference (i.e. the underlying distinction between infrastructure and content) is also germane to the question of which ICT deserve protection as "democratic institutions". Most everyone would likely agree that public authorities must guaranty the security of polling stations, voting machines, and official election returns. In other words, they are expected to prevent direct intervention that is contrary to the rule of law. This is represented by the United States' "infrastructure-centric" view of cyber security that was highlighted by the DNI. Content poses a much more complicated challenge.

Framing Question 3: Is the national government responsible for ensuring the confidentiality, availability, and integrity of all media resources that can influence a democratic electorate? Why not?

The discussion about where to draw the line regarding indirect influence quickly becomes muddied, as we regularly see with proposals for campaign finance reform. Managing the impact of informational content pits two democratic values against one another, namely freedom and equality. How much leverage should freedom of expression permit wealthy individuals and companies to exert on democratic processes? Is every mass media outlet or social media platform to receive a critical infrastructure designation because they can be utilized to influence public opinion? Which entities are "entitled" to special protections and/or restrictions? Each of those questions is a public policy dilemma.

Figure 2: Examples of Civilian Infrastructures that Impact Democratic Elections

	VOTING SYSTEMS	INFORMATION RESOURCES

PUBLIC	Government administered polling stations and officially monitored vote tabulation. Susceptible to corruption by ruling party.	National television, radio, print, and online media outlets. Subject to selective coverage and preferential treatment by ruling party.
PRIVATE	Hardware and software for voting systems and registration databases developed by commercial companies. Susceptible to supply chain and/or remote penetrations.	Independent mass media and online social media platforms. Subject to censorship by government as well as disruption and/or manipulation by third parties.

The status of political parties and their proprietary resources also raises very difficult legal and policy questions. If the compromise of an entity like the Democratic National Committee or the Republican National Committee in the United States is deemed a national security concern, then what level of governmental oversight and regulation of (i.e. access to) that party's ICT networks is appropriate in the national interest? Does that level change depending on whether that party is currently in power? Should smaller political parties be exempt from such regulation if they are not likely targets for foreign intervention? Once again, these cyber challenges are pitting core democratic values against one another (e.g. privacy versus national security) and policy trade-offs are inevitable.

Framing Question 4: Can private data be treated as a national asset against the will of its owner?

Social media represents a uniquely influential and vulnerable feature of modern politics. Its impact during the Arab Spring was noted by governments and demonstrators alike around the world. Since then, the use and manipulation (e.g. "astroturfing" to generate the semblance of broader support) of social media has become an instrumental part of political campaigns, opposition movements, and foreign influence operations. It is possible, at least to a certain degree, to reveal such social media manipulation (e.g. by technically determining the provenance of posted information, detecting automated programs for "re-tweeting" and "liking" posted information, and identifying patterns of coordinated "trolling"), but that requires analysis of large tranches of proprietary data, including both content and technical meta-data. In democratic societies, private ICT companies have no *ex ante* obligation to make their databases available to government authorities for speculative research.

Figure 3: Examples of Information Propagation to Induce Political or Economic Behavior

	INTENTIONAL MESSAGING	UNWITTING EXPLOITATION
INFORM	The 2007 airborne delivery of leaflets over Afghanistan by the U.S. military in order to deter insurgent activity by the Taliban.[ix]	In 2016, Twitter suspended thousands of suspected terrorist accounts that promoted violence and/or spread propaganda.[x]
DECEIVE	Adoption of the title "Bolshevik" (i.e. "one of the majority") by a party faction that was numerically inferior.[xi] The ironic naming of "Greenland" by Erik the Red to encourage emigration to a new colony that was less temperate.[xii]	The Syrian Electronic Army's false "tweet" disseminated from the Associated Press's Twitter account, which led to temporary fluctuations in U.S. stock markets in 2013.[xiii] False news items posted on Facebook during the 2016 U.S. presidential campaign.[xiv]

Data Integrity

As Figure 3 illustrates, many forms of media have been used to spread both information and disinformation for political effect. History is certainly replete with examples of interest groups "marketing" their views to the public – such as the U.S. founding fathers' ascription of the moniker "Anti-Federalists" to their opponents in order to impute a negative connotation – but social media platforms present a new challenge whereby they host content that is neither of their own creation nor necessarily attributable to physically identifiable third-parties. Accordingly, they become enablers for all sorts of online activities that can foster or undermine democratic institutions. That schizophrenia is perhaps best characterized by the hacker consortium Anonymous, which has both thwarted sovereign governments and also publicized child pornographers and corporate fraud.[xv]

Framing Question 5: Is the "common carrier" model the right legal analogy for social media outlets?

All of the themes aforementioned in this paper (e.g. espionage, influence operations, quantitative change, qualitative distinctions, public versus private infrastructure, freedom of expression, national security, etc.) coalesce around the key issue of data integrity. Because democracies rely on the ability of their populaces to make informed decisions, increased dependence on insecure ICT poses considerable threats. How can the public ever differentiate truth from falsehood with certainty?

In fact, international humanitarian law (aka the law or armed conflict) struggles with a similar conundrum when it distinguishes between perfidy (i.e. the illegal intent to betray confidence) and ruses of war (i.e. permissible deceptions not based on garnering false status).[xvi] Interestingly, though, “misinformation” is listed as a ruse vice perfidy; moreover, the relevant treaty distinctions explicitly do not “affect the existing generally recognized rules of international law applicable to espionage.”[xvii] Thus, cyber operations premised on exerting indirect influence are particularly problematic – especially when they only reveal true information.

Framing Question 6: Can two “rights” make a “wrong” … that is, should espionage (which is accepted in international relations) that exposes the truth (a core democratic value) be prohibited?

Ultimately, the most nefarious threat to democratic institutions is the corruption of the integrity of information. The pervasive introduction of false data into mainstream media could erode public confidence and destabilize society. That is, of course, exactly what authoritarian regimes are (i) highly concerned about happening to themselves, and (ii) well-practiced in perpetrating against their adversaries. Yet, democracies pride themselves on permitting their citizens to hold and publicize contrarian (or even counterfactual) opinions, and modern ICT permit foreign voices to participate in domestic dialogues.

It seems then that the most conceptually disturbing challenge for democratic institutions regards digital, highly efficient, indirect, foreign, misinformation campaigns that can neither be prevented nor easily identified. Furthermore, it is unclear what kind of governmental institutions (domestic or international) and/or private sector initiatives could resolve that difficulty, for this seemingly new cyber concern tautologically reduces to the well-known game theory paradox of “who guards the guardians”?

iSee http://www.atlanticcouncil.org/images/files/publication_pdfs/403/121311_ACUS_FiveCyberFutures.pdf ; See generally, David Kennedy, A WORLD OF STRUGGLE: HOW POWER, LAW, AND EXPERTISE SHAPE GLOBAL POLITICAL ECONOMY, Princeton University Press (2016).

ii See http://news.nationalpost.com/news/canada/canadian-politics/electoral-fraud-did-take-place-in-2011- federal-vote-but-it-didnt-affect-outcome-judge-rules

iii See https://www.theguardian.com/politics/2016/apr/22/barack-obama-brexit-uk-back-of-queue-for-trade-talks

iv See e.g., https://en.wikipedia.org/wiki/French_constitutional_referendum,_1800

v See https://www.nytimes.com/2016/12/13/us/politics/russia-hack-election-dnc.html?_r=0 vi See http://www.defensenews.com/story/defense/policy-budget/cyber/2015/06/27/opm-attack-hack-china- cybersecurity-personal-data-suspect-espionage-verifiable-/29341789/; See https://www.the-american- interest.com/2015/06/16/former-cia-head-opm-hack-was-honorable-espionage-work/

vii See https://www.dhs.gov/news/2017/01/06/statement-secretary-johnson-designation-election-infrastructure- critical

viii James R. Clapper, Statement for the Record, Worldwide Threat Assessment of the US Intelligence Community, Senate Select Committee on Intelligence, March 12, 2013

ix See http://www.af.mil/News/Article-Display/Article/127729/operation-achilles-leaflet-airdrop-delivers-message- to-taliban/

x See https://www.wired.com/2016/08/twitter-says-suspended-360000-suspected-terrorist-accounts-year/

xi See https://www.britannica.com/topic/Bolshevik; See http://www.historytoday.com/richard-cavendish/bolshevik-menshevik-split

xii See http://news.nationalgeographic.com/2016/06/iceland-greenland-name-swap/; See also, https://www.scientificamerican.com/article/proof-on-ice-southern-greenland-green-earth-warmer/

xiii See https://www.washingtonpost.com/news/worldviews/wp/2013/04/23/syrian-hackers-claim-ap-hack-that- tipped-stock-market-by-136-billion-is-it-terrorism/?utm_term=.f575e36dfcd2

xiv See http://www.reuters.com/article/us-usa-election-facebook-idUSKBN1380TH

xv See https://sg.finance.yahoo.com/news/Anonymous-exposes-visitors-afpsg-2809071407.html; See http://asia.nikkei.com/Business/Trends/Hackers-turn-stock-advisers-as-Anonymous-targets-China-Inc?page=1

xvi See Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (hereinafter Protocol 1), Article 37, June 8, 1977; See also, Protocol 1, Article 39

xvii Protocol 1, Article 39(3); See Protocol 1, Article 37(2)

COLUMBIA|SIPA
School of International and Public Affairs

Normative Restraints on Cyber Conflict

March, 2017

By Joseph S. Nye, Jr.

1. Introduction

Where does the world stand in the development of norms to restrain conflict in cyber space? Elsewhere I have compared learning about cyber security with the way states learned to cooperate in regard to nuclear weapons. ("Nuclear Lessons for Cyber Security," *Strategic Studies Quarterly*, Winter, 2011). While cyber and nuclear technologies are vastly different in their characteristics and effects, at a meta level, the processes of how societies and states learn to cope with a highly disruptive technology have interesting similarities. In terms of chronology, it took states about two decades to reach the first cooperative agreements to limit conflict in the nuclear era. If one dates the cyber security problem not from the beginning of the Internet in the 1970s but from the period since the late 1990s when burgeoning participation made the Internet a substrate for economic and military interdependence (and thus vulnerability), cooperation in cyber is now at about the two decade mark.

The first efforts in the nuclear era were unsuccessful UN centered treaties. In 1946, the US proposed the Baruch plan for UN control of nuclear energy, and the Soviet Union promptly rejected locking itself into a position of technological inferiority. It was not until after the frightening Cuban Missile Crisis, that a first arms control agreement, the Limited Test Ban Treaty was signed in 1963. The NPT followed in 1968 and the bilateral Strategic Arms Limitation Treaty in 1972. In the cyber field, in 1999, Russia proposed a UN treaty to ban electronic and information weapons (including propaganda). With China and other members of the Shanghai Cooperation Organization, it has continued to push for a broad UN based treaty. The US resisted what it saw as an effort to limit American capabilities, and continues to view a broad treaty as unverifiable and deceptive. Instead, the US, Russia and thirteen other states agreed that the Secretary General should appoint a Group of Government Experts (UNGGE) which first met in 2004. It initially had meager results, but by July 2015 it issued a report which proposed norms for limiting conflict as well as confidence building measures that was endorsed by the Group of 20 summit. Groups of experts are not uncommon in the UN process, but only rarely does their work rise from the basement of the UN to a summit of the twenty most powerful states. The success of this group was above the ordinary.

2. The UN Group of Government Experts

The GGE issued reports in 2010, 2013 and 2015 that have helped to set the negotiating agenda for cybersecurity, but despite this initial success, the GGE has limitations. The participants are technically advisors to the Secretary General rather than fully empowered national negotiators, and although their number has increased from the original 15 to 20 to 25, most nations do not have a voice. According to one diplomat who has been central to the process, some seventy countries have expressed interest in participating. But as the numbers expand, the problems of reaching agreement increases. Some observers worry that entropy will set in and they express concern whether this process can continue to succeed.

To understand the GGE, it helps if one puts it in a broader context of normative constraints upon states. The three canonical sources of international law are treaties, customary international law, and expert juridical opinion. Some observers draw a sharp distinction between international law and international norms. The Tallinn Manual, for example, is an important effort by a group of international lawyers to write down what is agreed to be international law. it is clear that lawyers do not always agree, but on many matters they do agree on law that is supposed to be binding on states. A norm, as distinguished from law by Martha Finnemore and Duncan B Hollis,("Constructing Norms for Global Cybersecurity," 110 *American Journal of International Law*, 2016) is a collective expectation of proper behavior of actors with a given identity. Norms apply to multiple actors and are not legally binding. "Laws can serve as a basis for formulating norms, just as norms can be codified by law."(p442) Norms play a role in constituting new roles as well as constraining existing ones. The "oughtness" of their constraints can grow out of law, politics and cultures.

Parsing the differences between laws, norms and other types of constraints is sometimes useful but it is not my purpose here. By lumping together a wide range of normative constraints, I want to illustrate nine potential arenas for action in the following matrix. Horizontally, in terms of formalism, normative constraints on states range from formal treaties to conventional state practice to codes of conduct and norms. Vertically, in scope of membership, the groups thus constrained can range from global, to plurilateral, to bilateral. Such groups can include both states and non-state actors. The totality can also be described as a regime complex.

3. Normative Constraints on States and Non-State Actors

	Agreements	State Practice	Norms and codes
Global	ICANN	Routing practices and exchanges	UNGGE
Plurilateral	Budapest Convention	Like minded groups	G 20, OSCE Regional orgs.
Bilateral	US/China on commercial CNE	Entanglement and self restraint	CBMs, US-Russia hot line

Non-state actors can be constrained by domestic law, punishment, culture, but in a world without overarching international government, why do sovereign states themselves sometimes let normative considerations constrain their behavior? Among the considerations, one reason is fear. Another is external reputation. A third is domestic political pressure.

4. Fear, Prudence and Norms

What can history tell us about the effectiveness of these normative instruments of policy in other areas? In the decade after Hiroshima, tactical nuclear weapons were widely regarded as "normal", and the U.S. military incorporated nuclear artillery, atomic land mines and nuclear anti-aircraft into its deployed forces. In 1954 and 1955, the Chairman of the Joint Chiefs of Staff told President Dwight Eisenhower that the defense of Dien Bien Phu in Vietnam and the defense of offshore islands near Taiwan would require the use of nuclear weapons, but Eisenhower rejected the advice in part because of fear of unintended consequences. (See my "Deterrence and Dissuasion in Cyber Space," *International Security*, Winter 2017).

Over time, this prudence developed into a norm of non-use of nuclear weapons which has added to the cost that a decision maker must consider before taking an action to use them. The Nobel Laureate economist Thomas Schelling argued that the development of a norm of non-use of nuclear weapons was one of the most important aspects of arms control over the past 70 years. Ironically, Eisenhower (and other leaders) was unwilling to sign onto a formal norm of no-first use of nuclear weapons because the residual uncertainty of potential use was needed to deter Soviet superiority in conventional forces. It was not until the era of Gorbachev and Reagan that leaders were willing to agree that nuclear war could not be won and must never be fought. The norm of non-use has had an inhibiting effect on leaders of major states, but for new nuclear states like North Korea, one cannot be sure whether the costs of breaking the taboo would be perceived as outweighing the benefits.

In cyber, fear of destroying the benefits reaped from the Internet (which are increasingly important to economic growth) may constrain attacks on the Domain Name System or the IANA function. In addition, the very newness of cyber war and fear of unforeseen consequences in unpredictable systems may contribute to prudence that could develop into a norm of non-use or limited use or limited targets. As Brandon Valeriano and Ryan Manness point out in *Cyber War vs. Cyber Reality* (Oxford University Press, 2015), on a number of occasions when faced with a choice in wartime, political and military leaders have preferred the predictability of kinetic weapons. Sometimes fear of unintended consequences can lead to prudence which can develop into a norm.

5. External Reputation

After World War I, a consensus taboo developed about poisons, and the 1925 Geneva Protocol prohibited the use (though not possession) of chemical and biological weapons. They existed but were not used in World War II because of deterrence through fear of retaliation. Then in the 1970s, two treaties were negotiated that prohibited the production and stockpiling of such weapons. That meant that there is a cost associated not only with their use but even their very possession. Verification provisions for the Biological Warfare Convention are weak (merely reporting to the UN Security Council), and such taboos did not prevent the Soviet Union from cheating by continuing to possess and develop biological weapons in the 1970s. The Chemical Weapons Convention did not stop either Saddam Hussein or Bashir al Assad from using chemical weapons against his own citizens, but they did have an effect on the perceptions of costs and benefits of actions, such as the international dismantling of most Syrian weapons in 2014. With 173 states having ratified the Biological Warfare Convention, states that wish to develop biological weapons have to do so secretly and illegally and face widespread international condemnation if evidence of their activities leak. External reputational harm, along with uncertain benefits in use, appear to be the main reasons that norms seem to have limited possession such weapons.

Normative taboos may become relevant in the cyber realm as well, but not against mere possession of weapons. The difference between a computer program that is a weapon and a non-weapon depends on intent, and it would be difficult to forbid the design, possession, or even implantation for espionage of particular programs. In that sense, cyber arms control cannot be like biological arms control or the nuclear arms control that developed during the Cold War which involved elaborate detailed treaties regarding verification. Unlike physical weapons, it would be impossible to reliably prohibit possession of the whole category of cyber weapons.

A more fruitful approach to normative controls on cyber arms is not to focus a taboo against *weapons* but against *targets*. The United States has promoted the view that the internationally recognized Laws of Armed Conflict (LOAC) which prohibit deliberate attacks on civilians apply in cyber space. Accordingly, the U.S. proposed not a pledge of "no first use" of cyber weapons, but a pledge of no use of cyber instruments against civilian facilities in peacetime.

This approach to norms was adopted by the GGE. The taboo would be reinforced by confidence building measures such as promises of forensic assistance and non-interference with the workings of Computer Security Incident Response Teams (CSIRTs). The GGE report of July 2015 focused on restraint on attacks on certain civilian targets rather than proscription of particular code. At the 2015 summit between American President Barrack Obama and China's President Xi Jinping, the two leaders agreed to set up an expert commission to study the GGE proposal (as well as a separate agreement limiting cyber espionage for commercial purposes). As noted above, the GGE report was endorsed by the leaders of the G-20 and referred to the UN General Assembly. On the other hand, an attack on the Ukrainian power system occurred in December 2015, and was widely attributed to Russia, a GGE member (though Russia might argue that given its hybrid war with Ukraine, it was not bound by a peacetime norm.) Similarly, in 2016, the U.S. accused Russia of using cyber means to interfere in the American election. Despite the fact that the US had added electoral processes as a 17th item on its list of critical infrastructures, Russia clearly did not include the election process in the U.S. as a critical civilian infrastructure covered by the taboo. At this point the development of normative controls on cyber arms remains a slow and incomplete process. In general, the multi-lateralization of norms helps raise the reputational costs of bad behavior. It is worthy of note that the Missile Technology Control Regime and the Proliferation Security Initiative began as voluntary measures and gathered momentum, members, and normative strength over time.

6. Domestic Factors

There is a third process which can lead to statesmen accepting normative constraints on their actions and that arises out of domestic politics. In cyber as in other domains, theorists like Martha Finnemore and Kathryn Sikkink ("International Norm Dynamics and Political Change," *International Organization* 1998) have hypothesized that norms have a life cycle starting with norm entrepreneurs, tipping points into cascades, and then internalization which translate their effects into beliefs that have domestic costs that deter external actions. If one looks at the historical development of norms against the slave trade in the 19th century or in favor of human rights in the second half of the 20th century, one can see that some states are constrained by the effect of norms on domestic opinion. Of course, one would expect such constraints to be stronger in democracies than in authoritarian states (though not totally absent in the latter – witness the effects of Basket Three of the Helsinki Process). Today, in cyber norms the world is largely at the first stage with the GGE as one of a number of important norm entrepreneurs. Perhaps norms are beginning to enter the second phase of a cascade. But the internalization of norms remains weak and limited to narrow elites. Moreover, there is no metric for measuring time in this hypothesized cycle, and indeed no guarantee of a cycle at all. For example, if relations between states become bitter over all, retrogression is certainly possible.

7. Next Steps

There is a wide range of views about the next steps for the GGE process. A first draft of a new report existed at the beginning of this year, but it was a long way from agreement. At the February 2017 Munich Security Conference, the current chair argued that the group should not try to rewrite the 2015 report, but should say more about the steps that states should take in peacetime. Some states suggested new norms dealing with data integrity and maintenance of the core structures of the Internet, but other states believed such expansion would open up a Pandora's box. There was general agreement about more discussion of confidence building measures and of capacity building, but also concern about how states will implement what has already been agreed.

If the GGE norms are to "cascade", states must raise awareness in a broader public. It is noteworthy that the Ukrainian disruption was not flagged and debated as possibly contrary to the GGE report of 2015. A representative of a small country argued that international law was crucial to small states without power, and made the case for more attention to the Tallinn Manual 2.0. The representative of a major power said the GGE should dig deeper on questions such as what is meant by civilian processes. A UN under-secretary argued that the norm development process had to be broadened to include more countries to increase its legitimacy among the 193 UN members, and should relate cyber to other issues such as arms control in space and terrorism. In his view, the 5th GGE should dig deeper and then the 193 members of the UN should debate the report and task the next GGE to examine specific areas.

The GGE process reflects the positions of the states that nominate the experts and their strong views on state sovereignty. Certain normative issues are not discussed. The questions of contents and human rights are finessed by saying that all states agreed to the Universal Declaration of Human Rights though they interpret and implement it in different ways. Further progress on such subjects would probably be limited to plurilateral discussions among like-minded states rather than universal agreements. Other norms that may be ripe for discussions outside the GGE process could include a protected status for the core functions of the Internet; supply chain standards and liability for the Internet of Things; treatment of election processes as protected infrastructure; and more broadly norms for sub-LOAC issues such as crime and information warfare. All these are among the topics that may be considered by the new informal International Commission on Stability in Cyberspace announced by the Dutch Foreign Minister at Munich.

As member states contemplate next steps in the development of cyber norms they are faced with the dilemma of maintaining the effectiveness of the GGE while expanding participation in order to develop a broad legitimacy for norms that will help them to cascade and internalize. The answer may be to avoid putting too much burden of a burden on any one institution like the GGE. Norms are affected by their institutional homes, and in the long run many homes may be better than one. Progress on the next steps of norm formation may require simultaneous use of many of the nine cells for action identified in the matrix above. It will also require a strategy for mutual reinforcement among the cells. For example, the bilateral agreement between China and the US on cyber espionage for commercial purposes was taken up by the G20 as well in bilateral negotiations between China and a number of other states. In some instances, development of norms among like-minded states can lead to norms to which others may accede at a later point. In other instances, norms for security on the Internet of Things may benefit from codes of conduct where the private sector or non-profit stakeholders take the lead. And progress in some areas need not wait for others. The development of a regime complex may be more robust when linkages are not too tight. (See my “The Regime Complex for Managing Cyber Activities,” Research Paper #1, The

Global Commission for Internet Governance, 2014). Such flexibility would be incompatible with an over-arching UN treaty at this point. Expansion of participation is important for the acceptance of norms, but progress on norms will require action on many fronts. We are still in the early stages in the formation of normative constraints on cyber activity.

Section 3:
Digital Economy

Digital Trade, E-Commerce, the WTO and Regional Frameworks

Merit E. Janow & Petros C. Mavroidis

1. Digital Trade and Trade Agreements

The digitalization of trade is a reality, and yet the regulation of the world trading system as embedded in the World Trade Organization (WTO) only tangentially, if at all, touches upon this issue. True, digitalization of the economy, the fourth industrial revolution as it is colloquially referred to, is a recent phenomenon, and to some extent post-dates the conclusion of the Uruguay round agreements (1994). True also, however, is the reality that the the world trading system has shown a remarkable inability to adjust to modern business realities in its multilateral rule architecture. To the extent that these transformations are reflected in new rules, such are being introduced in regional or bilateral frameworks, albeit in an incomplete fashion. It is also the case that the world is witnessing several different regimes around data and information developing in the world today—most notably in the US, Europe and China. As always, part of the reason that international frameworks have not been born stems from the fact that international rules rarely occur before domestic regulatory and legal regimes are well developed.

In a world where the regulation of information and digital within national systems is in flux, and major jurisdictions are in tension, international frameworks might be expected to develop between jurisdictions that have the most confidence in each other, the most experience or the greatest trade flows. At the same time, multilateral frameworks can offer the greatest transparency and predictability to the largest number of countries and stakeholders. The essays in this volume help us consider those possibilities. We start, however, with the observation that nothing much has happened at the WTO on digital trade with the exception of a recurring decision, adopted during practically every

Ministerial Conference, to continue exempting from trade restrictions products traded electronically (the so-called “e-commerce” decisions). In essence, there is a renewed commitment every two years to continue negotiating, and every two years a recognition to the effect that nothing much has happened in the previous two years. Moreover, in very recent months, in a period when the WTO is facing several institutional crises about its future operations--there are a few new signs of interest around digital matters. For example, there has been an expansion of coverage on digital trade in the recently negotiated new North America trade agreements, and new discussions among some WTO members about expanded WTO efforts around digital trade.

What do WTO rules and cases tell us today? We believe there are important foundations. For example, technological neutrality, a principle now well-established in WTO case law, obliges WTO members to treat like services in an even-handed manner irrespective of the means of supply. In US-Gambling, the Appellate Body found that, the absence of specific commitment with respect to Internet gambling notwithstanding, the United States still had an obligation to accept Internet gambling since it had promised to impose no barriers under Mode 1.

Yet this hardly answers the question of how to think about the many forms of digital commerce and consequences. What about 3D printing and all other products that come into being through digital cross border transference? As things stand, there is at least uncertainty as to their treatment under WTO law. WTO rules are predicated on an absolute dichotomy between those dealing with goods- and those focusing on services trade. And yet, 3D printing tests the legitimacy of this distinction. Think of a US company engaging in 3D for a client in Switzerland. Is a service being exported when recourse to 3D is made, or a good being imported? Digital trade, more than anything else, tests the legitimacy of the distinction between the GATT and the GATS. Staiger (2018) expressed recently, similar thoughts on this score.

The papers included in this volume cover many important dimensions of digital trade—from an assessment of the scale and scope of digitalization and cross border developments to particular rules covered in regional or other arrangements. Our invited authors explain what Marsh (2012) first termed “the new industrial revolution”, and the challenges it poses for the world trading system, as we now know it. It took time for the digital economy to have an impact on the real economy, but it is now being increasingly felt.

The digital economy is altering patterns of production and patterns of trade. The WTO has yet to address the issue of digital trade in comprehensive manner. For now, all we have at the WTO level is a Moratorium exempting goods and services traded digitally from duties. The absence of multilateral movement, however, does not mean that countries have given up negotiating trade agreements and, in fact, digital trade provisions are appearing in various frameworks—mostly regional. Free-trade areas (FTAs) continue to multiply, and digital trade consistently figures across the subject areas negotiated therein, albeit with varying degrees of specificity and coverage. Remarkably, it features not only when FTAs are being negotiated across “homogeneous”, advanced players (those possessing digital technology), but also across “heterogeneous” players, contributing thus to the narrowing of the “digital divide”.

Some of the papers in this collection focus on the significant conceptual issues that are being triggered as a result of digitalization and cross border trade. For example, Anupam Chander and Joshua Meltzer emphasize the privacy and security issues that are raised by digital trade and the emerging world of connected devices and the internet of things. Of these two authors, Chander may be somewhat more explicitly optimistic about the adaptive characteristics of the WTO system. Meltzer’s focus is more explicitly on the significance of the issues for domestic regulatory regimes. A number of other papers in the collection discuss the regulation of digital trade in the realm of FTAs, both with respect to their specific features and, by implication, the potential consequences for the world trading system. This type of comparison is important because the experimentation on rule frameworks is most robust through these bilateral and regional frameworks. Wolfe observes, for example, that digital trade is an example of how states are in fact learning how to solve the problem of state responsibility while allowing 21st century commerce to flourish. Our readers may be left wondering why the discussions at the WTO level remain nascent. The essay by Usman Ahmed, offers some insight and direction.

The papers in this volume discuss the challenges for the WTO regime, but do not deal in comprehensive manner with the reasons behind the inability to establish a multilateral rule framework at the WTO or elsewhere. Indeed, this could be a quixotic task as the reasons vary from stakeholder preferences within WTO members to the more general challenge that digital trade rules (and many other issues) are hostage to a certain a generalized inertia around negotiations and reform. With this in mind, we now turn to a more detailed presentation of the featured papers.

2. Presentation of the Papers

We divide the papers into three categories: those dealing with conceptual issues, those detailing the regulation of digital trade in FTAs, and finally, the papers aiming to draw lessons from this discussion for the WTO.

2.1 Conceptual Issues

Anupam Chander underlines the need for regulating the internet. In his view, even while offering substantial improvements in our lives, the Internet of Things will require significant regulatory oversight. He explains how ubiquitous smart objects will raise questions of privacy, security, standards, and interoperability. The coming of this smart world, he argues, will also put pressure on trade law, as dispute settlement mechanisms are invoked to assess whether a particular government measure is legitimate regulation or simply disguised protectionism. The coming of the Internet of Things complicates the elegant distinctions at the heart of international trade law, particularly between goods and services, while it reveals that trade law always recognized the complexity of a world where goods embedded services. And it further reveals, he claims, that the international trade regime may yet prove more adaptable than might have been expected.

Joshua Meltzer explains why the inherently global nature of the internet is in tension with regulation that is typically focused on domestic goals: the regulatory challenge is to find ways for cross-border data flows and local regulation to co-exist, as governments' willingness to commit to digital trade rules will be affected by the impact of such rules on the achievement of domestic regulatory goals. In the privacy context, for example, the author notes that the uncertainty as to how EU (European Union) personal data is protected in third countries has led to the restrictions on transfers of personal data outside of the EU. The EU approach to privacy has led to hesitancy by the EU in accepting commitments to cross-border data flows in the Trade in Services (TiSA) negotiations and in the U.S.-EU Transatlantic Trade and Investment Partnership (TTIP) negotiations. EU hesitancy with digital trade rules is also reflected in the recently finalized Japan-EU FTA, which does not include commitments to cross-border data flows, but does include a commitment to revisit this issue within three year of entry into force of the agreement. [1] He observes that balancing between international trade commitments and their impact on domestic regulatory flexibility is not a challenge specific to digital trade, of course. Indeed, this has been a key challenge for WTO jurisprudence over the last 20 years. In the digital trade context, the challenge is at once more acute and less bounded. It is more acute because cross-border data flows are happening more rapidly and in greater quantities than has occurred with trade in goods or services. This raises the prospect for regulators that cross-border data flows have greater potential to undermine the achievement of domestic regulatory goals than traditional trade in goods. This is a challenge to which the WTO eventually will have to respond.

Norman Zhang poses a hypothetical WTO challenge to the Passenger Name Records (PNR) Transfer Agreements the European Union has signed with the United States (as well as Australia and Canada). These agreements ask, head on, the question of

[1] http://trade.ec.europa.eu/doclib/press/index.cfm?id=1891

whether national security-related concerns can adequately be taken care of within the current WTO regime, as embedded in the GATS (General Agreement on Trade in Services). The focus, in his view, will be on a possible citation of GATS Art. XIV National Security Exception by the EU, and the viability of such a defense. Because of the absence of case law on this issue so far, Zhang in his paper attempts to synthesize an acceptable standard for assessing a GATS National Security Exception citation.

2.2 Digital Trade in FTAs

Robert Wolfe states that it is a truth universally acknowledged that every ambitious 21st century trade agreement is in want of a chapter on electronic commerce. In this context, one of the most politically sensitive and technically challenging issues is personal privacy, including cross-border transfer of information by electronic means, use and location of computing facilities, and personal information protection. States are learning to solve the problem of state responsibility for something that does not respect their borders while still allowing 21st century commerce to develop. He then performs a comparison of the Canada-European Union Comprehensive Economic and Trade Agreement (CETA) and the Trans-Pacific Partnership (TPP). In his view, this comparison allows us to see the evolution of the issues thought necessary for an e-commerce chapter, since both include Canada. The comparison also allows us to see the differing priorities of the U.S. and the EU, since they are each signatory to one of the agreements, but not to the other. He concludes by identifying a few important generalizations about why we see a mix of aspirational and obligatory provisions in free trade agreements. Wolfe suggests that the reasons are that governments are learning how to work with each other in a new domain, and learning about the trade implications of these issues.

Evan Kim discusses the e-commerce chapters in South Korea’s FTAs, which cover a wide range of issues, ranging from non-discrimination to electronic signatures. Across the agreements, the country’s provisions on consumer protection, paperless trading, and data protection are uniquely consistent, while those on other issues are not. With the aid of a framework (Framer v. Follower) that captures the dynamics of bilateral negotiations, he argues that in Korea’s case, the more consistent the particular set of provisions is portfolio-wide, the more likely it was for Korea to have prioritized the relevant issue and actively pushed its preferred terms in the FTAs. He provides an overview of Korea's e-commerce chapters, and he explains his Framer-Follower framework that drives his main analysis. He then analyzes each issue included in Korea's e-commerce chapters using the framework.

2.3 Lessons for the WTO

Usman Ahmed deplores the current state of affairs at the WTO. In his view, there is real urgency to begin to tackle some of the challenges of digital trade through regulatory cooperation. Trade negotiators are trying to tackle some of these problems, but are limited by the structure, the processes, and the history of the trade regime. Regulatory cooperation arrangements no doubt have shortcomings, but when they result in a mutual recognition agreement they can address fundamental regulatory challenges with a concrete and enforceable regime. He argues that mutual recognition agreements are not easy to achieve due to concerns about information asymmetry, but technology, transparency, and vigilance can help to improve trust and enable domestic regulators to approve the processes of their foreign counterparts. The proliferation of mutual recognition agreements on thorny issues related to digital trade could help bring certainty to the digital ecosystem and unlock the full growth potential of the digital economy.

Neeraj R.S. attempts to explore the challenges in and possibility of situating a multilateral digital trade agreement within the legal framework of the WTO. He first discusses the broad challenges that digitization poses for the international legal framework for trade regulation. He argues that the traditional classification of products into goods and services under the WTO system is structurally incompatible with the digital economy. He also argues that striking the appropriate balance between trade liberalization and the pursuit of legitimate public policy objectives in a digital trade agreement will be uniquely challenging, since certain features that are intrinsic to the digital industry and market structure require a treatment that is fundamentally different from the "balancing methods" used in other multilateral agreements that the WTO has facilitated. He then surveys the efforts that have been undertaken to regulate digital trade as manifested in FTAs, as well as proposals made by WTO members under the WTO Work Program on Electronic Commerce. Acknowledging that the Trans-Pacific Partnership (TPP) agreement is being used as a benchmark while developing rules to regulate digital trade, he argues that future negotiations for a multilateral digital trade policy will not benefit from using the TPP as a benchmark. The TPP does not, in his view, reconcile systemic tensions between the digital economy and the extant WTO system, or address the domestic regulatory challenges that are unique to the digital ecosystem while trying to achieve a balanced outcome.

3. Brief Concluding Remarks

The post second world war history of multilateralism has seen a gradual expansion of international rules beyond tariff reductions to increasingly internal areas of economic activity. Digital trade is now implicating still more areas which, importantly, are sensitive for nations and places where national regulatory approaches differ markedly and are in flux. This is occurring in the context of the failure to advance even the well-identified WTO negotiating agenda for the post Uruguay Round period, let alone the areas such as digital trade and ecommerce which are well recognized but raise complex and new conceptual and practical issues. The essays in this volume help us consider some of the areas around which there might need to develop greater shared understandings and approaches if multilateral rules are to develop. Importantly the essays also clarify some of the experimentation that is occurring in regional arrangements. These essays underscore the urgency of action, the characteristics of frameworks developed to date, and by implication the importance of these for the multilateral system.

References

Marsh, Peter. 2012. The New Industrial Revolution: Consumers. Globalization, and the End of Mass Production, Yale University Press: New Haven, Connecticut.

Staiger, Robert W. 2018. On the Implications of Digital Technologies for the Multilateral Trading System, p. 150 in World Trade Report 2018, The World Trade Organization: Geneva, Switzerland.

TRADE RULES FOR THE DIGITAL ECONOMY: CHARTING NEW WATERS AT THE WTO

NEERAJ RS[*]

Abstract

This Article attempts to explore the challenges in and possibility of situating a multilateral digital trade agreement within the legal framework of the World Trade Organization (WTO).

Part I of the article discusses the broad challenges that digitization poses for the international legal framework for trade regulation. I argue first that the traditional classification of products into goods and services under the WTO system is structurally incompatible with the digital economy. I also argue that striking the appropriate balance between trade liberalization and the pursuit of legitimate public policy objectives in a digital trade agreement will be uniquely challenging since certain features that are intrinsic to the digital industry and market structure require a treatment that is fundamentally different from the "balancing methods" used in other multilateral agreements that the WTO facilitated.

Part II surveys the efforts that have been undertaken to regulate digital trade as manifested in Free Trade Agreements (FTAs) and proposals made by WTO members under the WTO Work Program on Electronic Commerce. Acknowledging that the Trans-Pacific Partnership (TPP) agreement is being used as a benchmark while developing rules to regulate digital trade, I argue that future negotiations for a multilateral digital trade policy will not benefit from using the TPP as a benchmark. The TPP does not reconcile systemic tensions between the digital economy and the extant WTO system or address the domestic regulatory challenges that are unique to the digital ecosystem while trying to achieve a balanced outcome.

[*] LL.M Candidate, Columbia Law School; Email: nr2647@columbia.edu. The author wishes to acknowledge Professor Merit Janow, Dean School of International and Public Affairs, Columbia University and Professor Petros Mavroidis, Edwin B. Parker Professor of Foreign and Comparative Law at Columbia Law School for their guidance and helpful insights on the article. All errors and omissions that remain are the author's own.

Working Paper, March 20, 2018

THE INTERNET OF THINGS: BOTH GOODS AND SERVICES

Anupam Chander[*]

International trade law, organized around the goods-services dichotomy, is about to meet the Internet of Things. How will rules written for a world of 1994 fare in a world of talking teapots and connected cars? How do we fit smart objects within the classification schemes devised a quarter-century ago?[1]

With the advent of the Internet of Things, not only do things cross borders, so do streams of data over the lifetimes of those things. Accordingly, the Internet of Things implicates both physical and virtual borders—both customs clearance and information regulation. Should international trade law treat the Internet of Things (IoT) strictly as goods, whatever their intelligence or capabilities? Should the data flows of IoT be seen as communications, and not services? The answer to these questions will have significant impact on the course of trade and the global distribution of manufacturing and services in the years to come.

The Internet of Things has been defined as "a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies."[2] It consists of devices that interconnect with the world electronically, transmitting and receiving data and modifying their actions accordingly. The IoT represents the emergence of a smart environment, where robots and inanimate objects can monitor, interpret, and affect our physical surroundings. It

[*] Director, California International Law Center, and Martin Luther King, Jr. Professor of Law, University of California, Davis School of Law. A.B. Harvard College; J.D. Yale Law School. Thanks to Merit Janow and Petros Mavroidis for inviting me to participate in their Columbia University digital trade seminar where I received insightful comments. I am grateful for a Google Research Award that supported this work. The work builds on earlier research conducted for UNCTAD. The views expressed herein, and any errors, are my own.

[1] In fact, the country schedules principally rely on a classification system that dates to 1991, when the United Nations published the provisional Central Product Classification scheme. Rolf H. Weber & Mira Burri, Classification of Services in the Digital Economy 19 (2012).

[2] Recommendation ITU-T Y.2060.

entails the embedding of intelligence into everything from "streetlights to seaports."[3] Estimates suggest some 20 billion IoT devices (what I will call "smart objects") will be deployed by 2020.[4] Soon smart objects will far outnumber humanity (though their uneven distribution will unfortunately create a new digital divide). Increasingly, the goods that have long been the subject of international trade are becoming embedded with tiny computers that sense their surroundings and communicate about what they see.

Governments across the world have embraced the Internet of Things, with countries from Brazil to India committing to smart cities. Even as they recognize the possible benefits of IoT, many governments are also beginning to observe that the IoT presents significant privacy and security issues, as well as questions regarding standards and interoperability.[5] Governments, acting upon these important concerns, will thus find it necessary to regulate the Internet of Things. Such regulatory activity will create opportunities to favor local businesses over foreign providers, and thus bring to bear scrutiny of such regulations for compatibility with international trade law.

While no trade dispute thus far has involved IoT, international conflicts around IoT are brewing. Take three examples of nations that have taken steps against foreign IoT manufacturers. In August 2014, China banned its ministries and federal agencies from purchasing Apple iPads and MacBooks, a ban that seems to have been

[3] Daniel Burrus, The Internet of Things Is Far Bigger Than Anyone Realizes (2014), available at https://www.wired.com/insights/2014/11/the-internet-of-things-bigger/. For examples of IoT deployment in a variety of settings, see US Gov't Accountability Office, Internet of Things : Status and implications of an increasingly interconnected world 62-64 (2017), https://www.gao.gov/assets/690/684590.pdf.

[4] https://www.gartner.com/newsroom/id/3598917. Additional estimates available at http://spectrum.ieee.org/tech-talk/telecom/internet/popular-internet-of-things-forecast-of-50-billion-devices-by-2020-is-outdated; International Telecommunications Union (ITU) and Cisco, Harnessing the Internet of Things for Global Development 11 (2016).

[5] Federal Trade Commission, Internet of Things: Privacy and Security in a Connected World (2015); U.S. Department of Commerce, Fostering the Advancement of the Internet of Things (2017), https://www.ntia.doc.gov/files/ntia/publications/iot_green_paper_01122017.pdf. http://meity.gov.in/sites/upload_files/dit/files/Draft-IoT-Policy%20%281%29.pdf.

rescinded later.[6] China represents Apple's second-largest market, after the US.[7] The exclusion was apparently based not on concerns about the hardware, but followed upon a review of security and privacy issues. In January 2017, the United States Federal Trade Commission brought a lawsuit against the Taiwanese smart object maker D-Link for alleged inadequate security in its devices, alleging that D-Link advertised "Advanced Network Security" but failed to adequately secure its wireless routers and internet cameras, leaving their consumers at risk of hacking.[8] In August 2017, the United States Army banned the use of drones made by leading Chinese drone-maker DJI over security concerns.[9] In each case, governments have taken adverse measures against foreign IoT suppliers based not on the hardware, but on the digital features of the products.

Classification is critical to the application of World Trade Organization (WTO) agreements. Classification determines what trade rules can be brought to bear on any controversy involving IoT trade.[10] If we determine that IoT consists in goods, then the General Agreement on Tariffs and Trade (GATT), as well as the Agreement on Technical Barriers to Trade (TBT), will discipline trade barriers to the flow of goods. If we determine that IoT consists in services, then the General Agreement on Trade in Services (GATS) will apply, though generally to different barriers than those covered by GATT. I will argue here that IoT consists in both goods and services, therefore calling into application multiple WTO disciplines, with the specific agreements that are applicable dependent on the particular measure subject to challenge. While my focus is on the WTO, much of the arguments will apply, *mutatis mutandis*, to bilateral and regional free trade agreements, which also adopt the goods/services dichotomy.

[6] *China Said to Exclude Apple From Procurement List*, BLOOMBERG NEWS, Aug 6, 2014, http://www.bloomberg.com/news/2014-08-06/china-said-to-exclude-apple-from-procurement-list.html; Charles Clover, *China bans federal officials from buying Apple products*, FINANCIAL TIMES, Aug. 6, 2014.

[7] China accounted for 20% of Apple's global revenues in the last calendar quarter of 2017 (which Apple labels its first fiscal quarter of 2018). https://www.apple.com/newsroom/pdfs/Q1_FY18_Data_Summary.pdf.

[8] https://www.ftc.gov/news-events/press-releases/2017/01/ftc-charges-d-link-put-consumers-privacy-risk-due-inadequate. The FTC alleged "hard-coded" login credentials integrated into D-Link camera software, software flaws that enable remote attackers to control customer's devices, and the failure to encrypt login credentials on D-Link's mobile app.

[9] Lily Hay Newman, The Army Grounds Its DJI Drones Over Security Concerns, Wired, Aug. 8, 2017, https://www.wired.com/story/army-dji-drone-ban/.

[10] Farrokh Farrokhnia & Cameron Richards, E-Commerce Products Under the World Trade Organization Agreements: Goods, Services, Both or Neither?, 50 J. World Trade 793, 800 (2016) ("the issue of classification is one of practical significance since it would determine thee nature of the trade regime for relevant products").

The analysis proceeds as follows. Part I motivates the inquiry by observing how IoT raises important concerns about privacy and international standards. Part II then turns to an examination of how international trade law will approach these new hybrid subjects of international trade. It begins by asking how international trade should classify IoT. It then assesses how to determine which WTO discipline to call to bear with respect to a particular dispute involving a smart object.

I. REGULATORY CHALLENGES OF THE INTERNET OF THINGS

The rapid deployment of the Internet of Things worldwide will lead man governments to scrutinize this international trade more closely. This Part discusses some of the legal issues raised by the global deployment of the Internet of Things. In its 2017 Information Economy Report, the United Nations agency UNCTAD noted the digital economy requires us to consider "data security risks, data localization pressures, as well as data collection and privacy concerns."[11] The Internet of Things raises questions regarding the abuse of private information, the deployment of insecure devices, and the need for standards and lack of interoperability.

A. PRIVACY AND SECURITY

IoT devices already outnumber the number of people in the world.[12] As UNCTAD observes," IoT devices "silently listen, watch and record location and activity in the household, the workplace, and/or in public places to assist individuals with their lives or help companies or governments improve their goods or services or tailor advertisements."[13] This creates both privacy and security risks, as the information might be abused or compromised. The race to deliver IoT devices cheaply may result in devices that are insufficiently secure. Indeed, hackers exploited security vulnerabilities to take control of IoT devices, allowing them to use these "zombie" devices to deliver a massive denial of service attack.[14] A report from the United States Federal Trade Commission notes that IoT presents a variety of security risks by: "(1)

[11] UNCTAD, 2017 Information Economy Report at __.

[12] http://www.zdnet.com/article/iot-devices-will-outnumber-the-worlds-population-this-year-for-the-first-time/.

[13] UNCTAD, Information Economy Report 2017 at 5.

[14] Lily Hay Newman, The Botnet That Broke the Internet Isn't Going Away, Wired, https://www.wired.com/2016/12/botnet-broke-internet-isnt-going-away/; Shackelford et al., 2017; "A new era of internet attacks powered by

everyday devices", *New York Times,* 23 October 2016.

enabling unauthorized access and misuse of personal information; (2) facilitating attacks on other systems; and (3) creating risks to personal safety."[15]

Even used as intended, the devices raise significant privacy concerns because of their immense data collection capacities and their ubiquitous deployment. Traditional methods used to notify individuals about data gathering are not readily available because these devices often lack screens to transmit such information.[16] Furthermore, the fact that multiple people might encounter an IoT device during its lifetime means that whoever installs and configures that device effectively makes privacy determinations for others.

UNCTAD's Global Cyberlaw Tracker reveals that 107 countries have established data protection/privacy legislation.[17] These laws will presumably apply to IoT manufacturers and service providers, including foreign manufacturers and service providers providing such services from abroad. Privacy requirements can incentivize security measures that reduce privacy risks.

Governments must protect the privacy and security of their citizens' information whether the information is held by a domestic or a foreign IoT provider. The United Nations General Assembly has recently reaffirmed the right to privacy as set out in article 12 of the Universal Declaration of Human Rights and article 17 of the International Covenant on Civil and Political Rights.[18]

B. STANDARDS AND INTEROPERABILITY

IoT today is characterized by the emergence of often proprietary, incompatible ecosystems rather than open, interoperable networks. This means that these "operational technology systems work largely in silos."[19] The fragmentation manifests itself in multiple ways, including different manufacturers, different operating systems, different versions of software, different types of connectors, and different

[15] Federal Trade Commission, supra note 5, at ii.

[16] Scott Peppet, 2014.

[17] http://unctad.org/en/Pages/DTL/STI_and_ICTs/ICT4D-Legislation/eCom-Data-Protection-Laws.aspx.

[18] UN General Assembly, The Right to Privacy in the Digital Age, G.A. Res. 68/167, U.N. Doc. A/RES/68/167 (Dec. 18, 2014), available at http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/68/167; for a discussion, see Anupam Chander & Molly Land, Introductory Note to United Nations General Assembly Resolution on the Right to Privacy in the Digital Age, International Legal Materials Vol. 53, No. 4 (2014), pp. 727-731.

[19] World Economic Forum, *supra* note 4.

communications protocols.[20] Maximizing the value of networked devices will involve increasing the compatibility of communications and other protocols at different layers of the network stack. The European Commission, for example, declares interoperability between devices and services a key to its Europe 2020 Strategy: "The EU must enhance the interoperability of devices, applications, data repositories, services and networks."[21] Interoperability can reduce the winner-take-all result of network industries; if other companies can participate in the network without needing the permission of a particular provider, it opens room for competition. At the same time, "imposing standards across devices could curb investment and innovation."[22]

Because smart objects must communicate their information to the outside world, they often depend on access to telecommunications networks. Local regulations might be written in ways to disadvantage foreign smart object suppliers' access to local networks. The GATS Annex on Telecommunications seeks to prevent such actions. Article 5(a) of the Annex mandates that foreign service suppliers must have "access to and use of public telecommunications transport networks and services on reasonable and nondiscriminatory terms and conditions" for the supply of a service that is listed in that country's schedule of liberalization commitments.

C. DATA LOCALIZATION

The Internet of Things would not be possible without global data flows. Communication is at the heart of IoT, with machines talking to people or to other machines.[23] IoT devices must communicate with their manufacturers or third parties selected by the manufacturer for data services and software updates. Smart objects depend on a remote infrastructure, receiving, storing, and processing information through that infrastructure. Most IoT manufacturers either build that data infrastructure locally near their home jurisdiction, or contract with cloud service providers like Alibaba, Amazon, Google, or Microsoft to provide scalable servers. As the smart objects are sold across borders by their manufacturers, the manufacturers

[20] Altimeter, Interoperability: The Challenge Facing the Internet of Things (2014), available at https://www.prophet.com/thinking/2014/02/interoperability-the-challenge-facing-the-internet-of-things/.

[21] Digital Single Market, Europe 2020 Strategy, available at https://ec.europa.eu/digital-single-market/en/europe-2020-strategy.

[22] Brandie Nonnecke, Mia Bruch, & Camille Crittenden, IoT & Sustainability: Practice, Policy and Promise (June 3, 2016).

[23] One study places information at the heart of IoT (noting that "information lies at the heart of IoT, feeding into a continuous cycle of sensing, decision making, and actions"), but it also observes that "It is imperative for things to have the capability of communication – exchanging data over a network between them and/or with the cloud backend services." ENISA, Baseline Security Recommendations for IoT in the context of Critical Information Infrastructures 18-19 (Nov. 2017).

and other related service providers rely on cross-border data flows to power these devices.

The European Commission has recognized the importance of facilitating "the flow and transfer of data," observing that IoT involves the generation of data, the transfer of data, storage of data, processing of data, and the provision of data services, all of which must flow across borders.[24] At the same time, many countries are increasingly demanding that data not be taken out of their country on privacy and security grounds.[25]

Data localization mandates take a variety of forms. Nigeria, for example, requires information and communications technology companies to host all subscriber and consumer data locally within the country.[26] Australia requires that personally identifiable health information must not leave the country without the consent of the individual to whom it pertains. British Columbia and Nova Scotia prevent personal information held by government agencies from leaving Canada without the consent of the data subject.[27] Consent requirements pose special difficulties for IoT as devices often interact with multiple persons, not just the individual installing the device. Because IoT relies on remote storage and processing of information, restrictions on cross-border data flows significantly interfere with the ability to create global IoT products and services.

Data localization requirements mean that IoT manufacturers must either establish or lease local data facilities in every country with such requirements, substantially raising the costs of supplying IoT across the world. Such requirements not only raise costs, they also slow down global sales and add additional security risks because of the need to secure additional computer servers. Of course, some IoT manufacturers might ignore data localization obligations altogether because such laws are difficult to enforce. As the World Bank has pointed out, "some countries are using these [data localization] barriers to protect local firms."[28] Rules that hinder data flows across borders are facially discriminatory against foreign providers of data services. Data localization requirements effectively disfavor foreign IoT manufacturers who are

[24] European Commission, Advancing the Internet of Things in Europe 13 (2016).

[25] For a roundup of data localization obligations, see Anupam Chander & Uyên P. Lê, Data Nationalism, 64 Emory L.J. 677 (2015).

[26] Nigerian Law Intellectual Property Watch, Guidelines for Nigerian Content Development in Information and Communications Technology (ICT) (section 12.1), available at https://nlipw.com/guidelines-nigerian-content-development-information-communications-technology-ict/.

[27] Usman Ahmed & Anupam Chander, Information Goes Global: Protecting Privacy, Security, and the New Economy in a World of Cross-Border Data Flows (2015).

[28] World Bank, Reaping Digital Dividends Leveraging the Internet for Development in Europe and Central Asia at 145.

less likely to have local data infrastructures. Thus, data localization requirements may violate commitments to liberalize trade in goods and services.[29]

II. APPLYING TRADE LAW TO THE INTERNET OF THINGS

How should trade law understand a smart object or its ongoing operations? Should it see smart goods and their operations as a good or a service, both or neither? If we answer "both good and service" as I will suggest here, then when do we apply GATT and when GATS and when both? The next two sections consider these questions in turn.

A. BOTH GOOD AND SERVICE

Smart objects have been with us since the dawn of computing. If we saw them as simply the evolutionary successor to computerized objects such as the Casio smartwatches of the early 1980s, we might conclude that we should treat them simply as goods, whatever their purported smarts. After all, the Casio Databank watch stored an address book and calendar, alongside calculator functions. It certainly held a computer chip.[30] But the smart objects of today are more Dick Tracy than Casio Databank. Today's smartwatches connect user information to the Internet, storing and accessing information held on Internet servers around the world. They can monitor our heartbeats, perhaps even predicting high blood pressure through machine learning-based artificial intelligence applied to the data gathered by the device.[31] Extensive ongoing data services also generally characterize today's Smart Objects as well: the continuous, real-time, evolving information flows emanating from and to the Internet of Things distinguish them from most earlier computerized objects. While computers have long been embedded in devices, from Casio smartwatches to Tickle-Me-Elmo dolls, the new devices also continuously communicate with the world, collecting and evaluating information.

The 1980s Casio smartwatch can be seen as providing a service—telling time, remembering your calendar, storing your contacts, or doing calculations. But this construction might transform all goods into service providers—a fan can be seen as a cooling service provider; a stool becomes a sitting service provider; a tractor, a plowing

[29] For an important overview of trade law's discipline of data regulation, see Mira Burri, The Governance of Data and Data Flows in Trade Agreements: The Pitfalls of Legal Adaptation, 51 UC Davis L. Rev. 65 (2017). For an examination of data localization measures from a trade law perspective, see http://e15initiative.org/wp-content/uploads/2015/09/E15-Policy-Brief-Crosby-Final.pdf.

[30] http://en.wikipedia.org/wiki/Calculator_watch.

[31] https://www.wired.com/story/ai-can-help-apple-watch-predict-high-blood-pressure-sleep-apnea/.

service provider; or a car, a transportation service provider. Transforming all goods into services eliminates the goods/services distinction without any useful effect.

How should we then identify whether a particular good entails a simultaneous service? To date, WTO discussions related to the Internet have largely focused either on e-commerce used to enable trade in goods or services, but have not considered the growing challenges of the Internet of Things.[32]

International trade law has proved reticent in seeking to define a service with precision. GATS merely offers a recursive definition: "For the purposes of this Agreement, trade in services is defined as the supply of a service…."[33] For its part, the Dispute Resolution Body has not sought to define services abstractly, but rather simply identified a particular measure as one affecting services when faced with real world challenges that required a classification. Given technological changes that are creating new kinds of services or enabling for the first-time international trade in existing kinds of services, such reluctance to preordain a strict definition and thereby leave this question to future developments seems prudent.

Even the expansion of the Information Technology Agreement at the Ministerial Conference in 2015 has not sought to clarify these issues. In 1996, many WTO nations agreed to phase out duties on the imports of a variety of information technology products.[34] In 2015, the now-larger membership of the Information Technology Agreement agreed to expand the list of duty-free information technology products,[35] but did not clarify the application of the WTO agreements to the Internet of Things.

The classic WTO case exploring the intersection between a good and a service is *Canada—Periodicals*.[36] There the United States challenged a special Canadian tax on periodicals that adversely affected United States periodicals such as *Sports Illustrated*. The United States argued that the Canadian tax violated Canada's national treatment obligation for U.S. products under GATT. Canada countered that the tax was directed

[32] WTO, "Work Programme on Electronic Commerce" adopted on 25 September 1998 (WT/L/274). In 2013, the General Council expanded its discussions as follows: "the Work Programme should continue to examine the trade related aspects of, inter alia, enhancing internet connectivity and access to information and telecommunications technologies and public internet sites, the growth of mobile telephony, electronically delivered software, cloud computing, the protection of confidential data, privacy and consumer protection." WTO, Ministerial Decision of 7 Dec. 2013, WT/MIN(13)/32, WT/L/907, 11 December 2013.

[33] GATS Art. I:1.

[34] Information Technology Agreement.

[35] https://www.wto.org/english/thewto_e/minist_e/mc10_e/briefing_notes_e/brief_ita_e.htm.

[36] Appellate Body Report, *Canada–Periodicals*, WT/DS31/AB/R (Jun. 30, 1997).

towards advertising in the magazines, and thus was a measure affecting a *service*, not a *good*. Since Canada had not promised national treatment for advertising services under GATT, the Canadian characterization of the measure as one affecting a service would have defeated the U.S. challenge to the discriminatory tax. The Appellate Body rejected this argument, observing, "The entry into force of the GATS … does not diminish the scope of the application of the GATT 1994."[37] The Appellate Body concurred with the panel's view that the "obligations under GATT 1994 and GATS can co-exist." [38] The Appellate Body found that the periodical in question implicated services—but that the final product was a good which comprised services: "[A] periodical is a good comprised of two components: editorial content and advertising content. Both components can be viewed as having services attributes, but they combine to form a physical product -- the periodical itself."[39]

Applying *Canada—Periodicals* to smart objects, should we not conclude that a smart object is a good, which is comprised in part of services? Are not smart objects best understood simply as goods, the successor to computerized objects such as the Casio smartwatches of the early 1980s?[40] After all, the Casio Databank watch stored an address book and calendar, alongside calculator functions. It certainly held a computer chip.

But today's smartwatches connect user information to the Internet, storing and accessing information held on Internet servers around the world. This is true generally of today's smart objects as well: the continuous, real-time, evolving information flows emanating from and to the Internet of Things and the robots of today distinguish them from most earlier computerized objects. While computers have long been embedded in devices, from Casio smartwatches to a Tickle-Me-Elmo, the new devices also continuously communicate with the world.

Even if they communicate with the world, does that necessarily involve a service? Perhaps we should consider the data flows as communications, not as services at all? While it is easy to see the "good" aspect of a smart object, it can be more difficult to recognize the services embedded within. Services now provided across borders include such abstract concept as thinking, analyzing, recommending, and remembering. In many cases, the data flows entailed by these products cannot be found in traditional tariff classification schemes.[41]

[37] *Id.* at 19.

[38] *Id.*

[39] *Id.* at 17.

[40] http://en.wikipedia.org/wiki/Calculator_watch.

[41] *Cf.* Fiona Smith & Lorna Woods, *A Distinction without a Difference: Exploring the Boundary between Goods and Services in the World Trade Organization and the European Union*, 12 COLUMBIA J. EURO.L. 463, 510 (2005/06) ("[N]ew products may not fit easily into the existing coding

In *China – Electronic Payment Systems*, the WTO panel embraced a broad view of data operations as services. Consider the wide array of functions performed electronically that the panel recognized as services:

> *The Panel recalls that the services at issue, as defined in the panel request, consist of a "system" that "typically includes" five elements, namely (i) the processing infrastructure, network, and rules and procedures that facilitate, manage, and enable transaction information and payment flows and which provide system integrity, stability and financial risk reduction; (ii) the process and coordination of approving or declining a transaction, with approval generally permitting a purchase to be finalized or cash to be disbursed or exchanged; (iii) the delivery of transaction information among participating entities; (iv) the calculation, determination, and reporting of the net financial position of relevant institutions for all transactions that have been authorized; and (v) the facilitation, management and/or other participation in the transfer of net payments owed among participating institutions.*[42]

The data storage and processing required for a smart object seem of a kind with the operations recognized as services in *China – Electronic Payment Systems*. Rather than supporting financial transactions, the data services from a smart object might support health monitoring and analysis, or usage rates and times, etc.

Some of the data flows from smart objects are easy to recognize as services. Take, for example, the home monitoring service offered by makers of modern surveillance cameras. The Nest home surveillance system offers a $199 camera, a major feature of which—permitting the user to rewind and see who visited the premises the previous day --only works with a $5 per month video recording service.[43] That service consists in cloud recording and replaying of the video.

In many cases, the economic value of the service will over the long term overwhelm the value of the good. Again, this is evident in smart objects such as a Nest, for which the monthly video recording service cost will far exceed the cost of the camera over the lifetime of the device.

But what of a Samsung home monitoring camera, which offers an option to send the video home surveillance recording to the user's Google drive account?[44] (This

systems with disagreement arising over the correct classification of the product. There is a risk of discrepancies arising in two contexts: either products can be classified differently within the HS or W/120/CPC code, or, more radically, products can be classified as goods in one scheme and services in another. This problem is acute for products traded online although more established products, such as those of the communications industry, have also given rise to problems.").

[42] *China – Electronic Payment Systems*, para. 7.41.

[43] https://store.nest.com/product/camera/NC1102ES (advertising $199 camera and a $5/month or $50/year service for "continuous recording, intelligent alerts.").

[44] https://www.samsungsmartcam.com/manual/android_en.pdf ("A 30-second video clip is uploaded automatically to the user's Google Drive account."). Samsung earlier offered

might well involve the flow of data from a house in California to a data server in South Korea and then back to Google's data servers on the West Coast.) And all of this for free. Perhaps the *sine qua non* of a service should be whether it is provided for a cost? Under such a rule, Wikipedia would not be a service under international trade, even though it largely replaced the expensive encyclopedias of earlier generations. For smart objects like the Samsung camera, it seems better to treat the service as bundled with the good itself at the point of sale. Indeed, one of the key selling points distinguishing the Samsung home surveillance camera is the fact that one does not have to pay ongoing fees for the monitoring service, by employing free services instead. Thus, rather than seeing the data services provided for the lifetime of the object as free, we might see them instead as prepaid. After all, it costs Samsung money to provide the data processing for such cameras.

Thus, it makes sense to see a Smart Object as both a good and an ongoing service, and any regulation thereof thus subject to both GATT and GATS disciplines. In *China – Audiovisual,* the Appellate Body affirmed that "a measure can regulate both goods and services and that, as a result, the same measure can be subject to obligations affecting trade in goods and obligations affecting trade in services."[45]

In sum, it seems likely that the Dispute Resolution Body would conclude that smart objects are goods with embedded services, subject to both GATT and GATS disciplines.

B. GATT OR GATS?

The fact that a smart object may be subject to GATT and GATS disciplines simultaneously does not answer the question as to which treaty to apply in any particular challenge to a specific measure.

Again, the case of *Canada—Periodicals* is instructive. There, the Appellate Body had to decide whether GATT or GATS should be applied, with Canada arguing for the application of GATS, and the U.S. arguing for the application of GATT. The critical question, as the Appellate Body saw it, to determine whether to apply GATT or only GATS to the dispute turned not simply on an examination of the good itself, but on *the measure at issue*. The Appellate Body wrote, "The measure at issue in this appeal, Part V.1 of the Excise Tax Act, is a measure which clearly applies to goods." The Appellate Body continued,

to upload video to the user's private YouTube channel, but discontinued that in 2014. https://www.samsungsmartcam.com/web/cmm/board/view.do?idx=151&currPage=1&lastPage=1.

[45] Appellate Body Report, *China – Audiovisual*, paragraph 194, WT/DS363/AB/R (Dec. 21, 2009).

> An examination of Part V.1 of the Excise Tax Act demonstrates that it is an excise tax which is applied on a good, a split-run edition of a periodical, on a "per issue" basis. By its very structure and design, it is a tax on a periodical. It is the publisher, or in the absence of a publisher resident in Canada, the distributor, the printer or the wholesaler, who is liable to pay the tax, not the advertiser.[46]

If the measure at issue had been directed at the advertiser in the periodical, then it might have been appropriate to characterize the measure as directed towards the regulation of the service. This is consistent with the opening mandate of GATS, set forth in Article 1:1: "This Agreement applies to measures by Members affecting trade in services."[47]

Thus, the answer to the question of GATT or GATS does not depend on the nature of the economic transaction central to the dispute, but rather the measure at issue and to what it is applied. This approach recognizes that services and goods are often conjoined in a particular economic activity. Determining whether to apply GATT or GATS turns on what the measure is applied to—the good or the service in the economic activity. The Appellate Body's recognition in *Canada—Periodicals* that goods can have services embedded in them seems especially apt with respect to the Internet of Things.

Confirmation of this approach can be found in another Appellate Body decision in the case of *China—Audiovisual*, a dispute which also involved the consideration of the intersection between goods and services. The Appellate Body repeated its observation in *Canada—Periodicals* that "particular measures 'could be found to fall within the scope of both the GATT 1994 and the GATS,' and that such measures include those 'that involve a service relating to a particular good or a service supplied in conjunction with a particular good.'"[48] It continued, "a measure can regulate both goods and services and that, as a result, the same measure can be subject to obligations affecting trade in goods and obligations affecting trade in services."[49]

China argued that its measures concerning films and audiovisual products did not regulate goods, but rather the content of films shown in China—and thus were not covered by its trading rights commitments, which were limited to goods. The Appellate Body, however, concluded the regulation limiting content necessarily limited who could import goods, and that therefore the measure implicated China's trading

[46] Id. at 18.

[47] GATS, Art. 1:1.

[48] Appellate Body Report, *China – Audiovisual*, paragraph 193, WT/DS363/AB/R (Dec. 21, 2009).

[49] Id. at paras. 196-198.

rights commitments.[50] In order to determine whether goods commitments were implicated, the Appellate Body asked whether the challenged measure affected a foreign supplier of goods. Finding that it did, the Appellate Body assessed whether the measure violated the country had violated its commitments with respect to those goods.

This approach explains why even a case involving bananas might bring to bear the GATS. In *EC — Bananas III*, Ecuador and a number of other countries brought a variety of claims against a European licensing regime for banana imports, distribution, and sale.[51] In addition to the principal GATT claims, Ecuador argued that the European measures violated that region's GATS commitments to permit Ecuadorian "wholesale trade service" providers. The European Communities insisted that their measures did not implicate GATS because the measures were focused on the licensing of goods, and thus should be examined under GATT exclusively. The Appellate Body sided with Ecuador, concluding that European measures that prevented the Ecuadorian companies from buying or selling certain bananas interfered with Europe's GATS commitments.[52] The Appellate Body concluded, "It is difficult to conceive how a wholesaler could engage in the 'principal service' of 'reselling' a product if it could not also purchase or, in some cases, import the product."[53]

This approach offers a sensible means to determine whether to apply GATT or GATS (or both) in any particular dispute. If the measure is directed towards the regulation of the service, then GATS disciplines should apply; if directed towards the regulation of the good qua good, then GATT disciplines should apply. If directed towards both, both would apply. Rather than beginning with the particular item of international trade to determine whether it consists in a good, service, or both, we examine the challenged measure to see whether it is targeting a good, a service, or both.

What if a measure regulates both a service and a good simultaneously? In *EC – Bananas III*, the Appellate Body offered a way to determine how to apply GATT and GATS when a measure implicated both. In that case, the Appellate Body had in mind "measures that involve a service relating to a particular good or a service supplied in conjunction with a particular good."[54] Such a measure, it noted, "could be scrutinized

[50] Id. at para. 188 ("where the content of a film is carried by physical delivery materials, Article 30 of the *Film Regulation* will *inevitably* regulate who may import goods for the plain reason that the content of a film is expressed through, and embedded in, a physical good.").

[51] EC – Bananas III.

[52] Id. at para. 244.

[53] Id. at para. 226.

[54] Id. at para. 221.

under both the GATT 1994 and the GATS."[55] However, while the same measure could be scrutinized under both agreements, the specific aspects of that measure examined under each agreement could be different: "Under the GATT 1994, the focus is on how the measure affects the goods involved. Under the GATS, the focus is on how the measure affects the supply of the service or the service."[56]

Some have argued that the WTO framework needs to be revised in light of contemporary technologies. Farrokh Farrokhnia & Cameron Richards argue that e-commerce products could conceivably constitute both a good and a service. They seem to have in mind products that are intangible, such as movies or music in digital form. They maintain that intangibility alone does not necessitate any conclusion as to whether a particular product is either a good or a service because neither GATT nor GATS "says anything about tangibility or intangibility."[57] They note the possible "need to create a new category for e-commerce products in the WTO framework."[58]

Lucian Cernat and Zornitsa Kutlina-Dimitrova propose such a new category—but for a different type of economic activity. Arguing that the existing GATS framework is inadequate to the increasing role of services in manufacturing, they propose a new mode 5 for services that are incorporated into products.[59] They have in mind goods involving design or similar services, or goods embedded with software. The introduction of mode 5 would mean that countries would need to identify whether they intend to make commitments to national treatment and/or market access for each good/service category combination—a rather demanding requirement given the long-running impasse in multilateral liberalization.

CONCLUSION

Even while offering substantial improvements in our lives, the Internet of Things will require significant regulatory oversight. In particular, ubiquitous smart objects will raise questions of privacy, security, standards, and interoperability. The coming of this smart world will also put pressure on trade law, as dispute settlement

[55] Id.

[56] Id. at para. 221.

[57] Farrokhnia & Richards, supra note __, at 810.

[58] Id. at 815.

[59] Lucian Cernat and Zornitsa Kutlina-Dimitrova, Thinking in a Box: A 'Mode 5' Approach to Service Trade, 48 Journal of World Trade 1109 (2014) ("The [existing] GATS four modes of services supply … do not account for the fact that a substantial and increasing share of services is being embodied in products and traded around the globe.").

mechanisms are invoked to assess whether a particular government measure is legitimate regulation or simply disguised protectionism.

The coming of the Internet of Things complicates the elegant distinctions at the heart of international trade law. At the same time, it reveals that trade law always recognized the complexity of a world where goods embedded services. And it further reveals that the international trade regime may yet prove more adaptable than might have been expected.

Trade Commitments and Data Flows: The National Security Wildcard Reconciling Passenger Name Record Transfer Agreements and European Union GATS Obligations

Norman Zhang

Abstract

This paper poses a hypothetical WTO challenge to the Passenger Name Records (PNR) Transfer Agreements the European Union has signed with the United States (as well as Australia and Canada). The focus will be on a possible citation of GATS Art. XIV National Security Exception by the EU, and the viability of such a defense. Because of the absence of caselaw, this paper will also attempt to synthesize an acceptable standard for assessing GATS National Security Exception citations

Working Paper 20180509

Learning about digital trade: Privacy and e-commerce in CETA and TPP

Robert Wolfe

An earlier version of this paper was presented at the Columbia Law School Trade Seminar Series on November 6, 2017. I am grateful for insightful comments from Susan Ariel Aaronson, Henry Gao, Bernard Hoekman, Petros C. Mavroidis and Roy Santana, for the able research assistance of Grace Tahan, Sifat Syeda, and Maria Bridgemohan; and for confidential interviews in Geneva and Ottawa.

Abstract

It is a truth universally acknowledged that every ambitious 21st century trade agreement is in want of a chapter on electronic commerce. One of the most politically sensitive and technically challenging issues is personal privacy, including cross-border transfer of information by electronic means, use and location of computing facilities, and personal information protection. States are learning to solve the problem of state responsibility for something that does not respect their borders while still allowing 21st century commerce to develop. A comparison of the Canada-European Union Comprehensive Economic and Trade Agreement (CETA) and the Trans-Pacific Partnership (TPP) allows us to see the evolution of the issues thought necessary for an e-commerce chapter, since both include Canada, and to see the differing priorities of the U.S. and the EU, since they are each signatory to one of the agreements, but not of the other. I conclude by seeking generalizations about why we see a mix of aspirational and obligatory provisions in free trade agreements. I suggest that the reasons are that governments are learning how to work with each other in a new domain, and learning about the trade implications of these issues.

COLUMBIA|SIPA
School of International and Public Affairs

Trade Regulation, and Digital Trade

May, 2017

By Petros C. Mavroidis[i]

1. The WTO: Neither Transactional, Nor Policy-Oriented

In 1998, the WTO (World Trade Organization) established a Working Group on Electronic Commerce (e-commerce).[ii] Almost twenty years later, the group has nothing to show in terms of achievements, other than a few papers discussing the general, potential applicability of multilateral rules on some forms of digital trade. True, even the minutes reflecting the outcome of WTO Ministerial Conferences include a few lines on “e-commerce”, but this is where the buck stops.[iii]

The WTO attitude is neither transactional, nor policy-oriented, as we explain in more detail later. It is haphazard. One cannot understand when going through all this mass of information regarding e-commerce, that the WTO has made publicly available, what the WTO-think on digital trade is. In the meantime, digital trade is progressing fast. According to data provided by the McKinsey Global Institute in 2016, the growth is explosive: international data flows are forty five times higher in 2014 than they were in 2005.[iv]

Under the circumstances, one might wonder whether international rules are necessary at all. Digital trade grows fast anyway. And yet, a number of issues arise that impede further progress, and that require solutions preferably at the multilateral level: data localization, geo-blocking are the latest in a series of examples on this front. The WTO Work Programme has not managed to address similar issues head on. It has not managed to integrate them in a wider thinking about digital trade either.

Some free trade areas (FTAs) have managed to fare better on this front. There are, of course, a number of reasons why this has been the case ranging from homogeneity of players involved (who share similar concerns) to negotiating costs. It is submitted that one reason why FTAs succeed where WTO has failed lies in that it is easier to bring together the trade and regulatory communities in a forum consisting of like-minded players. Digital trade is not about trade exclusively. There is an important regulatory dimension that covers issues such as privacy, security etc. This issue must be considered as well. The trading community will discuss how it applies to infra-firm flows for which there is no associated payment flow. We will end up thus, with a PPM (process and production method) analogue set of issues. Production function matters in this discussion (e.g., is data secure? How ensure security? etc.). The regulatory community will be discussing this latter set of issues.

In Section 2, I briefly discuss where WTO stands now on digital trade. In Section 3, equally briefly I discuss some illustrative FTA-examples, and finally, in Section 4 I provide scaffolding for a more structured discussion on digital trade in the WTO.

2. Multilateral Regulation of Digital Trade

I divide this discussion in two parts: what is the coverage of digital trade at the WTO-level as rules now stand, followed by a brief discussion of he Work Programme. I kick off this Section with semantics.

2.1 What is Digital Trade

Official WTO documents use the term "e-commerce" (instead of digital trade), which is routinely defined as

Production, distribution, marketing, sale or delivery of goods and services by electronic means.

Thus expressed, the term covers not only end-to-end delivery of services, like internet and other telecoms, but also other services that can be transmitted in digitized form. The legal regime applicable to these transactions is that provided in the various national schedules of commitments under GATS (General Agreement on Trade in Services). Recall nonetheless, that in US-Gambling, the Appellate Body (AB) endorsed "technological neutrality", that is, the means of supply of a service does not matter. Digitally transmitted services are covered by commitments entered even when digital supply was not an option at the moment when the commitment had been entered.

And what about goods sold on the internet? Well, it all depends on their characterization as goods or services. A book sold say on Amazon will be subjected to the tariff concessions of the importing state. Panels have yet to decide whether a song sold on Amazon, if downloaded and saved, should be characterized as good or service.

Finally note that, n literature, the term "digital Trade" seems to be associated with a wider coverage than "e-commerce" as explained above. Branstetter (2016) for example, includes the following definition.

> *… the full range of electronic commerce issues, from online commercial transactions to the ancillary aspects of protection of intellectual property rights, privacy, and the protection of national interests.*

This wider understanding of the term is more in line with expressed business interests.

2.2 As Things Stand

WTO does not regulate head on e-commerce (or digital trade) but electronically transmitted services are covered by the GATS to the extent that commitments to liberalize the pertinent service sector have been made.[v] Indeed, WTO adjudicating bodies have resolved disputes dealing with electronically transmitted services.

In US-Gambling, the AB held that the US was violating its commitments regarding the supply of internet gambling. In China-Publications and Audiovisual Products, it was upheld that the electronic distribution of music was covered. In China-Electronic Payment Services, the AB held that the Chinese electronic payments regime was in violation of nondiscrimination. Finally, in Mexico-Telecoms, the Panel held that Mexico was violating its commitments on telecoms by imposing supra-competitive termination rates.

The TRIPs (Trade-related Intellectual Property Rights) Agreement as well, is relevant to this discussion. IP rights have typically a territorial dimension, and it is precisely this characteristic of IP rights that might obstruct supply of digitalized services. Since TRIPs embeds a minimum standard of protection of IP rights, WTO members remain free to enact higher standards of protection to the extent that they observe nondiscrimination. Nothing of course, stops WTO members from signing agreements to by-pass national idiosyncratic elements.

2.3 Work Programme

The Work Programme aims to bring e-commerce under the multilateral disciplines. At the moment of writing, it is clear that we are far away from even a modest agreement.

Since the end of the Uruguay round agreement, the ITA (Information Technology Agreement, I and II) have been concluded. This agreement has liberalized trade by eliminating duties in products such as computers, semiconductors, or telecommunications equipment. Note that the number of the initial participants (29) grew significantly and reached 81, accounting for about 97% of world trade in IT goods.

3. FTAs and Digital Trade

Digital trade occupies space in the majority of free trade areas (FTAs) signed in the post-Uruguay round era.

3.1 Here, There and Everywhere

Take the European Union (EU) FTAs for example. Its agreement with Canada (CETA), Korea (KOREU), but also its agreements with more heterogeneous partners (like EU-Vietnam) all contain chapters dealing head on with digital trade (e-commerce).

The EU is not alone in this. US follows a similar path. The now (almost) defunct TPP, for example, contains provisions aiming to facilitate digital trade. There are some obvious starting points, like the provision to abolish duties on digital goods. There are also some more hotly debated issues that found their way into the text. The TPP, for example, takes a strong stance against data localization (not allowed to require the establishment of local computing facilities as a condition of doing business) .

TiSA (Trade in Services Agreement), the most ambitious plurilateral agreement[vi] negotiated between a few WTO members outside the confines of the WTO, when finalized, will include an Annex on E-Commerce, which would cover open networks, unsolicited commercial communication, interactive computing, and wider international cooperation in this area.

3.2 Advantage FTAs

FTAs go thus consistently further than the multilateral regime does when it comes to addressing digital trade.[vii] Issues like data localization for example, which have not entered the WTO jargon, are commonpce in the regulation of digital trade under the aegis of FTAs.

Why are trading partners prepared to do things bilaterally (or plurilaterally) and not multilaterally? After all, standard theory would suggest that deals should be easier when there is more to exchange. Regulation nevertheless, unlike tariffs cannot be dwindled down. To the extent that it exists for good reasons, it is nonnegotiable. The key is thus, to bring around the table regulators and the trading community. To sensitize the former to the trade impact of their measures, and the latter to the well-founded of the intervention.

This is what a close-knit group of like-minded players can do. Examples abound: from the US-Regulatory Cooperation Council to the instruments for regulatory cooperation in CETA. [viii]

4. A Role for the WTO

WTO should change course. Mindful of its limits, it should approach this discussion in functional manner, working on its strengths rather than embarking on a Work Programme with no compass where to go.

4.1 Advantage FTAs

WTO should attempt to address three questions:

- What is being delivered?
- Who delivers electronically?

These two questions will help identify the relevance of the WTO on digital trade, both with respect to the GATT and the GATS.

4.2 Next Steps

The next question for WTO should be what can be done to further liberalize digital trade. In that, WTO should function originally as complement to FTAs, and substitute for their efforts when gains can be multilateralized.

4.2.1 Building Bridges to the Hothouses of Regulation

Cutting edge issues are easier discussed across like-minded players. Think of the discussion on consumer privacy encryption, which has been taking place in TPP, for example, but is not in the radar screen of the WTO Work Programme.

Think also of the data localization issue for example. TiSA negotiations almost collapsed because of this issue. The EU, because of legal constraints, could not subscribe to the recipe advanced.[ix] This issue is being discussed in various bilateral fora, and has yet to find its way into the WTO Work Programme.

And then there are issues, which have not been resolved even in more intense integration processes. Geo-blocking has been plaguing the EU quest for a unified digital market in the European continent. Recently, the Commission has proposed a regulation that will constitute the first step only towards eliminating obstacles to market integration.[x]

The WTO has a lot to learn from these discussions. How do that?

4.2.2 Complements and Substitutes

WTO could complement these efforts by designing an osmosis mechanism. Issues that for example, have found similar or identical solutions in various FTAs could be debated as potential multilateral regulation. In doing that, WTO could become the multilateral substitute for regulation at the FTA-level.

In the meantime, it can provide an information-exchange regime, where good ideas and regulatory solutions agreed at the FTA-level could find a forum to be discussed by potentially interested players. Those keen could mimic the best regulatory examples. Others would have additional food for thinking their next regulatory interventions.

Reference

1. Bollyky, Tom, and Petros C. Mavroidis. 2017. Trade, Social Preferences and Regulatory Cooperation, the New WTO Think, Journal of International Economic Law, 20: 1-30
2. Branstetter, Lee. 2016. TPP and Digital Trade, pp. 72-81 in Jeffrey J. Schott and Cathleen Cimino-Isaacs (eds.), Assessing the Trans-Pacific Partnership, vol. 2, Innovations in Trading Rules, Peterson Institute of International Economics: Washington DC.
3. Hoekman, Bernard M., and Petros C. Mavroidis. 2015. WTO 'à la carte' or WTO 'menu du jour'? Assessing the Case for Plurilateral Agreements, European Journal of International Law, 26: 319-343.
4. Mishra, Neha. 2017. The Role of the Trans-Pacific Partnership Agreement in the Internet Ecosystem: Uneasy Liaison or Synergistic Alliance, Journal of International Economic Law, forthcoming.

[i]Edwin B. Parker Professor of Law at CLS, New York City, and Professor of Law at the University of Neuchâtel (Switzerland). For helpful discussions, I am indebted to Bernard M. Hoekman, and Robert Wolfe.

[ii] WTO Doc. WT/MIN(98)/DEC/2 of 25 May 1998.

[iii] One can find all these documents (both members' proposals, as well as WTO Secretariat background papers) in the WTO webpage https://www.wto.org/english/tratop_e/ecom_e/ecom_e.htm

[iv] See also Information Economy Report (2015), Unblocking the Potential of E-Commerce for Developing Countries, UNCTAD: Geneva, Switzerland.

[v] In light of our discussion above, there is no need to elaborate any further on the regime regarding trade in goods.

[vi] Hoekman and Mavroidis (2015) discuss the workings of plurilateral agreements.

[vii] Mishra (2017) for example discusses TPP

[viii] Bollyky and Mavroidis (2017) discuss this issue in more detail.

[ix] In C-362/14 (Max Schrems v. Facebook), the Court held that a Commission decision to allow for free flow of data cannot undermine the powers of national supervisory authorities to review whether transfer of data complies with the requirements of the EU directive regulating this issue (and essentially discuss the consistency of transfer with protection of fundamental rights). Eventually, in February 2016 the EU and US reached an agreement, the implementation status of which is still uncertain. Branstetter (2016) refers to CGE studies quantifying the negative impact on investment resulting from data localization requirements.

[x] COM (2016) 289.

E-commerce in South Korean FTAs:

Policy Priorities and Provisional Inconsistencies

S. On International Trade Regulation Issues: Digital Trade

Evan Y. Kim

Abstract

The e-commerce chapters in South Korea's FTAs cover a wide range of issues, ranging from non-discrimination to electronic signatures. Across the agreements, the country's provisions on consumer protection, paperless trading, and data protection are uniquely consistent, while those on other issues are not. With the aid of a framework (Framer v. Follower) that captures the dynamics of bilateral negotiations, this paper argues that in Korea's case, the more consistent the particular set of provisions is portfolio-wide, the more likely it was for Korea to have prioritized the relevant issue and actively pushed its preferred terms in the FTAs.

Part 1 provides an overview of Korea's e-commerce chapters. Part 2 explains the purpose of the paper and Framer-Follower framework that drives the main analysis. Part 3 analyzes each issue included in Korea's e-commerce chapters using the framework. Part 4 lays out the findings, their implications, and possible impact on Korea's national interests.

Working Paper
For circulation at the 2017 Global Digital Futures Forum
May 5, 2017

COLUMBIA|SIPA
School of International and Public Affairs

Framing Conversation: What Would Internet Fragmentation Mean for the Digital Economy?

By William J. Drake

1. Introduction

The theme of this year's Forum is very timely, as the question of Internet fragmentation has been the focus of a good deal of discussion of late in both generalist and specialist policy circles. But before we can explore the potential impact of Internet fragmentation on the digital economy, global governance, and global trade, it would be useful to step back and consider what we mean by the term in the first place. Some references in popular media seem to suggest that the term connotes a singular phenomenon on which there is broad agreement so we can simply invoke it and move on from there.

In fact, Internet fragmentation remains a contested concept. A cursory review of its usage in various publications and public pronouncements suggests that people often speak of it when discussing a variety of problems and tensions that arise on the Internet that do not all originate from the same source. For example, some in the business community have used the term as a generalized reference to variations in national policies that add to the cost of doing business globally. While some such policies may indeed be related to fragmentation, many other simply reflect differences in national legal systems, policy traditions, and so on that may antedate and arguably do not fragment the Internet. Similarly, some people have described the increasing linguistic diversity of cyberspace as an example of fragmentation, when of course this is simply a matter of a diverse humanity getting on line.

Another tendency among at least some observers is to suggest that the Internet is in imminent danger of falling apart. Because there is so much variation in national policies and practices, it is said, the Internet is likely to "break up" into a series of disconnected islands. This seems to be an overly dramatized misreading of some troubling trends. In fact, no cataclysm is around the corner; the underlying infrastructure remains stable and secure in its foundations, and it is incorporating new capabilities that open up new horizons, from the Internet of Things and services to the spread of block chain technology and beyond. But there are fragmentary pressures accumulating which, if left unattended, could reduce to varying degrees the Internet's enormous vitality and contributions to the world.

Conversely, while the examples just mentioned concern overly broad applications of the term, other observers tack in the opposite direction and say that "fragmentation" can only be properly used in reference to the Internet's underlying infrastructure rather than the creation of significant closed digital spaces. In one variant of this thinking, fragmentation would only happen if there was a massive defection from the unified Internet to entirely separate and non-interoperable systems running off different zone files. Since such a defection does not appear to be likely in the near future, voilà, there is no fragmentation, and people who argue to the contrary are needlessly hyperventilating, perhaps in hopes of looking prescient.

With these conditions in mind, in this memo I will briefly address three matters in the hope of helping to frame the conversation. First, I will advance working definitions of Internet fragmentation drawing on a white paper I wrote with colleagues for release at the World Economic Forum's (WEF) Annual Meeting in Davos in January 2016. [1] Second, I will highlight the variability and fluidity of Internet fragmentation in order to underscore that we are not talking about a simple binary condition that flicks on or off like a light switch. Third, I will conclude by raising a few concerns about the potential impacts of fragmentation on the evolving global digital economy.

2. Defining Internet Fragmentation

A useful starting point is to consider what we mean by an unfragmented Internet. What is the baseline from which fragmentation departs and against which it can be assessed? From a technical standpoint, the original shared vision guiding the Internet's development during the research and education era was that

[1] William J. Drake, Vinton G. Cerf, and Wolfgang Kleinwächter, *Internet Fragmentation: An Overview* (Geneva: The World Economic Forum, January 2016). A few bits of this memo derive from that earlier paper.

every willing endpoint on the Internet should be able to exchange data packets with any other endpoint that was willing to receive them. Universal "connectivity among the willing" was the guiding objective, and it could be achieved if autonomously controlled and even separately designed networks were internetworked and made interoperable via a shared protocol stack, TCP/IP, and related standards and protocols. Such interoperability needed to be to be seamlessly coherent on an end-to-end basis and consistent, so that users' actions would yield the same responses irrespective of location or the service providers involved.

These core features of universal, consistent interoperability and communicability between consenting end points were fundamental from a design standpoint. Every end point that wanted to send and receive bits with any other should be able to do so, so that the network of networks functioned as a free and open system. Actions or conditions that impaired this seamless functioning and blocked users from reaching each other could be said to constitute fragmentation.

Imagine, by way of analogy, an international telephone network on which people in country A could communicate with people in countries B, D, and F but not with people in countries C, E and G, while people in country B could communicate with people in countries A, C and E but not D, F and G, and so on across 196 countries. If humanity's ability to reach the full range of willing correspondents were this barrier-laden and segmented into go and no-go zones, would we characterize the global telephone network as open and unfragmented? Probably not. But on the Internet this sort of highly variable geometry of communicability is fairly standard and taken for granted, especially if one considers the infinite substantive variety of the bits that could be shared if allowed. We know that over 700 million users in China cannot access major platforms that are used by billions of people elsewhere; that billions of downloaders encounter messages like "the content you requested cannot be displayed;" that the transfer of certain classes of data out of certain countries is blocked or requires government permission; and so on, endlessly.

My contention, which like others is certainly contestable, is that the pervasive limitations on users' abilities to freely access, create, and dissemination information indicates an endemic condition of Internet fragmentation. The Internet is not a wide-open medium in which "anything goes," popular characterizations notwithstanding. It is certainly far more open than any global medium we have ever had before, and the limitations on its openness are frequently the focus of efforts to bypass or reverse by various actors, but they are there. And, as Eli Noam has argued in a provocative essay, they were inevitable.[2] There was simply no chance that the conditions that obtained in the early years when the Internet was a vehicle for the non-commercial sharing of research and educational information among computer scientists in various organizational settings could survive the transition to the Internet becoming a global mass medium used for an endless variety of social, commercial and political information sharing and resource discovery. Inevitably, governments were going to work to embed the Internet in frameworks of public authority that involved a wide variety of prescriptions and proscriptions, and companies were going to work to monetize peoples' access to and use of different kinds of contents by erecting a wide variety of enclosures and requirements. At the same time, with millions of technical people around the world working to deploy new capabilities, increase security and various other objectives, conditions could develop that, often unintentionally, had the effect of reducing or at least complicating the seamless functioning and interoperability of the infrastructure.

Hence, from this standpoint, it makes little sense to pose questions like "will the Internet fragment?" The Internet has long been fragmented to varying degrees in varying ways. A better question might be, will "Internet fragmentation increase in a manner that becomes much more problematic for a much wider range of uses and users?" Such a formulation turns our attention to the direction of change, rather than whether change might commence.

While Internet fragmentation has a common root---limitations on the ability of every willing endpoint to exchange data packets with any other willing endpoint---it is not a singular phenomenon. Fragmentation varies in its sources and manifestations in ways that are worth assessing separately on their own terms. Hence, in the above-mentioned paper for the WEF, my co-authors and I advanced three different "working definitions," so-called because the paper was an initial exploration and mapping and we were cognizant that more precise formulations might be desirable after our colleagues in the field

[2] Eli M. Noam, "Towards a Federated Internet", *InterMEDIA* (41, 4, 2013), pp. 10 –13.

kicked us around a bit on points that needed rethinking. We began from the proposition that a single "narrow definition" focused only on conditions in the underlying infrastructure would not capture how people use and experience the technology in order to construct digital social formations and engage in information, communication and commercial transactions, or by extension the sorts of political and economic forces that may impede their abilities to do so. We therefore amended the standard four-layer characterization of the Internet based on the TCP/IP Protocol Stack by adding a fifth

Content and Transactions layer to capture the substantive information exchanged and the interactions and behaviors involved.

Figure 1: Internet Layers

5. Content and Transactions Layer
4. Application Layer
3. Transport Layer
2. Network/IP Layer
1. Physical/Link Layer

Beginning from this amended baseline, we advanced the following working definitions of fragmentation:

- *Technical fragmentation:* conditions in the underlying infrastructure that impede the ability of systems to fully interoperate and exchange data packets and of the Internet to function consistently at all end points. These generally pertain to layers 1-4 of the model above.
- *Governmental fragmentation:* Government policies and actions that constrain or prevent certain uses of the Internet to create, distribute, or access information resources. These generally are targeted at the 5th layer in our model, but they may involve actions taken at the lower technical layers as well.

- *Commercial fragmentation:* Business practices that constrain or prevent certain uses of the Internet to create, distribute or access information resources. These generally are targeted at the 5th layer in our model, but they may involve actions taken at the lower technical layers as well. [3]

As is evident, one of our concerns was to distinguish sources and locations of fragmentation based in part on the question of intentionality. Technical fragmentation of the underlying physical and logical infrastructure is a complex evolutionary process that has unfolded slowly but is gathering pockets of steam in the contemporary era. Some of it has been intentional and motivated by operational and other concerns, but more often it has been the unintended by-product of actions taken with other objective in mind. In contrast, governmental and commercial fragmentation usually have been due to the intentional efforts of these third parties to establish limitations on users' abilities to create, distribute or access information. As a general matter, one could argue that such limitations are much more problematic and difficult to remediate than technical problems, for which engineers often can devise "fixes." In contrast, governmental and commercial fragmentation can be difficult to engineer "work arounds" for with lasting effects, e.g. people confronting censorship may rely on virtual private networks to mask their locations, but then governments figure out ways to block and monitor these and another technique must be found, at least until that too is found out.

[3] Drake, Cerf and Kleinwächter, p. 14.

3. Variability and Fluidity

Just as fragmentation is not singular in its form or the domain of its effects, the extent of fragmentation within and across the three categories also is highly variable. One could imagine a number of dimensions on which such variation could be found, but here are just three that merit consideration.

Occurrence: The first and most fundamental consideration is whether a given form of fragmentation exists. This is not an entirely straightforward question; as is noted above, fragmentation as a systemic property is not a simple binary condition that is either present or not present, and a specific instance of fragmentation in some domain may involve gradations with different values along a continuum. In some cases those values can be precisely quantified (e.g. the number of websites or other information resources to which access is fully blocked), but in others the best we can do is to devise ordinal measures. Similarly, there can be variations in duration. Fragmentation may be a short-term phenomenon that is rectified fairly quickly, as with recovery from some an attack that blocks access to resources, or it can be sustained as a long-term condition. In time sensitive situations, even short-term fragmentation can be very damaging to users or transactions. In general though, we should be most concerned with sustained fragmentation that has ongoing consequences.

A final issue here is that fragmentation does not need to be currently present to be of concern. That is, in many of the instances that people cite when worrying about the matter, what is at stake is the emergence of tendencies and pressures that could give rise to something significant in the future. As in any policy arena, we need not wait for a problem to become full blown and wreaking havoc for awareness and action to be well advised.

Intentionality: Fragmentation, particularly in the technical arena, may be the unintended by-product of decisions and actions guided by unrelated objectives. People who deploy or fail to deploy a particular technology in addressing a localized operational challenge may not be setting out to fragment the Internet. Nevertheless, their actions, especially if replicated by others, could come to have cumulative effects. Divergences between individually rational choices and systemically suboptimal consequences are a standard feature of collective actions problems generally and the same logic can apply to the openness or fragmentation of the Internet.

Alternatively, fragmentation may be intentional. The character of these intentions obviously matters quite a bit. On the one hand, organizations, communities and individuals may seek to separate themselves somewhat from the open public Internet for entirely defensible reasons. Installing a firewall to limit access and communication to only authorized and consenting parties and to protect resources from unwanted interference is a benign act of self-separation. In our WEF discussions last year, some participants argued that self-separation, such as the construction of firewalls or the use of encryption on a network, could be thought of as "positive fragmentation." I tend to think of this as being a different sort of activity that may involve some protective segmentation but is not preventing willing end points from communicating, since one end point is choosing to mediate its boundaries.

Of more concern, and more properly a matter of fragmentation in my view, is when actors such as governments seek to shape, constrain or fully block the activity of others who have not consented to this. Imposing limitations on others is a malign act of forced separation. Both unintentional and intentional fragmentation can be problematic, but the best approach to remediation may vary accordingly.

Impact: Fragmentation may be deep, structural and configurative of large swaths of activity or even the Internet as a whole. Consider, for example, the implications if significant categories of data flows were to be widely blocked around the world, or if an alternative root system with its own name space were to be established with the backing of powerful governments or organizations. The scope of the processes, transactions and actors impacted by such breakage would be substantial. But fragmentation also can be more shallow, malleable and applicable to a narrowly bounded set of processes, transactions and actors. The impact could be significant for some people but go unnoticed by others.

As with the other dimensions just mentioned, it can be difficult to measure the intensity of fragmentation and say with certainty exactly where on the continuum a given instance lays. Even so, in considering examples, we should be mindful that fragmentations are not all created equal in terms of magnitude and import. Indeed, a number of the examples one could mention are relatively low-impact or low-intensity matters – bothersome and concerning enough to engineers and operators that attention to them is merited, but not so significant that they endanger the fundamental integrity, openness and utility of the Internet. In contrast, some other action are higher-impact and arguably in need of concerted responses.

Given the above, from a systemic standpoint fragmentation is something of a shape shifter. It is always with us, particularly at the fifth level of content and transactions, but its specific manifestations are highly fluid and variable in scope, depth and duration. What should be of most concern are intentional forms that are deep, structural and configurative of large swaths of activity or even the Internet as a whole.

4. Implications for the Global Digital Economy

Some forms of fragmentation of this character are of relevance to the opening session on the digital economy. For example, with regard to technical fragmentation, if governments engage in widespread blocking of new generic top-level domains, opportunities for additional economic growth and social empowerment would be foreclosed. A massive defection by a leading country or countries to another root system, while presently unlikely, undoubtedly would have a very pronounced negative impact on the global digital economy. In general though, technical fragmentation at present does not seem likely to take on the sort of character that would in any dramatic way spoil the party.

Commercial fragmentation probably raises greater risks. There is growing concern today as to whether divergent corporate preferences may result in inadequate technical standardization of the emerging Internet of things. The adoption and locking in of proprietary standards in key arenas like this could produce fragmenting effects, with important products and processes not working well across corporate boundaries and national borders. The current push in the United States to abandon network neutrality as an organizing principle, driven in particular by traditional network operators and government ideologues, could result in widespread discrimination against applications and entrepreneurs and produce a fragmentary, multi-speed environment. Overly expansive and rigid intellectual property rules could curtail entrepreneurial dynamism as well as free expression and human empowerment. And as we move ever further into a platform-dominated online economy that absorbs an increasing share of advertising dollars and economic activity, the ways in which terms of service are constructed, the possibilities for anti-competitive behavior, and the prevalence of "walled garden" strategies may alter the character of the digital economy in ways that attenuate existing inequalities. Arguably, this may be particularly a concern with respect to the participation of developing countries in the digital trade arena.

Finally, and most importantly, governmental fragmentation of a structural nature seems to be a particularly pressing concern. The widespread "securitization" of Internet policies and the growth of so-called "cyber-sovereignty" strategies is already producing trends toward more widespread censorship and digital protectionism. These measures can be very difficult to roll back, and can impose significant costs on global companies and national economies and citizens alike. The potential scope of the challenge is underscored by the current trend toward forced data localization policies and the erection of barriers to cross-border data flows, which are the subject of a follow-up study to the above-mentioned fragmentation paper that will be released later this year. [4] In the opening session we may wish to delve into these and related questions.

[4] William J. Drake, *Data Localization and Barriers to Cross-Border Data Flows: Toward a Multistakeholder Approach,* (Geneva: The World Economic Forum, September 2017).

Working Paper
For circulation at the 2017 Global Digital Futures Forum
May 5, 2017

COLUMBIA|SIPA
School of International and Public Affairs

The Fragmentation Mismatch:

Deficiency of Dealing with Fragmentation through Trade Policy

By Hosuk Lee-Makiyama

1. The context to fragmentation

As we are two decades into the digitalisation, data is an established concept in trade policy. Yet fragmentation of the internet is still a matter of great urgency: In pursuit of "re-territorialisation" of digital economic space, 86 data localisation measures are applied in at least 36 jurisdictions (a number that has quadrupling in fifteen years).[1] The eagerness to regulate every new innovative use of data have created regulatory divergences between the economies. Even the trade agreements that are supposed to curb these divergences are fragmented and impose different standards due to irreconcilable policy objectives.

Internet is not the first time in history where a pre-existing model of global governance is caught in a dilemma between maintaining an open economic order, and a sovereigns' right to regulate. But the mismatch between internet and global economic governance is a unique challenge: The rule based system is based on a "bottom-up" approach, that integrates national markets through various instruments of cooperation between them. However, internet was already an open and seamlessly global architecture by the time it became relevant to the trading and financial systems. Hence, bilateral or regional integration (perhaps best exemplified by the Digital Single Market in the European Union) could lead to fragmentation by atomising an open structure that was already global at onset.[2]

This note illustrates how fragmentation occurs across several layers of the economy, serving national objectives on security, political authority and market stability. Such objectives go beyond historical pretexts for economic protectionism. So far, 'hard', strategic objectives have

[1] For a full catalogue of data localisation measures, see ECIPE *Digital Trade Estimates*, accessed at: http://ecipe.org/DTE

[2] Legrain, Lee-Makiyama, *Open Up: How to Fix the Flaws in the EU's Digital Single Market*, OPEN, 2017

trumped the self-punitive damage brought by fragmenting the internet, where data localisation generate net economic losses from 0.7 to 1.7% of GDP, from severe productivity losses.[3]

With few other incentives, digital trade barriers are difficult to address even amongst jurisdictions with similar interests and sensitivities. Negotiations amongst like-minded countries do not necessarily generate positive outcomes. This policy-induced balkanisation is therefore unlikely to be addressed in existing forums for economic cooperation and in the prevailing climate of economic diplomacy.

But fragmentation does not just restrict new services – it is an undoing of the existing framework and revocation of existing liberalisation achieved in trade, investment and taxation, and here lies culmination of the mismatch between internet and governance:

- As 56% of international trade in services relies on access to data,[4] market access in offline services (typically banking, professional services, transports and retailing) can be revoked by simply restrict access to data, despite prior commitments to liberalise such services. This condition has achieved a roll-back of existing GATS and FTA schedules.[5]
- Similarly, notion of 'digital presence' allow tax authorities to withdraw from the territoriality principle on taxation and tax entities that are outside their jurisdiction.[6] As market access via commercial presence (mode 3 in trade parlour) is far more restrictive than cross-border modes of supply, extraterritorial taxation impels towards less cross-border economic exchange;
- On investments, the current provisions against performance requirements in BITs can be easily circumvented through privacy and financial regulations, forcing investors to place their operations in the host country.

2. Taxonomy of fragmentation – extraterritoriality, technical, regulatory and commercial fragmentation

[3] Bauer, Lee-Makiyama, van der Marel, *The Costs of Data Localisation: A Friendly Fire on Economic Recovery*, ECIPE, 2014
[4] Based on assumption used first by *UNCTAD Information Economy Report*, UNCTAD, 2009
[5] Lee-Makiyama
[6] OECD Addressing the Tax Challenges of the Digital Economy, OECD, 2014; see critique thereof, Lee-Makiyama, Verschelde, OECD BEPS: Reconciling global trade, taxation principles and the digital economy, ECIPE, 2014

The conflict between the global nature of internet and the territorial nature of law has led to disputes between different state jurisdictions, producing conflict of forums or inconsistent results. The internet has become subject to a myriad of overlapping jurisdictions and conflicting obligations. Unlike other aspects of international law (e.g. law of the high seas) domestic laws are routinely enforced extraterritorially on online activities. Extraterritorial jurisdiction is often based on the nationality of the legal subject, i.e. a natural person who is a citizen, or a corporation is headquartered in the jurisdiction.

For example, the US tax code is based on worldwide income, that created the current problems of deferment of profit remittances from abroad. Similarly, US Department of Justice has claimed – albeit unsuccessfully – its jurisdiction over e-mail data stored on Microsoft's servers overseas based on the Stored Communications Act (18 U.S.C. §§ 2701) in a criminal investigation.[7]

But the most consequential case of extraterritorial jurisdiction over online space is found in the EU, which typically avoided extraterritoriality.[8] But the General Data Privacy Regulation (GDPR) is applied worldwide for personal information on any European citizen:[9] Applicability of GDPR is not territorially limited, and prohibits international transfers of personal information. Exceptions are limited to jurisdiction that the EU deems to have 'adequate protection', or by using legal instruments (binding corporate rules and model contracts) that impose strict liability for data processors and controllers that transfer the data.

Europe's fragmenting approach is beginning to establish a template for privacy regulation worldwide. In contrast to Europe, China goes extraordinary lengths to avoid extraterritoriality – yet produce similar results. The Great Firewall of China (or Golden Shield, as it is called within China) was initially a technical gateway for monitoring and controlling all internet traffic passing through Chinese borders. The Great Fire Wall balkanised the internet *technologically* rather than through extraterritorial applications of Chinese security laws to the rest of the world. Numerous other examples of *technical fragmentation* exist, such as the long-practiced online censorship in

[7] *Microsoft Corporation v. United States of America*, 829 F.3d 197 (2d Circ. 2016); rehearing request by US Department of Justice *en banc* denied, No. 14-2985, 2017 WL 362765 (2d Cir. Jan. 24, 2017)

[8] Blocking of sales of Nazi memorabilia in *Yahoo v LICRA*, TGI de Paris, 2000; US video streaming of a fashion show where certain logotypes were visible in a manner that violated French copyright laws, but falling under fair use in the US in *SARL Louis Ferarud v Viewfinder*, 489 F 3d 474, New York, 2007

[9] General Data Protection Regulation, Regulation 2016/679

some religiously conservative countries, to the more recent political censorship of Wikipedia and social media in Turkey.[10]

However, China's case differs greatly from Turkey. China had made several relevant commitments in its accession to the WTO for some of the most common online services,[11] and evidently had access to less-trade restrictive censoring techniques (thereby failing the two-tier test of GATS art XIV).[12] As a result, China has gradually moved towards a *regulatory* fragmentation rather than a technical one. China has introduced the Internet Content Provider (ICP) licence, a positive list of services that are deemed safe to use by the Chinese public, while other services may be subject to shut-downs. A licensing regime is more consistent with WTO rules thanks to its weak disciplines on domestic regulation (GATS art VI:4). Foreign investors were also restricted from operating wholly-owned e-commerce or voice over-IP services in China as such services require licenses for value-added telecom services (VAS). Clearly, such regulatory measures have both commercial and public security objectives. China's industrial policies on using indigenous, "secure and controllable" technologies and extremely strict requirements for participation in government procurement support the same dual objectives.

In other countries, the regulatory fragmentation supports objectives have justifications that appear equally uncompromising: A majority (58%) of data localisation measures are due to privacy regulations,[13] based on public perceptions of 'fundamental human rights',[14] an argument that has been proven to be difficult to counter by pointing to their economic costs. Other causes of regulatory fragmentation – such as copyright (disabling content portability across border) or banking regulations (financial supervisors demanding localisation of account data) are by their very nature national instruments confined to their jurisdiction. Such cases of localisation are even exceptions of supranational entities like the EU, addressing geo-blocking only for *pro tempore* cross-border use.

But even in the case where fragmentation does not serve 'hard' national objectives, digital protectionism differs from traditional protectionism, making them more complex to address. The post-war industrial policy engaged in regulatory protectionism to foster national champions,

[10] Turkeyblocks.org, *Facebook, Twitter, YouTube and WhatsApp shutdown in Turkey* and *Wikipedia blocked in Turkey*, 2017, accessed at: http://Turkeyblocks.org
[11] Online processing services (CPC843)
[12] Hindley, Lee-Makiyama, *Protectionism Online: Internet Censorship and International Trade Law*, ECIPE, 2009
[13] See note 1
[14] See *inter alia* EU GDPR, art 45 for international transfers

but online protectionism does not always follow that logic. To start, traditional protectionism would be pointless for the digital economy that rewards economy of scale in demand (ability to aggregate users), not production (a large factory that enable cheap production and exporting the surplus). For example, Germany's Industrie 4.0 strategy is built on a logic that the country must slow down competition through restrictive intermediary liability to cope with necessary reforms to protect its manufacturing supremacy and domestic media ownership – not necessarily to develop German search engines or social media.

Similarly, some of China's online protectionism is often linked to SOEs as they happened to be a fiscal income source for Chinese provinces, which are prohibited by the central government to raise taxes. Sectors where SOEs were absent (e.g. car-sharing, e-commerce) have been largely left unregulated, or the first sectors be liberalised for foreign ownership. Inability to decentralise China's fiscal structure thereby defers online reforms. Similarly, protectionism of online payments and *fintech* is linked to lack reforms of Chinese capital account and its banking sector that are constantly on the verge of systematic collapse.

Aside from such examples of commercial *objectives* for protectionism, *commercial* fragmentation by abusing pricing and other commercial terms. Absence of fair, reasonable and non-discriminatory (FRAND) terms for interconnection between a foreign and domestic telecom operator bars infrastructural and business services to provide a global service.

Commercial fragmentation by telecom operators often involves telecom SOEs, or wholesale prices that are set a national regulator (as in the *Telmex* case).[15] But non-state commercial entities could achieve same degree of fragmentation, if one local provider is allowed to dominate a market, or if all local telecom operators are colluding. Such allegations have been made against the US telecom and internet markets by foreign entities.[16] Such barriers are horizontal antitrust issues between private players. Similarly, network prioritisation is dominance abuse by an upstream player against a downstream one.

In this context, it should be noted that commercial fragmentation is the only kind of fragmentation that has been reasonably addressed using existing instruments: Antitrust laws

[15] Mexico — Measures Affecting Telecommunications Services, DS204
[16] FCC, WC Docket 16-143 and Docket 05-25, filed by the European Delegation to the United States, accessed at: https://ecfsapi.fcc.gov/file/10419110631001/Ma419.pdf

generally afford national treatment to foreign complainants, and effective WTO remedies against horizontal anticompetitive practices exist in the GATS Telecom Annex, albeit underused.

3. Whither trade governance?

In absence of other effective remedies, extraterritoriality is the new international customary law. This is particularly true for privacy law, an area which is forcefully advocated by the EU. But indirectly, the US is also arguing the case for data localisation and much more fragmenting privacy laws in Russia, Vietnam, China and India. Meanwhile bilateral instruments like adequacy decisions, only enforce existing extraterritorial regimes, rather than become a construct of free internal exchange amongst the signatories, as data is not allowed to flow to a third country. In that regard, they are similar to the limited reach of bilateral tax agreements.

Mutual legal assistance and extradition treaties (MLATs) could have curbed the need for extraterritoriality to address cybercrime, terrorism and privacy violations. However MLATS are today largely discounted. There is a lack of expediency, trust, and a great difficulty in achieving normative harmonisation on privacy and criminal law, making them impractical tools – which was demonstrated between two like-minded countries like Ireland and United States in *Microsoft v. United States*. This is also why harmonisation of privacy laws in international forums like APEC have its natural limits: As regulatory divergences are simply too wide, they contend to best endeavour guidelines based on minimum standards and proportionality. Enforceable rules under the WTO or other multilateral forums seem far off: After all, this is a world where even the 82 signatories of the ITA agreement cannot agree on the most basic non-tariff measures for electrical interference.[17]

As the economic and judicial cooperation fails to address fragmentation, trade disciplines against data localisation and data flows have been singled out as the only way forward – at least to deal with *regulatory* fragmentation. But FTA/RTA negotiations on these matters are effectively about expanding the exceptions, in particular for privacy, security and politically sensitive sectors: A hypothetical renegotiation of GATS art XIV and GATT art XX would most certainly lead to worse results than today.

[17] Electro-magnetic interference and compatibility (EMC/EMI) have been reformed to self-declaration of conformity (SDoC) practice.

Moreover, final TPP texts left generous exceptions for financial services, while the EU is keen to exempt privacy from the two-tier test – or move the burden of proof to the complainant. There are far-reaching consequences of such reversal as securing evidence of bad faith and behind a privacy law, or to prove that its intent is mere disguised protectionism, ought to be impossible. Any data localisation measure currently in place stand a scrutiny against such lax standards.

Given the sensitivities on personal information, one could foresee an argument that such information can be separated from other data *objects*, such as industrial data. The argument is that trade agreements could at least liberalise industrial use of data for the time being. Nonetheless, over 75% of all data online is user-generated,[18] making the majority of data flows personal information by default; the ‘industrial use of data' also involves personal data like delivery addresses, information on customers or personnel, as human operators are often logged in while collecting, processing or uploading machine data.

Given the very broad definition of personal data in recently enacted privacy laws, almost any industrial and business data could fall under its scope. All forms of data are also integrated and collated in a data object (say, a file): There are no technical or legal means to separate non-personal information (numbers in a spreadsheet) from non-personal information (author of the spreadsheet embedded in the code). This is the very much the purpose of regulatory fragmentation – to create discretionary powers for an executive to act as gatekeepers to the market by selectively enforcing burdensome rules. Fragmentation has now established “license to operate” regimes, where the executive sets up a positive list of commercial entities that are allowed on the market hinged on nationality or performance requirements.

4. Conclusions

With over 1300 barriers identified affecting the digital economy in a sample of just 65 countries, one could soon argue that we are a *fait accompli,* as there are too many barriers for international treaty negotiations to handle. Economic argument does not seem to sway ‘hard' objectives, such as security or fundamental rights. Economic arguments are sometimes even

[18] Austin, Upton, Leading in the Age of Super-Transparency, *MIT Sloan Management Review*, Winter 2016

futile for economic objectives – a draconic online tax law is paid through loss of GDP, in other words corporate revenues and consumer welfare, while governments may actually see their tax base increase. Public choice dilemmas arise as there are different incentives between the public authorities and its subjects.

Third countries find it difficult to incentivise against fragmentation, as balkanisation are consequences of unique structural problems in the underlying economy or the political system. This is the case of the fragmentation caused by both the EU and China.

However, this note is not to provide a justification to fragmentation just because they are uncompromising – but to map why traditional economic diplomacy has so far failed.

In the new political dimension of trade negotiations post-TPP and TTIP, like-mindedness is no longer a recipe for ambitious EPA/FTA outcomes. In fact, similarity is an impediment to successful conclusion of FTAs: Homogeneity (the extent barriers are imposed in same areas) lead to weak outcomes in intra-EU cooperation such as DSM. Regulatory divergences amongst the signatories of TTIP and TISA were narrower than TPP where parties imposed high barriers in completely different regulatory areas.

With no effective cooperation instruments for global openness and rule of law, the global governance system is at a lose-lose situation. As the actors cannot offer credible incentives or threats, and they are left with very few policy options but to block their own economy on reciprocal basis, and thereby contribute to further fragmentation.

Section 4:
Internet Governance

The Future of Global Cyber Trust:

Fragmentation v. Universality Tradeoffs[i]

May, 2017

By Dr. Laura DeNardis[ii]

1. Introduction

Commerce, speech, social life, and every imaginable industrial sector are now digitally mediated and therefore contingent upon the security and integrity of Internet infrastructure. Emerging technological advances such as cyber physical systems, cryptocurrencies, and artificial intelligence raise the stakes of network stability significantly. What are the implications of these trust dependencies on modern society and the Internet itself? Until societies experience economic or social upheaval, the role of trust in maintaining societal stability exists as a taken for granted background context of daily life. Individuals trust that financial institutions will secure their bank accounts, cars will not malfunction, airplanes will stay in the sky, and medical test results remain confidential. Democracies depend upon the integrity of voting systems and commercial transactions rely upon trust between buyers and sellers. What has changed in recent decades is that all of these trust dependencies now also depend upon the integrity and security of underlying digital infrastructure.

Even while societal dependencies on digital infrastructure mount, there is evidence of some loss of trust in this very infrastructure and its governing institutions. Some of this loss of trust stems from actions in the political realm, whereby governments establish policies, such as data localization laws or national cybersecurity measures, to enhance national sovereignty or address privacy concerns about foreign intelligence gathering practices. Loss of trust among Internet users arises from rising awareness of government surveillance and private sector data gathering practices, as well as high-profile cybersecurity breaches, including the massive data breaches at Yahoo!, Target, and the US Office of Personnel Management (OPM).

The 2017 CIGI-Ipsos Survey on Internet Security and Trust, polling more than 24,000 users in 24 countries, found that a majority of respondents were more concerned about privacy than they had been in the previous year, partly related to cybercrime but, increasingly, also due to concerns about their own governments (CIGI-Ipsos 2017). The poll indicated that only half of respondents trust their governments to act responsibly online.

Trust has always been a requirement for keeping the Internet operational, but society is approaching a tipping point in which significant improvements in digital trust are necessary to sustain a global digital economy and public sphere. Indeed, many of the most contentious global policy issues in the cyber arena involve struggles over trust: in the stability of infrastructure, voting systems, digitally mediated news, the security and privacy of user data, the authenticity of information and users, and commercial transactions. Not surprisingly, considerable policy and scholarly attention has focused on these issues, and especially, the close association between cybersecurity technologies and trust policies (Schneider 1998, Singer & Friedman 2014, Hampson & Jardine 2016).

Constructions of trust in cyberspace will affect whether the Internet continues to expand into a universal network or fragment into segments enclosed by geopolitical borders or proprietary market ecosystems. A great deal of policy and scholarly attention has examined tensions between Internet universality and fragmentation (Werbach 2008, Force Hill 2010, DeNardis 2016, Drake et al., 2016, Mueller 2017). What has been addressed less is the more narrow policy intersection between cyber trust and fragmentation. Can digital trust and Internet universality co-exist in the long term in light of technological and geopolitical changes facing the Internet? There is a moment of opportunity to examine intersections between digital trust and fragmentation and explore which future solutions – public policy, market approaches, civil society interventions, and technical design – can foster the trust necessary for the stability and security of digital systems while also enabling a universal Internet supporting digital trade, freedom of expression, and access to knowledge.

2. Digital Trust Points as a Precursor to Internet Universality

The Internet is not a single network but an interconnected collection of mostly privately owned networks able to interoperate because they adhere to common sets of standards for formatting and exchanging information. Trust between network operators has always been a requirement for this interconnection, just like trust between trading partners is necessary for the global digital economy to function. Each autonomous system advertises the routes (i.e. collections of Internet Protocol addresses) reachable through that network using Border Gateway Protocol (BGP). Historically, network operators have trusted adjacent networks to advertise accurate routes, although security breaches certainly occur at these borders. The ability to access information on a website from anywhere in the world similarly depends upon trust in the Internet's Domain Name System (DNS), the globally distributed system that translates domain names into corresponding Internet addresses locating information online. Trust in the DNS is a necessary precursor for the Internet to globally operate. Technical infrastructure trust mechanisms such as public key cryptography authentication are increasingly engineered into these systems.

Even though the digital economy has experienced tremendous growth – the Internet has more than 3 billion Internet users and contributes more than $4 trillion USD to the global economy – the Internet is not yet universal. Viewed through the lens of physical infrastructure and bandwidth, nearly half the world still does not have access and, among those who do, access speeds vary considerably (ITU 2015). At the logical, software-defined layer of the Internet, there is also fragmentation, such as the use of the DNS to carry out censorship and other content controls. At the application and content layer, the Internet is not yet universal because of language differences, including barriers to universal accommodation of internationalized domain names (IDNs) that incorporate non-Latin characters such as those used in Arabic, Chinese, and Cyrillic text. Regional policies block content locally, such as the Right to be Forgotten in the European Union, the geo-IP restriction of Netflix in Canada, and systems of censorship and blocking in China and elsewhere. Fragmentation of networks for security reasons, via firewalls and virtual private networks, is of course the norm for most corporate networks. This choice to create fragmentation for security reasons is quite distinct from fragmentation that is not a user choice. Overall, the Internet has continued to expand globally because of trust among networks, between websites and browsers, and in common technical standards and systems of routing and addressing.

3. Geopolitical Trust Tensions Are Creating Fragmentation

Despite the historical growth trajectory of the Internet, several geopolitical trust problems are creating digital fragmentation. Values of privacy, security, and national sovereignty increasingly conflict with values of universality and the free flow of information across borders. Some of these conflicts arise from problems of jurisdiction, as well as incongruities between technological and nation-state boundaries. The virtual architecture of the Internet and the cross-border nature of data flows are often incommensurable with political borders. While routers make decisions about the flow of information based on engineering optimization rather than geography, what counts as privacy, hate speech, indecency, and freedom of expression, differs greatly across geopolitical borders. Legal authority over citizens and institutions within borders does not comport well with the cross-border and distributed nature of cyberspace. Interoperability and harmonization of Internet policies across borders can prevent Internet fragmentation, but cultivating cultural and political agreement on many Internet policy issues can be an intractable problem, even in areas such as intellectual property rights enforcement and cybercrime. The jurisdictionally complex task of enforcing laws often falls to private intermediaries, creating a privatization of governance unprecedented in the contemporary era.

A trust-related example of attempts to harmonize national borders with virtual borders involves the introduction of data localization laws placing constraints on how private companies (e.g. banks, retail, or technology companies) handle customer data, including requirements that data be stored on servers within a nation's borders (Chander and Le 2015). The rationales for these policies often cite concern about customer privacy in the context of foreign surveillance, even though concentrating data in a fixed location can facilitate efficient surveillance and create a host of technical complexities and economic costs (Bauer, et al. 2016).

Governments increasingly view control of Internet infrastructure as a proxy for state power, whether motivated by national security, cyber war concerns, censorship, or economic objectives. China and other countries seeking greater control over information flows have advocated for top-down, bordered, government-centric cyber sovereignty approaches that supplant traditional private sector led governance approaches in the name of cyber order (DeNardis, Goldstein and Gross 2016). Some of these efforts to assert cyber sovereignty arise from lack of trust in the institutions that govern the Internet and raise the possibility of fragmentation not only of digital networks but of the global governance structures tasked with keeping networks operational.

4. Emerging Trust Terrains: IOT, Currency, and AI

Emerging technological innovations raise the stakes of digital trust and also challenge some prevailing assumptions that the goal of a universal Internet is always in the public interest. Internet of Things (IOT) projections envision the ability to interconnect an estimated 50 billion objects to the global Internet. The diffusion of the Internet into material objects - remote sensor devices, health monitoring devices, home appliances, traffic systems, and networked vehicles – raises the stakes for digital trust. For example, a disruption of a network-connected cardiac implant threatens human safety rather than simply the ability to communicate. Digitally dependent and digital-only cryptographic currencies also continue to gain traction, often outside of traditional regulatory frameworks. What trust mechanisms are necessary to preserve confidence, integrity, and security in financial systems? As decisions about how information is organized and how data is analyzed move to machine learning and artificial intelligence systems, new systems of accountability and human safety will be necessary to instill trust in digital environments.

5. Framing Questions for the Panel

Fragmentation as a Context-Dependent Value. Given threats from cyberattacks, cybercrime, and geopolitically motivated Internet conflict, and considering that the cyber realm now includes industrial control systems, medical devices, vehicles and other human safety-related contexts, is fragmentation necessarily something that should be minimized? Conversely, in highly trust dependent areas, under what conditions is fragmentation actually desirable?

The Tension between Privacy/Security and Universality. Can values of privacy and security, and the trust solutions necessary to sustain these values co-exist with norms of Internet universality?

Trust as a Precursor for Universality. Where Internet universality has positive economic and social effects (e.g. freedom of expression, global commerce), what are the most pressing trust dependencies necessary for the growth of the global digital economy and digital public sphere?

Trust Solutions. What solutions - in technical architecture, market approaches, government policies, and international agreements – hold the most promise to create trust conditions necessary for an appropriate balance between Internet universality and fragmentation?

Emerging Trust Dependencies. What policy solutions of today can address emerging technological phenomena such as artificial intelligence, cryptographic currencies, and cyber physical systems?

Reference

1. Bauer, Matthias, Martina Ferracane and Erik van der Marel (2016). “Tracing the Economic Impact of Regulations on the Free Flow of Data and Data Localization,” *Global Commission on Internet Governance Papers Series* No. 30, May. Accessed at https://ourinternet-files.s3.amazonaws.com/publications/ gcig_no30web.pdf.
2. Chander, Anupam and Uyen Le (2015). “Data Nationalism,” *Emory Law Journal*, Vol. 64, No. 3.
3. CIGI-Ipsos Global Survey on Internet Security and Trust, April 2017. Accessed at https://www.cigionline.org/internet-survey.
4. DeNardis, Laura (2016). One Internet: An Evidentiary Basis for Policy Making on Internet Universality and Fragmentation, Global Commission on Internet Governance, GCIG Paper No. 38, July 19. Accessed at https://www.cigionline.org/publications/one-internet-evidentiary-basis-policy-making-internet-universality-and-fragmentation.
5. DeNardis, Laura, Gordon Goldstein, and David A. Gross (2016), “The Rising Geopolitics of Internet Governance: Cyber Sovereignty v. Distributed Governance,” Columbia SIPA Working Paper, November 30.
6. Drake, William J., Vinton G. Cerf and Wolfgang Kleinwächter (2016) “Internet Fragmentation: An Overview.” World Economic Forum Future of the Internet Initiative White Paper, January. Accessed at www3.weforum.org/docs/WEF_FII_ Internet_Fragmentation_An_Overview_2016.pdf.
7. Force Hill, Jonah (2012). "Internet Fragmentation: Highlighting the Major Technical, Governance and Diplomatic Challenges for U.S. Policy Makers,” Belfer Center for Science and International Affairs, Harvard Kennedy School. Accessed at http://belfercenter.ksg.harvard.edu/files/ internet_fragmentation_jonah_hill.pdf.
8. Hampson, Fen Osler and Eric Jardine (2016). Looks Who’s Watching: Surveillance, Treachery and Trust Online, Center for International Governance Innovation (CIGI) Press.
9. ITU (2015). International Telecommunication Union (ITU), “ICT Facts and Figures – The World in 2015,” 2015. Accessed at http://www.itu.int/en/ITU-D/Statistics.
10. Mueller, Milton (2017). *Will the Internet Fragment? Sovereignty, Globalization and Cyberspace*, Polity Press.
11. Schneider, Fred B., ed. (1998). *Trust in Cyberspace*, National Academy of Science Press.

12. Singer, P.W. and Allan Freidman (2014). *Cybersecurity and Cyber War: What Everyone Needs to Know*, Oxford University Press.
13. Werbach, Kevin (2008). "The Centripetal Network: How the Internet Holds Itself Together, and the Forces Tearing it Apart," *University of California Davis Law Review*.

[i] Background Thought Piece - Digital Futures Policy Forum Panel: Developing Trust and Assurance

[ii] Adjunct Senior Research Scholar, Columbia SIPA, Professor, American University

08. **Doomed to Fragment? Addressing International Security Challenges While Avoiding Internet Fragmentation**

Nikolas Ott and Hugo Zylberberg

Introduction

Considering the recent spike in news coverage on ransomwares, hacks, cyber attacks, data breaches and intrusions, it is easy to forget the significant economic and social opportunities that digital transformation can provide on a global scale. New innovations, as well as ubiquitous connectivity around the world, are reshaping technology and its role in people's daily lives. In turn, digital transformation opens up new economic and social opportunities: the sharing economy, decentralised crowdfunding platforms, and accessible global communications have the potential to increase political stability worldwide.

This is true both in the developed and in the developing world. The digital economy "contributed $2.3 trillion to the G20's GDP in 2010 and an estimated $4 trillion in 2016, [and] is growing at 10% a year - significantly faster than the overall G20 economy."[1] Moreover, there is evidence that connectivity drives growth in a development context. As the World Bank report on digital dividends states: "For businesses, the internet promotes inclusion of firms in the world economy by expanding trade, raises the productivity of capital, and intensifies competition in the marketplace, which in turn induces innovation. It brings opportunities to households by creating jobs, leverages human capital, and produces consumer surplus. It enables citizens to access public services, strengthens government capability, and serves as a platform for citizens to tackle collective action problems."[2] As connectivity becomes a crucial factor for economic development, the security-development nexus is increasingly being recognised as a key sustainability factor.[3]

As the 2016 World Economic Forum (WEF) report on internet fragmentation correctly outlines, these economic and social outcomes rely on the "Internet [remaining] stable and generally open and secure in its foundations."[4] Yet, the model for cyberspace governance can hardly be that of one uniform internet. In the spirit of the inventors of the internet,[5] states should aim at producing interoperable policy frameworks allowing the possibility of governance across stakeholders, while leaving states in charge of implementing these frameworks at the national level.[6] The WEF report identifies 28 issues of current or potential fragmentation along three buckets: technical, commercial, and governmental fragmentation. This paper focuses on governmental fragmentation, which refers to governmental rules that hinder the introduction or further development of international policy guidelines, or that affect the perception of a unique network.

When it comes to both national and international security concerns, the international institutions governing cyberspace face a dilemma: they cannot fully satisfy all relevant stakeholders at the same time. Finding the right balance between the interest of states, the private sector, and citizens is a delicate process that is deliberated in various multistakeholder fora such as the Internet Governance Forum (IGF). The discussions touch on themes as diverse as data protection, privacy, freedom of expression and law-enforcement responsibilities. However, so far, addressing security challenges in cyberspace through such fora has had limited results. Instead, influential states, such as the United States (US), Russia, China and France, are trying to address such challenges through national legislation with extra-jurisdictional reach. While national legislation might seem easier for states, they tend to worsen governmental fragmentation and further complicate the creation of international procedures addressing global cyber security challenges.

Recently, much of this governmental fragmentation appears to be driven by security concerns: be it in France where a filtering system for jihadist websites was implemented in 2015,[7] in Germany where a recent law forced platforms to remove obviously illegal hate speech[8] in a context when fake news and the security of election infrastructures is under question, or in a series of other countries proposing to ban end-to-end encryption. This is by no means limited to authoritarian states and has become an issue that concerns policymakers around the world, as reflected in the joint anti-encryption opinion piece penned by the Manhattan district attorney, the Paris chief prosecutor, the commissioner of the City of London Police, and the chief prosecutor of the High Court of Spain,[9] and more recent declarations of the UK Home Secretary against terrorist usage of end-to-end encryption.[10] Taking stock of the security rationale for such government policies that will further increase internet fragmentation, this paper argues for the establishment of interoperable policies, through a holistic "fragmentation impact assessment" and increased involvement in international security discussions to limit what this paper labels as "security-based fragmentation": governmental fragmentation related to international and national security in and through cyberspace, which also includes security incidents relying on legitimate uses of cyberspace.

States' Westphalian Notion of "Sovereignty" in the Digital Age

The current internet governance structure (a multistakeholder governance framework) is ideologically and conceptually at odds with the Westphalian notion of states' sovereignty in its current understanding and practice. While in most states, multinational technology companies have a crucial role in ensuring the accessibility and the maintenance of cyber infrastructure, this does not automatically give them a role within the international policy decision process. One could rightfully argue that the states' permission to integrate technology companies and civil society in these negotiations is an exercise of their sovereignty.[11] Indeed, many non-state actors are now involved in the practical application of international law to cyberspace, through 'Track 1.5' dialogues[12] or efforts such as the Tallinn Manual,[13] where leading academics assess how existing international law apply in cyberspace. Despite many states being uncomfortable with this development, the fact that a large part of the infrastructure is owned and operated by the private sector and loose communities of researchers makes their participation crucial to advancing international discussions.

To understand the challenges that states are facing, it is necessary to further clarify the different concepts surrounding cyber security. Broadly speaking, these can be captured in four categories:[14] international security, national security, device security and data security.

1. **International cyber security** focuses on interstate issues of cyber conflict. Policies in this category include: exchanging national security doctrines, creating communication channels, and reviewing the applicability of international law in cyberspace. The most active fora for these policy discussions are the United Nations Group of Governmental Experts (UNGGE)[15] and the Organization for Security and Co-operation in Europe (OSCE),[16] though other fora, such as the Organization of American States or the Association of Southeast Asian Nations (ASEAN) Regional Forum are contributing to these discussions as well.

2. **National cyber security** addresses the challenges of intelligence agencies, law enforcement, policing and other entities that are responsible for addressing crimes committed in and through cyberspace. In addition to national entities, the Budapest Convention on Cybercrime and INTERPOL, too, play a crucial role in facilitating cooperation and information exchange between state entities.

3. **Device security** focuses on the integrity and stability of internet infrastructure and related cyber-physical systems: systems in which "operations are integrated, monitored, and/or controlled by a computational core."[17] Related efforts are mostly technical and led by national institutes for standards and technology, or offices for information security, within large multinational technology companies.

4. **Data security** mostly centres on maintaining security and privacy throughout the data lifecycle: collection, storage, treatment (or processing) and use. Few states have dedicated agencies for privacy issues but some have special commissioners or governmental representatives to assure proper inclusion of privacy concerns in related policy discussions.[18]

This piece focuses on the first category of cyber security: international cyber security. Unfortunately, few policy discussions draw upon these distinctions. One example is the current discussion about a digital Geneva Convention, brought forward by Microsoft. The current proposal covers several of the aforementioned categories at the same time, which makes it difficult for policymakers to properly address the proposed changes, given the lack of compatibility with existing policy structures. However, it is important to note that decisions taken within the realm of international security have both direct and indirect effects on the other categories. For example, a discussion between states can have an impact on multinational technology companies that operate globally and rely on internationally recognised procedures, certification standards, or treaties. At the same time, interstate negotiations affect the daily work of national law enforcement entities that rely on productive interstate relations.

Despite the inherent borderless nature of cyberspace, most policy solutions to date are tailored on a national (Russia/China) or regional (European Union) basis. States seem to 'muddle through' instead of working with non-state stakeholders towards suitably interoperable actions. This is especially true for international and national cyber security issues. However, more recently, multinational technology companies have been trying to contribute to this policy debate as they are increasingly affected by its outcome. This is reflected in ongoing legal discussions about the legality of access for states on data stored in another country. Microsoft's lawsuit against the US government over rightful access of data is only one out of many cases where companies come into conflict with government demands for access to data stored abroad.[19] As cloud computing is expanding drastically, it is reasonable to expect that overall technological developments introduced by the private sector have and will most likely continue to outrun the pace at which policy decisions are made. Therefore, multinational technology companies should continue to play an important role in the development and implementation of security policies that affect cyberspace.

Whether it is because states operate on the assumption that policies that increase fragmentation are necessary to maintain their security in cyberspace across all four aforementioned categories, or because fragmentation is an unanticipated second-order effect of their policies, it seems that this security-based fragmentation has indeed been on the rise. This paper now examines the assumption that fragmentation can lead to better security, before proposing a framework promoting interoperable policy frameworks to avoid it.

An Increase in Internet Fragmentation does not Necessarily Lead to Better Security

States' practice has shown that the restriction of cross-border data flows[20] for privacy or security reasons, and increased power to lawfully access this data is becoming more widespread, even as the extent of such restrictions and surveillance is being debated. The European Union's (EU) General Data Protection Regulation (GDPR),[21] associated with the negotiation of the Privacy Shield agreement, creates a framework whereby data flows are restricted towards countries where the data protection framework is too weak. Another example of this trend is the United Kingdom's Investigatory Powers Act,[22] which requires internet and phone companies in the UK to maintain the capability to intercept their customers' personal data; this is unlikely to be the case in other countries, including in Europe. The UN Special Rapporteur on the Right to Privacy, Joe Cannataci, recognised this growing trend in a recent report,[23] calling for an international treaty to protect people's privacy from unfettered cyber surveillance. However, such calls mostly address the fourth of the previously introduced categories, namely, data security. While ensuring citizens' privacy deserves significant attention, the increasing friction between states within cyberspace needs more attention as well.

The belief that a more fragmented internet—bringing borders to the digital realm—leads to a more secure interstate environment is flawed, for three main reasons:

First, it is currently much harder to secure a network than to attack it.[24] While this mostly affects device security, it also entices states to engage in deterrence-based cyber security strategies through the development of offensive cyber capabilities. As a well-resourced and motivated attacker always succeeds, digital borders at the national level will be bypassed just as physical borders, i.e. bypassing firewalls. This leads to a perpetual state of insecurity that can currently only be addressed through diplomatic means, such as confidence-building measures and legal agreements.

Second, tools to circumvent national borders (e.g. virtual private networks) will continue to appear and be used precisely by those actors who present the most serious security threats. Moreover, prohibiting or limiting the use of end-to-end encryption will take it away from regular people and companies that rely on such security measures. On the other hand, terrorists, criminals and other nefarious actors will eventually find new ways to avoid surveillance efforts. Therefore, efforts to limit the use of such tools are not just ineffective in the long term, as adversaries adjust, they also negatively affect data and device security in the short term.

Third, cyberspace is the domain, not the source of security threats. As countless government reports have argued, governmental shortcomings in the security realm do not come from a lack of institutional capacity to collect data, but from a lack of integration and coordination between law enforcement, the justice system, and the intelligence community. This is a long mission that the United States (US) started ahead of other countries in the wake of 9/11, by creating the Office of the Director of National Intelligence,[25] but the ongoing discussion on the proper division between military (US Cyber Command) and espionage (National Security Agency) activities shows that the debate is far from being concluded. Ultimately, it is important to highlight that security measures most often fail due to human, not computer, error.[26] While such concerns generally affect domestic cyber security policy, a lack of such domestic capacity significantly hinders the ability of states to engage in constructive interstate dialogues. Consequently, having a comprehensive national cyber security strategy is highly desirable to further increase the likelihood of successful international negotiations. Moreover, it is especially important to get national cyber security policies right, to be properly prepared for a cloud-based and borderless operational environment, where international cooperation on law enforcement and other issues are becoming even more important for properly addressing security challenges within this domain.

Calibrating A Multistakeholder Discussion on Security-Based Fragmentation

Even though more government control can help secure cyberspace in the short term, it is often unlikely to do so in the long run. As technologists weighing in on the debate over backdoors have shown,[27] short-term solutions (developing a system where law enforcement is able to access any system given judicial authority) can eventually be subverted by malicious actors for their own purposes, undermining global cyber security. While short-term issues are crucial in a world where serious security threats can put human lives at risk, any solution must take into account the consequences of enabling malicious actors to gain state-level mass surveillance capacities. Developing partnerships with the private sector is a crucial element of any potential solution. Without developing new infrastructure-enabling mass surveillance, security services can often find the data they need in existing privately-owned infrastructure. Therefore, some countries have now adopted the position that instead of laws requiring companies to give them access to their servers, they can be satisfied with a point person available at all times to help with urgent requests related to national security.

In addition to this balance between short-term and long-term concerns, international discussions on cyber issues need to consider their own impact on the security and stability of cyberspace. Indeed, policy choices affect cyberspace stability, and conversely, a state's evaluation of its stability affects its policy choices. Inspired by the recent publication of Laura DeNardis through the Global Commission on Internet Governance,[28] this paper suggests that in the same way that companies have to produce privacy impact assessments or human rights impact assessments, fragmentation impact assessments (FIA) could be developed for policies that appear to drive fragmentation in an excessive fashion. These FIAs could include an introduction to the policy being discussed, as well as an evaluation of its impact on the issues below.

Basic principles:

- **Protection of personal data:** All actors should respect fundamental data protection principles giving citizens—not states or companies—power over their personal data.
- **A neutral network:** No technical restrictions at the infrastructural level should restrict which applications the general public can or cannot use.
- **Network generativity:** Should there be any limits to innovation at the end nodes? Principles affecting the private sector:
- **Interoperability:** All services provided online should be interoperable.
- **Industry standards:** Technical standards should not be subverted for national security purposes.
- **Global commons:** Is there a subset of the internet that should be declared a global

commons?

Principles affecting states' behaviour:

- **Data sharing:** States should streamline data-sharing processes between law enforcement, judicial and national security institutions.
- **Integrity of data:** States should not alter the integrity of data, at rest or in motion.
- **Accessibility:** When is it legitimate to block content travelling to one state from another through whatever technical means?

Building interoperable policies regarding acceptable behaviour for states vis-à-vis access to data and public-private partnerships are key to limiting security-based fragmentation. International and regional efforts, such as the Global Commission for the Stability of Cyberspace, the UNGGE and the OSCE or ASEAN, provide platforms to identify common interests and acceptable standards of behaviour between states. Here again, stronger integration of the private sector during policy negotiations, despite the increased difficulty, is key to finding interoperable solutions that work in practice.

In parallel to building interoperable policy frameworks using FIAs, states should develop an understanding of when and where fragmentation can be legitimate. There is a need to find the characteristics of legitimate national regulation with limited externalities on internet fragmentation. Such a discussion could start with the following questions:

- Is there a "public core of the internet"?[29] Governments can agree on a limited set of targets that should be protected from both states and intervention, e.g. the Domain Name Systems or some fundamental internet routing protocols.
- Which components of the internet should be regulated on a national basis, and which ones on an international basis? In areas where states will continue to regulate on a national basis, how can this regulation be made interoperable with others to mitigate the economic cost incurred? International efforts in building policy frameworks in a transnational fashion must be encouraged so that legislation can continue to develop on a national basis but produce outcomes that are increasingly interoperable with neighbouring ones.
- Are there alternatives to satisfy states' security needs that include more or less policy fragmentation? More academic work to understand fragmentation can help states produce FIAs to measure the consequences of a specific policy proposal.
- Where and when does fragmentation matter most? Academic efforts taking stock of existing internet fragmentation, and asking when and where its consequences are most limited, are still lacking.

Conclusion

Despite growing concerns over security incidents in and through cyberspace, the internet still holds significant economic and social opportunities. The securitisation of the current debate compounded by a return of nationalism in the public debate of liberal democracies threatens these promises as well as the very values enshrined in the technical infrastructure and the governance mechanisms associated with the internet. However, this paper argues that some of this securitisation is based on the flawed premise that a fragmented internet with monitored digital borders matching physical ones is more easily defensible.

This paper concludes by recommending questions and characteristics for a global multistakeholder debate, the establishment of FIA, and increased involvement in the development of cyber security policies. Section three outlines how these three recommendations are intertwined and can support each other, namely, questions and characteristics for a global multistakeholder debate that, combined with FIAs and stronger involvement, can better inform policymakers and increase the chances of producing interoperable policy frameworks, thus limiting security-based fragmentation.

Acknowledgements:

For written comments on early drafts, the authors are indebted to Marina Kaljurand, William Drake, Ben Hiller, Henry Rıigas, Aude GØry, Jessica Zucker and Christoph Berlich. An earlier version of this paper was presented at the OSCE-ASEAN Inter-Regional Conference on Cyber/ICT Security.

November 2016

Columbia SIPA
School of International and Public Affairs

Tech & Policy Initiative

The Rising Geopolitics of Internet Governance: Cyber Sovereignty v. Distributed Governance

By: Laura DeNardis, Gordon Goldstein,
and David A. Gross

The Rising Geopolitics of Internet Governance

Cyber Sovereignty v. Distributed Governance

Laura DeNardis, Gordon Goldstein, and David A. Gross

Presented at Columbia School of International and Public Affairs
November 30, 2016

The Political and Economic Stakes of Internet Governance

Internet governance is at a crossroads. The 21st century has given rise to two incommensurable visions for the global Internet and how it is governed. One envisions a universal network that generally supports the free flow of information and whose governance is distributed across the private sector, governments and new global institutions in an approach that has historically been described as "multistakeholder" governance. This vision has materialized, albeit imperfectly, in how the Internet and its coordination has historically progressed and is an approach advocated by the United States government and many other countries. This is the model of Internet governance that has dominated throughout the past decade. The competing vision advocates for greater multilateral and top-down administration of the Internet in the name of social order, national cyber sovereignty, and tighter control of information flows. China and other countries interested in greater administrative control over the flow of information have been vocal proponents of a more multilateral approach to Internet governance. These visions are often debated using the language of abstract theoretical constructs but they involve actual policy choices that have arisen in particular historical contexts and whose future will have tangible effects on American foreign policy interests, American values of freedom of expression and innovation, the global digital economy, and the stability and resiliency of Internet infrastructure itself.

The Internet now ranks high on the policy agendas of governments in countries ranging from Russia and China to the United States and Brazil. In only a decade, concerns about governance of the Internet have transformed from being interesting only to the technical community and a subset of academics to becoming a national policy priority for G20 government leaders. Questions about the control and security of the Internet now rank in global importance alongside topics such as terrorism, climate change, and human rights. Governments and other stakeholders have elevated Internet governance on the global policy agenda because the Internet has become a strategic resource with unprecedented economic, political, and social implications.

The economic stakes of cyberspace are immense. The Internet now contributes more than $4 trillion USD to the global economy. More than this, every sector of the economy from financial services to transportation to health care is completely dependent upon cyber systems for basic day-to-day transactions and functioning. A collapse of the Internet would also be a collapse of the economy. The economic stakes will only increase in the context of the Internet of Things in which every device from cars to medical devices will be deeply integrated with and dependent upon digital systems.

Battles over Internet policy often reflect global interests around geo-economic competition. The Internet is arguably the greatest value creation engine in the history of civilization, with digital technology companies like Apple, Google (Alphabet, Inc.), Microsoft, and Facebook now consistently occupying the upper echelon (within the top ten) of multinational companies in terms of market capitalization. Eventually one will be have a trillion-dollar market capitalization. What complicates global Internet policy is the reality that there is such a diversity of strategic interests around the Internet. The global dominance of American companies emerged in part from the Internet originating in the United States. The globalization and spread of the Internet has seen the rise of even newer companies, such as Alibaba in China, which appropriated aspects of the business models of many US companies such as ecommerce and Internet search.

The rising economic dependency on the Internet exists in a much broader historical context. Globalization and technological change have created deep-seated economic uncertainty about the future of jobs and entire industries. History may demonstrate that the shift from an industrial society to an information society is even more consequential than the shift from agrarian societies to the industrial age. The combination of globalization and industrial transformation that has served as a background context for Brexit and for the election of U.S. President-elect Donald Trump also serves as the context for rising tensions over control of the digital systems on which all economic and political systems now depend. Government interest in cyberspace is in part a response to the rising importance of the Internet but also stems from an interest in providing a stabilizing response to citizen uncertainty over the accelerating pace of technological and industrial change.

The political and social stakes of the Internet may be even higher as control over cyberspace increasingly becomes viewed as a proxy for state power. The 21st century is marked by technical mediation of the public sphere and the condition that cyberspace underpins all functions of government and society. An outage in cyberspace now equates to an outage of government and society, as was first starkly demonstrated when cyberattacks disrupted government and industry systems in Estonia in 2007. Cyberspace is now the fifth domain of warfare, with the stability and security of critical digital infrastructures high on the policy agenda and control over cyberspace an emerging front for state power. Although data breaches against large multinational companies and government institutions such as Target, Yahoo!, and the United States Office of Personnel Management have potential economic consequences for individuals, cyberattacks are often abjectly political, such as the Sony hack or the alleged Russian cyberattacks breaching private email accounts of American Democratic party leaders during the 2016 U.S. presidential election. Individual civil liberties such as freedom of expression and privacy are now already shaped by arrangements of digital policy, but increasingly democracy itself, is becoming contingent upon the stability and security of cyberspace.

With economic stability and political and social systems increasingly dependent upon the Internet, tensions over control of the Internet are also rising. Any longstanding views about the Internet being ungoverned or ungovernable are simply not true. The Internet is, and has always been, governed, albeit not necessarily by traditional governance structures but by a combination of actors from the private sector, new global institutions that cross borders, and also traditional laws and policies. Internet governance is not one system, one control point, or one institution, despite how it is often described by the media and policymakers. Keeping the Internet operational requires a large cadre of mostly private companies and institutions performing various administrative oversight functions. Standards-setting institutions such as the Internet Engineering Task Force (IETF) and the World Wide Web Consortium (W3C) establish common technical protocols that enable interoperability among various systems and devices. Network operators make private contractual agreements to interconnect their networks and exchange traffic. Information intermediaries providing services ranging from social media to search establish conditions of free expression and privacy through user terms of service. New institutions like the Internet Corporation for Assigned Names and Numbers (ICANN) oversee critical Internet resources such as the domain names and Internet addresses that serve as globally unique identifiers for online resources. The growth and success of the contemporary Internet has been largely shaped by private sector-led innovations, investments, and administrative coordination.

Some of the most strident and important geopolitical debates over control of the Internet – such as battles over control of the critical Internet resources of names and numbers, interconnection regulation, and encryption – have involved an essential tension between private sector-led versus government-led coordination and control.

This geopolitical tension can be described as multilateralism and Internet sovereignty versus distributed and private-sector-led multistakeholder governance.

One of the most high-profile conflicts reflecting this tension has been the ongoing struggle for control over the Internet names and numbers and the transition of United States Department of Commerce oversight of the "IANA functions" (Internet Assigned Numbers Authority), and particularly the authorization of changes to the root zone file mapping Internet top-level domains to Internet addresses, the binary numbers computers use to route information to its online destination. The transition of U.S. oversight of these functions has taken place and the Internet continues to operate. After a decades-long deliberation, the baton has been passed to a global multistakeholder arrangement within ICANN itself. Unless governments like China and Russia, possibly via the United Nations structure, find a way to inject themselves into a future oversight function around the IANA functions, this transition appears to be at least a short-term win for multistakeholder governance of the Internet.

This paper argues that the transition of American oversight of Internet names and numbers is the beginning, not the end of Internet conflicts between distributed private-sector led governance and multilateral or even unilateral oversight. Battles over Internet governance are also global battles over political and economic power, not unlike control of sea lanes in the 18th and 19th centuries. The transnational and borderless character of the Internet challenges core notions of political power including government sovereignty. The borderless nature of cyberspace is anathema to many governments.

How can the traditional private-sector-led governance of the Internet co-exist in an era in which governments increasingly view control of cyberspace as a proxy for state power, whether for economic advantage, political control over the flow of information, national security, or as a strategic resource in warfare? This paper provides some historical context to the rise of distributed Internet governance, describes some of the key geopolitical conflicts that involve incommensurability between the ideology of national sovereignty and the technical topology and transnational characteristics of private Internet infrastructure, and argues for the preservation of private-sector-led multistakeholder governance rather than a shift to greater government control. Most importantly, the political and economic battles of the future will turn from control of content and largely symbolic power struggles to deeper layers of the Internet's infrastructure, including technical standards, systems of cryptocurrencies, the Domain Name System, cybersecurity systems, and the Internet of Things. It is in the interest of democratic governments to make policy decisions today that resist cyber sovereignty, focus on cybersecurity and Internet universality, and anticipate emerging battles over Internet infrastructure choke points.

The Evolution of Distributed Multistakeholder Governance

The historical context for the tension between distributed multistakeholder governance and cyber sovereignty models of governance begins with the evolution of the Internet itself, arising in an American setting and, even as the Internet commercialized and expanded internationally, generally reflecting First Amendment values of the free flow of information and private market competition. All technological systems, to a certain extent, reflect the prevailing societal values in which they arise and embed historically specific power structures. Technologies continually evolve and are reciprocally shaped by markets and political dynamics. So too has the underlying technical architecture of the Internet constantly evolved in response to market changes and demographic shifts.

How the Internet is governed is not fixed any more than Internet technical architecture is fixed. Yet, the Internet Society, the institutional home for the Internet Engineering Task Force, has suggested that, even as the Internet has continuously evolved, it has arisen from a set of fundamental design and coordination principles it describes as Internet invariants, the core values that have guided the trajectory of the Internet. While always contested and marked by conflicting interests, these principles have included 1. Global reach/integrity; 2. General purpose; 3. Supporting innovation without requiring permission; 4. Accessibility; 5. Interoperability and mutual agreement; 6. Collaboration; 7. Reusable (technical) building blocks; and 8. No permanent favorites. These principles have resulted in an Internet that, while not universal, has moved toward universality; while not always enabling competition, access, and the free flow of information, providing the potential building blocks for this occur.

Concepts of interoperability, permissionless innovation, no permanent favorites, and global reach may seem self-evident to some in the contemporary context, but were a radical departure from the business models that preceded the Internet in which networks were based on proprietary protocols and designed specifically to only work with equipment made by a single manufacturer (e.g. IBM, DEC, Apple). Even the rise of online consumer systems in the early 1990s, including America Online, CompuServe, and Prodigy, were closed systems designed not to be compatible. For a time, an individual on one system could not send email to an individual on another system. Before the World Wide Web gave rise to the popularization and growth the Internet, there was not yet interoperability, permissionless innovation, or the potential for a general purpose network. There is nothing preordained about the retention of these design principles. Indeed, it is the contention of this paper that they are being challenged in ways that could change the nature of the Internet itself, compromise its security and stability, and diminish the global flow of information and the pace of technological innovation.

The governance of the Internet, like its underlying infrastructure, also has a historical context. The Internet does not run itself but requires a great deal of coordination to stay operational. At one point in the Internet's history, and in the history of its predecessor ARPANET, the majority of Internet users were American, Internet infrastructure was primarily within US borders, and the coordination of Internet infrastructure was done by Americans. For example, a single individual, Jon Postel working at Stanford Research Institute (SRI) in Menlo Park, California and funded by the U.S. Department of Defense, distributed and kept track of the network's names and numbers, a task over time privatized and internationalized and now overseen by ICANN, regional Internet registries, and domain name registrars.

The term "multistakeholder governance" arose, in part, in conjunction with the evolution of how Internet names and numbers have been overseen. Names refer to domain names, the globally unique alphanumeric names, such as https://sipa.columbia.edu, that identify websites and other virtual locations online. While these are the names that humans use to exchange information online, routers use associated globally unique Internet Protocol (IP) addresses, binary numbers assigned either permanently or temporarily as a globally unique identifier locating an online resource. Each exchange of information on the Internet requires these unique binary numbers, similar to the unique role of a postal address in the material world. This requirement for global uniqueness for each name and number has necessitated a centralized system of coordination for assigning, allocating and tracking all of these virtual identifiers.

Because of the historical condition of the Internet originating in the US with Department of Defense funding, the US government has had a longstanding role in coordinating these resources. As the Internet rapidly grew and internationalized in the 1990s, the US government commenced a process of privatization and internationalization of these coordinating functions, formalized by the incorporation of ICANN and a 1998 memorandum of understanding between the US Commerce Department and ICANN, a non-profit coordinating institution incorporated in the State of California. The agreement formalized ICANN's role in coordinating names and numbers, while still retaining accountability of these functions to the US government during the process of trying to internationalize and privatize this role.

The Commerce Department also arranged a contract with ICANN to handle the tasks known as the IANA functions.

Since the formation of ICANN, the role of the US Commerce Department in holding the contract with ICANN and authorizing root zone changes has been one of the most contentious debates in global Internet governance. International pressure to transition American oversight of names and numbers marked the World Summit on the Information Society in Geneva in 2003 and in Tunis in 2005. Even the formation of the United Nations-sponsored Internet Governance Forum, an annual conference to discuss Internet governance, was a compromise designed to, among other things, allow for the discussion of the possible transition of American oversight.

The disclosures by Edward Snowden of the expansive surveillance practices of the US National Security Agency (NSA) escalated concerns about the unique role of the US government in overseeing Internet names and numbers. Even though this oversight role has no discernable connection to NSA surveillance practices, and despite the knowledge that so many other governments carried out similar surveillance for law enforcement and national security reasons, the disclosures contributed to a loss of trust in US information policy that carried over to name and number administration and, in particular, oversight of changes to the root zone file. An already decade-long international concern about American hegemony in Internet governance became heightened, and international pressure to transition US oversight intensified and was perhaps best reflected in a global gathering in Brazil in May of 2014 called NetMundial to discuss global Internet governance.

In March of 2014, the National Telecommunications and Information Administration of the Commerce Department announced that the US would transition its oversight if the global multistakeholder community could meet a strict five-part test, including ensuring that the IANA functions could not be controlled by another government or intergovernmental organization. Although the original date for the possible transition was extended from September of 2015, it ultimately occurred on October 1, 2016. Key US government oversight functions have essentially been turned over to new structural arrangements within ICANN itself. ICANN, it can be argued, is an example of a multistakeholder governance organization because its processes and structures involve multiple actors – from private industry, government, and civil society. Those concerned about transitioning US oversight of names and numbers to the multistakeholder Internet governance community were not necessarily opposed to the multistakeholder model but rather concerned about a possible future government takeover of names and numbers oversight possibly by the United Nations (including its affiliated entity, the

International Telecommunication Union), or concerns about undue influence of countries like Russia and China with repressive information policies. The question of concern is how oversight will potentially evolve in the future.

The longstanding and high-profile tension over US oversight of Internet names and numbers has sometimes created the misperception that these functions comprise the totality of all Internet governance tasks. Governance of the Internet is not at all a single function, although it is sometimes discussed in this way, but rather an entire ecosystem of tasks necessary to keep the Internet's infrastructure operational and to enact public policy around this infrastructure. There are many taxonomies that explain the various layers of Internet governance tasks. The administration of domain names and Internet addresses is just one area, which itself disaggregates into numerous tasks including the distribution of IP addresses, the assignment of domain names, the authorization of changes to the root zone file mapping top-level domains and Internet addresses, the operation of the Internet's root servers, the task of resolving billions of Domain Name System queries a day to translate domain names into numbers, and the approval of top-level domains (TLDs). This non-exhaustive list of tasks around name and number administration serves to explain the number of coordinating responsibilities required in this one area alone. Other activities of Internet governance include the establishment of technical protocols by standards-setting institutions, access and interconnection coordination, cybersecurity governance, the policy-making role of private intermediaries, and intellectual property rights enforcement. Cybersecurity governance alone involves many heterogeneous tasks ranging from cybersecurity regulation and enforcement, software patch management, routing and DNS security, encryption design, and the role of trust intermediaries authenticating websites.

What becomes obvious in describing these few tasks of Internet governance is that the administration of the Internet is distributed over a variety of actors, including the private sector, new global institutions, and traditional governance structures. Some tasks, such as international treaties or the establishment of information laws around intellectual property rights are the purview of governments, some, such as private interconnection arrangements, are private-sector led, and still others, such as the administration of names and numbers, are performed by institutions like ICANN that involve a combination of actors from civil society, private industry, and government. Taken together, these collective tasks that have developed over decades and necessary to keep the Internet operational, are called distributed "multistakeholder governance."

In some ways, multistakeholder governance is also the privatization of governance, with many functions formerly handled by the state in the material world now overseen by private industry in the digital world. For example, large information intermediaries like Google and Facebook establish the conditions of privacy and speech via their user terms of service, essentially private contractual agreements for how personal data and metadata is handled, collected, and shared, when user accounts are terminated on speech grounds, how to handle cyberbullying and hate speech, and how and when to comply with government requests to take down content or turn over user account information for law enforcement or other reasons. This technical mediation and, in many ways, privatization of individual civil liberties, does not exist in a vacuum because information intermediaries are constrained by the laws of the countries in which they operate, and influenced by market forces and civil society pressure. The Internet's technical community has an influence over architecture-based policy; global institutions like WIPO and the United Nations help shape approaches; civil society exerts pressure on the private sector, such as the boycotts over the Stop Online Piracy Act (SOPA) in the US; and governments have the ability to pass laws, including Section 230 of the Communications Decency Act.

Technically and institutionally mediated Internet policy decisions about conditions of speech, privacy, decency, and government requests, is significantly complicated by the geopolitical reality that cultural norms, technological capabilities, and statutory contexts vary widely from country to country. Navigating context-specific constraints, particularly around content, is complicated. For example, German law requires information intermediaries to block access to Nazi content. Brazil has strong hate speech prohibitions. The European Union has strong consumer privacy protections. The United States imposes strong intellectual property rights enforcement requirements, such as the notice and take down provisions of the Digital Millennium Copyright Act (DMCA). Lese-majeste laws prohibiting insults against a monarch or member of a royal family are in effect, for example, in Thailand and Malaysia. Private companies make determinations every day about how to comply with, or not comply with, requests to block information or turn over user data.

The multistakeholder model arose in the American Internet context but now the vast majority of Internet users are not in the US or even in the West. Overall, there are more than three and a half billion global Internet users, and this will soon reach five billion with the majority of growth in emerging markets and countries like India with most users accessing the Internet on mobile phones. The Internet is growing rapidly across the world, but the majority of Internet users are in China, with the number of

Chinese Internet users far exceeding the entire population of the United States. Both demographic changes and technocultural heterogeneity have provided the backdrop for the Internet governance conflicts of the modern era.

Inflection Points in Cyber Sovereignty vs. Distributed Governance

It has long been established that the distributed architecture and governance of the Internet is subject to national statutory contexts, but this is complicated to implement in practice. That legal borders matter became crystalized in the 2000 French court case *LICRA v. Yahoo!* addressing the sale of Nazi memorabilia on the Internet. Yahoo! was sued because French citizens were able to purchase memorabilia via this platform. Even though the company's servers (at the time) were housed in the US and the company incorporated in the US, and despite the strong US constitutional protections of free expression, the ability to purchase Nazi memorabilia via Yahoo! was ruled to be illegal under French law and would require technical blocking. This more than a decade-old case helps to capture the challenges private companies face in navigating heterogeneous legal contexts and how jurisdictional questions are complicated by the transborder nature of technical architecture. It also provides an example of an Internet policy issue quite distinct from the question of control of Internet names and numbers.

Media and policy discussions around the transition of United States oversight of names and numbers nevertheless often portray governance of the Internet as a single functional area – the administration of the Domain Name System - and view control struggles over this system through the prism of traditional governmental institutions. For example, United States Senator Ted Cruz framed the IANA transition as President Obama surrendering the Internet to authoritarian regimes and instead advocated that the Commerce Department retain its oversight. Operationally, governance of the Internet is not a single task that can be relinquished but an entire constellation of distributed responsibilities necessary for keeping the Internet operational and for establishing relevant public policy. The following describes some of the more high-profile global debates over control of the Internet that collectively portray Internet governance as much broader than control of the Domain Name System and that illustrate the essential tension between governmental versus distributed control.

Interconnection Conflict

One conflict between multistakeholder versus multilateral approaches to Internet governance emerged around the 2012 International Telecommunication Union

(ITU) World Conference on International Telecommunications (WCIT) convened in Dubai. The objective of the meeting was to revisit an international interconnection treaty on telecommunication pricing known as the International Telecommunication Regulations (ITR), an intergovernmental treaty dating back to the late 1980s to address cross-border operation of telecommunication carriers. The conference was convened by the ITU, a specialized sub-agency of the United Nations to address telecommunication policy. Regulatory questions regarding interconnection have historically addressed issues of compensation and pricing related to how telecommunication companies interconnect and exchange traffic.

At issue in Dubai was the question of a possible expansion of the multilateral treaty to include Internet connectivity (rather than primarily voice telecommunication issues, although the two overlap considerably), a greater call for government intervention in facilitating interconnection among private companies, and also the prospect of extending the treaty to include content-specific issues such as regulating spam. The conference exposed a fundamental divide between countries wanting greater government control and oversight of the Internet's infrastructure and of content, and those countries opposing an expansion of the ITRs to include Internet infrastructure, opposing greater government control of content, and generally wanting to preserve multistakeholder, rather than multilateral approaches to interconnection governance arrangements. This fundamental divide was captured by divisiveness leading up to and during the conference, as well in the fractured vote on the proposed new telecommunications regulations, in which 55 out of 152 countries opposed the proposed treaty changes, including the United States, Japan, Canada, German, India, the United Kingdom, and others.

Data Localization Policies

One area in which governments are seeking to assert national sovereignty in cyberspace involves so-called data localization laws, often framed in terms of technological sovereignty or data sovereignty. Data localization policies involve a range of specific requirements and prohibitions on how and where private companies may handle and store customer data. For example, they often mandate that content intermediaries store data within the country in which the customer resides. In other cases, they impose restrictions on how data "crosses borders," require consumer consent, or require taxation on "data exports." These laws are in place in many countries – from Russia to Brazil - that span a range of ideological approaches to information policy. For example, Russia's law requiring companies to store data of Russian citizens within the nation's borders took effect in late 2015. Some countries, like China, have industry specific data laws in areas such as restrictions on storing financial services and healthcare data offshore or requiring that cloud computing services housing government data reside within China.

Some of these policies, citing privacy concerns for citizens, arose in the aftermath of disclosures about the expansiveness of the NSA surveillance program. In other cases, motivations appear to be providing market advantages to home-grown companies. Some civil society groups have advocated for legal restrictions on the transfer of personal data across borders to comply with privacy and personal data laws.

From an engineering and business model standpoint, and possibly even a civil liberties standpoint, these laws create new challenges. They not only apply to technology companies, such as social media platforms, but usually to other industries that store data, such as banks and retail companies, and these multinational companies can be forced to retool their computer networks to comply with requirements. Newer companies can be shut out of markets because they might not be able to locate their physical infrastructure and servers within every potential market in which they do business. The concentration of data can also make it easier, not harder, to protect data privacy by concentrating personal information. Data sovereignty does not match up with the multinational market approaches of many companies, nor with the distributed technical design that crosses borders and could potentially store data in several locations, house a customer service center in another nation, and have a corporate headquarters in still another country.

Cyber Sovereignty and Multilateral "Cyber Order"

The Internet is not yet a universal network. Much of the world does not yet have access and, where there is Internet penetration, users have different experiences based on language, technical expertise, access speeds, and different technical characteristics. But the rapid growth and innovation of the network creates conditions in which, in most of the world, the network can be thought of as having at least a technological affordance of universality. In other places, like Cuba, Iran, China, and North Korea, networks are walled off from the universal Internet using a variety of control mechanisms. In Iran and Cuba, for example, there are firewalls that strictly control interoperability with other networks and control the flow of content, both in and out of the country, as well as within borders. The Great Firewall of China is perhaps the best example of an efficient, nation-wide system of content restrictions and censorship and a prime example of an exertion of cyber sovereignty in which governments require private industry, as well as the institutions of Internet governance, such as Internet Exchange Points (IXPs), standards bodies, and registries, to carry out various forms of restrictions to content and infrastructure.

An emerging narrative closely related to national Internet sovereignty is the discourse of cyber order, in which countries are advocating for stronger multilateral, rather than multistakeholder approaches as a way of bringing about social order. In late 2015, for example, this multilateral approach was a significant theme at the World Internet Conference convened by China. As the Internet has become more economically and politically important in China, the government has had an increasing interest in global Internet governance and has advocated for greater multilateral oversight in which governments primarily oversee the coordination of Internet infrastructure and the policies around this infrastructure. This philosophy is clearly evident in the June, 2016 *Joint Statement between the Presidents of the People's Republic of China and the Russian Federation on Cooperation in Information Space Development* which declares support for creating a multilateral Internet governance system in which the United Nations plays a very important role. The changing narrative from multistakeholder to multilateral Internet governance is a sea change in how the technical community, civil society, industry, and policymakers in the West have both carried out and talked about the administration of the Internet.

Cybersecurity Versus National Security

Points of pressure toward national Internet sovereignty also exists in Western countries, especially around questions of government access to online content, personal information and metadata about citizens for purposes of national security and law enforcement. The need for strong cybersecurity, particularly encryption, so economically necessary for instantiating trust in the digital economy, often comes into conflict with law enforcement and national security requirements for accessing data to fight crime and carryout intelligence functions. After learning about the extent of NSA surveillance, for example, the Internet's multistakeholder technical community, particularly the Internet Engineering Task Force (IETF) called for "hardening the Internet" with greater end-to-end encryption that makes it more difficult, or more expensive to carry out pervasive surveillance. Similarly, private industry has made encryption the default in email and web access.

One government response to trends toward greater encryption, most visibly in the US, has been a discussion of building in "backdoors" into encryption protocols and systems so that law enforcement can readily access data. The issue emerged in the wake of a terrorist attack in San Bernardino, California when authorities sought access to an encrypted Apple smartphone belonging to the attacker and Apple CEO

Tim Cook resisting attempts to circumvent device encryption because it potentially builds in an inherent security vulnerability that could be exploited by foreign governments, hackers, and cybercriminals. In the context of massive data breaches, such as potentially 500 million users' data compromised in the Yahoo! hack, the idea of building in security vulnerabilities for national sovereignty is in direct tension with the multistakeholder technical community's efforts to increase rather than decrease the security of the Internet infrastructure that underlies all industrial, political, and social systems in the modern era.

Four Problems with Cyber Sovereignty Models

These examples of conflicts involving the rise of cyber sovereignty share several characteristics that help explain the problems these trends present for the future of the Internet, the pace of innovation, and human rights online.

1. Cyber Sovereignty Tampers with Internet Technical Infrastructure and Business Models

The Internet's core infrastructure can be taken for granted because of the network's ongoing growth and success. Indeed, policymakers and users alike are often not aware of the massive infrastructure of telecommunication transmission systems, switches, interconnection facilities, and cloud computing server facilities comprising the global Internet. In using the Internet, one only sees content, applications, and the end device (e.g. smartphone, laptop) used to access the content and applications. More than 99% of the Internet's infrastructure lies beneath what is visible at these end points.

Internet infrastructure has grown organically in the contemporary era, led by private sector investment and innovation. There has not been a centralized or hierarchical system that has shaped infrastructure which, despite challenges associated with cyber security breaches, anti-competitive forces, and movements toward proprietary enclosure, has been relatively stable and secure. Decisions about infrastructure arrangements, while reflecting public interest concerns, have been shaped in part by consideration of engineering efficiency, redundancy, and security.

Cyber sovereignty approaches seek to tamper with technical architecture, raising questions about how these alterations will affect the Internet itself. For example, data localization requirements impose politically motivated constraints on configurations of technical architecture. Under the mantel of cyber order, government censorship efforts sometimes involve local DNS redirection techniques that compromise universally consistent name and number resolution. Referring back to the Internet Society's articulation of the design values that have shaped the Internet, the core design values of interoperability, global reach, and permissionless innovation, in particular, are challenged. Cyber order approaches seek to limit global reach, place politically motivated limits on interoperability and interconnection, require layers of permissions for the introduction of new services, and place restrictions on the organic configuration of networks based primarily on market demands or engineering efficiency. If these top down infrastructure modification requirements had arisen decades ago, it is very likely that the world would not have experienced the rapid growth and innovations of the Internet that have provided so many opportunities for freedom of expression and economic growth.

Part of this infrastructure tampering arises from incommensurability between national borders and transnational technology. For example, data localization requirements do not match how networks are designed. Engineering efficiency and performance objectives are often predicated upon distributing data closer to end points and creating distributed redundancy rather than having more centralized repositories. The flow of information crosses borders. Servers are distributed around the world. For a private company, a domain name can be registered with a registry in one area, a call center located in another part of the world, and the company incorporated in yet another region. While national laws apply within borders, cyber sovereignty models are often incompatible with how technology works in practice.

2. Cyber Sovereignty Models Impede Civil Liberties

Tensions between cyber sovereignty and distributed governance are often flashpoints that mediate what counts as human rights in the online environment. For decades, considerations about human rights online have included freedom of expression, privacy, the right to assembly, the right to participate in cultural life, the right to access knowledge, the freedom to innovate, and a host of economic liberties such as the freedom to innovate and participate in technological and scientific advancement. Civil liberties in the digital sphere have long also included the protection from online harms such as cyberbullying, censorship, unwarranted invasive surveillance, identity theft, and theft of intellectual property, among others. In many public policy areas, governments are viewed as necessary for creating the statutory and market conditions necessary for the promotion of human rights. The United Nations Human Rights Council has asserted that the same rights citizens are entitled to offline are applicable online.

In the online sphere, the historical record suggests that government interventions in cyberspace, under the guise of cyber sovereignty, are increasingly in direct conflict with human rights. For example, governments with repressive information policies use the Internet to monitor activists and censor information. They use distributed denial of service attacks to disrupt alternative media sources. They engage in expansive surveillance and information gathering practices that challenge basic norms of individual privacy. In the case of the Egyptian Internet outage, for example, they sometimes cut off citizen access entirely. The Internet is clearly recognized as a significant site of power in which economic and political objectives can be carried out.

Indeed, some tension points between cyber sovereignty models and distributed governance are actually tension points over human rights online. Models of cyber sovereignty in China include systems of filtering and censorship. Cyber order in Russia includes repression of free speech for LBGTQ and other communities. Attempts to weaken Internet security for expansive surveillance practices in the West raise profound privacy questions. Even data localization laws designed to protect privacy, concentrate data in a way that could make it easier to compromise privacy and make it easy for foreign surveillance to be carried out. Questions about the protection of human rights online are becoming increasingly complicated.

4. Cyber Sovereignty Approaches Weaken the Stability and Security of the Internet

Cyber sovereignty tools that tamper with the core infrastructure of the Internet often carry negative externalities for cybersecurity. Attempts to weaken or place limits on encryption are obvious examples of government interventions that, intentionally or not, can compromise the Internet's stability and security. Attempts to modify non-security-related core infrastructure of the Internet can also have affects. Politically motivated modifications to the Internet's Domain Name System are an example, such as local DNS redirection techniques that require local institutions residing within national borders, such as ISPs or non-authoritative DNS operators, to ignore the universally consistent record mapping names and numbers and instead modify address resolution data locally. In other words, when an Internet user attempts to access a website or page being blocked, the local DNS server would redirect the request to another site, or the lookup would fail. Local redirection has been used to block entire social media applications, such as when Iran banned Twitter.

Billions of address resolution lookups happen daily and the stable functioning of this system requires universal consistency. Local redirection can create problems when it does not remain local but rather cascades globally. The most well-known example occurred in 2008 when the Pakistan government order Pakistani Telecom to block YouTube using local redirection but the routing information was extended up the DNS technical hierarchy resulting in a more global inability to access YouTube.

Local redirection, and also DNS injection techniques that cause DNS servers to lie about associated IP addresses, create security complications, such as impeding the implementation of DNSSEC, an important protocol designed to cryptographically authenticate domain name lookup processes so that users can be assured that a server returns the webpage requested rather than a counterfeit or malicious site. DNSSEC would be unable to distinguish between local redirection and cybercrime designed to carry out identity theft, disseminate malicious code, or sell counterfeit products.

This example is emblematic of the values tensions between the economic and public interest need for cybersecurity and government interest in content control, whether for intellectual property rights enforcement, censorship, or other objective. Attempts to exert bordered, top-down policies on technical infrastructure must understand and consider the complex technical substructures that preserve Internet stability and security.

5. Cyber Sovereignty Approaches Fragment the Internet

A significant theme in global Internet policy discussions is whether the Internet will continue to grow into a universal network or fragment into networks divided by national borders, proprietary ecosystems or other divisions. Bringing the next billion users (and objects) online requires the ongoing growth of the Internet. Ongoing growth in the digital economy requires interoperability among systems and the free flow of information across borders. This potential for universality and interoperability has been a taken for granted assumption for the enjoyment of expressive rights and for economic development. The United Nations Human Rights Council statement on The Promotion, Protection, and Enjoyment of Human Rights on the Internet describes the global and open characteristics of the Internet as a driving force of development.

It is important to acknowledge that the Internet is not yet a universal network. There are barriers to access, digital knowledge divides, language divides, interoperability challenges among protocols, and other kinds of technical fragmentation. But the Internet has continued to become more universal. Cyber sovereignty models are attempts to overlay geographical borders on the cross border Internet, raising questions about the effects of these policy approaches on the question of whether the Internet will become more fragmented or move toward greater universality. Data localization laws, in particular, impose these national borders around the transnational Internet.

The universal nature of the Internet has brought about considerable economic growth and development and new opportunities for access to knowledge and expression. Imposing new techniques to assign national borders on this infrastructure may have significant implications for access to knowledge, business autonomy, and the digital economy.

Beyond cyber sovereignty approaches, there are also private industry efforts to limit Internet universality and interoperability through protocol fragmentation and the development of proprietary ecosystems designed to limit competition and interoperability for market advantage. Indeed, many of the problems caused by cyber sovereignty approaches would be similar with "industry sovereign approaches." It is exactly the relative balance of powers among various stakeholders in distributed Internet governance models that creates conditions for universality and innovation.

A Sea Change in Internet Governance

Cyber policy is a critical domain of foreign policy. The balance of control over the Internet is now inextricably linked to the state of the digital economy, critical infrastructure protection, human rights, and even job markets. The rising importance of cyber governance is increasingly recognized by governments as they acknowledge the Internet as a site of conflicts around political and economic power. From a foreign policy perspective, including in the U.S., Internet policy inherently contains tremendously conflicting interests, such as cybersecurity and the promotion of democracy and Internet freedom, on one hand, and intelligence and cyber warfare concerns, on the other.

The great myth of Internet governance in the past decade has been that it is a single system over which control can be wrested. The term Internet governance sounds singular. The media and policymaker overemphasis on the IANA transition and the stakes of the transition is an example of monolithic approach to control of the Internet. This sense of a cohesive, uniform Internet governance framework may have had some truth decades ago, but in the contemporary era, as the Internet becomes more globalized and heterogeneous and becomes so fundamentally important to the world, Internet governance has morphed into a large number of interrelated areas. Intersections between national security and critical infrastructure protection are one set of issues that have a certain set of players and its own set of different decision making processes. Concerns related to the free flow of information on the Internet, often framed as “Internet freedom” issues, embody a different set of foreign policy and public interest concerns. The Domain Name System and its global institutional governance structure embodies yet a different set of concerns. Politically, one of the contextual backdrops that created so much anxiety around the IANA transition is this discursive aggregation of more than a hundred Internet governance functions into a single control point. There is not one stop shopping for cyber issues and unpacking the various issues and their unique public interest and technical considerations most precede policy discourses. There are policy areas in which national sovereignty concerns, especially national security, trump other considerations, but in most areas, distributed multistakeholder governance approaches have contributed to the ongoing success of the coordinating functions necessary to keep the Internet operational.

The transition of American power over the DNS is the beginning, not the end of geopolitical conflict. The IANA transition resolves nothing about the broader geopolitical conflicts involving the essential tension between multilateral or sovereign control and the distributed multistakeholder approach to administering the Internet. There is a sea change in the philosophy of Internet governance around this tension. This paper has described a small subset of global inflection points that reflect this transformation. The same tension will play out on different game boards as technological, cultural and market forces continue to shape the Internet.

Internet governance questions now exist in a post-Washington-consensus world. Because of its historical origins in the US and the West, the Internet came of age in a certain political and cultural context. Liberalization, privatization, and globalization were the hallmarks of this Washington consensus world. There has been a sharp turn away from this, including increased interest in regulation, distrust of globalization, and movement away from privatization. In the same way a certain set of democratic and free market values shaped the constitution of the Internet in its opening decades, where will this emerging context have the greatest implications for the future of Internet architecture and governance?

The policy battles of the future will be around the co-option of infrastructure rather than control of content. Already, governments recognize infrastructure control points as points of power, whether seeking modifications to the Domain Name System, approaches to data storage, or interconnection agreements. This turn to infrastructure will only increase. The most complex and overwhelming of infrastructure issues, and one with profound implications for security, privacy and economic competition, exists around the Internet of Things. Already, Distribute

Denial of Service Attacks have exploited security vulnerabilities in IoT devices like baby monitors and surveillance cameras to carry out extensive and disruptive attacks on specific targets and even on the DNS itself. The Internet constantly changes and policy changes today have to anticipate the diffusion of the Internet into material objects of our social and economic systems. It would be a difficult process to create international norms and enforceable agreements to not attack the core infrastructure of cyber physical systems or to agree on common privacy laws around the Internet of Things. Getting ahead of future cybersecurity problems around the Internet of Things will require strong security and accountability, a public policy challenge because there may not be adequate incentives for individual users or markets to address security.

In the context of cyber physical systems, another major policy battleground likely to reflect tensions between distributed multistakeholder governance and cyber sovereignty will involve government surveillance. There will also be infrastructure battles over technical standards and the open question of whether cyber physical systems will rely on the universal Internet address space or seek alternative address spaces controlled by governments rather than by the global multistakeholder communities. The question of involvement of the ITU in technical architecture and governance questions around IoT standards and names may be as contentious a debate as the decades-long battle over control of the IP address space and associated standards. Will the Internet remain a universal system based on the common IP address space or transform to other name and number systems, and therefore new governance structures? There is nothing fixed about Internet governance arrangements in the same way there is nothing fixed about Internet architecture. The process of Internet governance has been contentious but fairly stable – however there is nothing preordained about this. Geopolitical tensions will only rise as states increasingly view points of control over Internet infrastructure as sites of power, even while every aspect of political, social and economic life increasingly depend upon this infrastructure to function. How these tensions are resolved and the values inherent in the way Internet governance operates will likely determine whether the Internet continues to be an engine for material economic, social, political and cultural change, or if the Internet begins to splinter and become just another technology that did not meet the hype and its promise to change the world permanently for the better.