Understanding Cyber Collateral Damage

Sasha Romanosky* & Zachary Goldman**

Introduction

In conventional (kinetic) U.S. warfare, there exists a standard methodology for identifying and assessing collateral damage (i.e. accidental damage to civilian targets). Indeed, the U.S. Department of Defense (DoD) relies on a governing document that defines the policy regarding unlawful military targets (no-strike targets), and methods for estimating collateral damage from kinetic military operations.[1] The definitions in this document are clear, and the harms against which it aims to protect are tangible because they relate to persons and property. The munitions in the military's arsenal are defined and well-known, and their properties—blast radius, amount of force delivered, and the like—are well understood. While accidents of course do occur, the anticipated effects of a kinetic operation (collateral or otherwise), are generally straightforward to anticipate, assess, and manage.

However, given the interconnectedness of cyber and cyber-physical systems, direct, indirect, and collateral effects can be much more difficult to predict, rendering ineffective traditional approaches to collateral damage estimation (CDE). Indeed, even the notion of clearly defining and considering "damage" within the cyber realm is challenging. For example, how does one estimate harms resulting from an outage of network connectivity caused when an attacker exploits a software vulnerability? How can one evaluate and weigh the collateral impact of a cyber intervention on incommensurable values, such as exposing the IP addresses of anonymous Tor users in order to arrest child pornographers, against international comity concerns that might be implicated by remotely searching foreign computers in contravention of traditional diplomatic and law enforcement norms?

We consider two main questions in this Article. First, how can traditional military doctrine be adapted to accommodate the unique challenges of estimating collateral damage in the cyber domain? And second, how can domestic U.S.

---

* Associate, RAND Corporation.

** Executive Director, Center on Law and Security, New York University School of Law.

Acknowledgements: We would like to thank Lily Ablon, David Aitel, Krista Auchenbach, Charles Brown, Bob Elder, Allan Friedman, Martin Libicki, Eric Jensen, Mark Sparkman, David Senty, Michael Warner, and Sean Watts for their valuable comments and insights. We would especially like to thank Cynthia Dion-Schwarz for her inspiration, and participants of the Legal and Policy Dimensions of Cybersecurity workshop at George Washington University School of Media and Public Affairs (Sept 28-29), 2016.

1. Chairman of The Joint Chiefs of Staff, No-strike and the Collateral Damage Estimation Methodology, Department of Defense, CJCSI 3160.01 (2009). Note, the version referred to within this document, obtained via a freedom of information act request by the ACLU, is unclassified and no longer for official use only (FOUO).

law enforcement agencies develop a similar conceptual framework for anticipating and evaluating collateral damage?

The purpose of this Article is not to reproduce existing literature regarding cyber war, military doctrine, or international laws of war, nor do we attempt to mathematically or empirically model computer dependencies. Indeed, we draw on these (and other) resources in order to understand how damage, and therefore collateral damage, may occur from cyber and kinetic operations in a range of contexts.

The fundamental question is whether unintended effects on data alone can constitute collateral damage requiring operational planners in the military and law enforcement context to weigh that inadvertent harm against lawful objectives during the mission planning and execution process. We answer that question in the affirmative, while recognizing that the precise contours of what constitutes collateral damage in cyberspace, relative to traditional canons, remain to be defined. That task will remain difficult while the vast majority of cyber operations remain secret and states remain unwilling to speak publicly about the process for planning and executing them. But as a greater number of such operations see the light of day and governments become less reluctant to divulge information, over time a more robust standard can be developed. For now, the main task is to identify the conceptual issues with which such a framework must grapple.

This Article will first define key terms for evaluating cyber collateral damage. We will then describe the analytical process for evaluating collateral damage in the kinetic context. Finally, we will present a framework for evaluating collateral damage relevant to cyber operations and show how that framework can apply to both law enforcement and military cyber operations.

I. DEFINITIONS AND BACKGROUND

For the key terms below, the definitions are drawn from the military context (as that is where the most mature framework resides). However, they are relevant in non-military situations as well and will therefore be used throughout this Article.

A. Cyberspace Operations

While formal definitions of "cyber" and "cyber operations" (or, "cyberspace operations") are evolving, for the purpose of this Article, we consider cyber operations to include the "(1) use [of] cyber capabilities, such as computers, software tools, or networks: [that] (2) have a primary purpose of achieving objectives or effects in or through cyberspace."[2] More specifically, U.S. military cyber operations consist of three types: offensive cyber operations (OCO),

2. OFFICE OF THE GEN. COUNSEL, U.S. DEP'T OF DEF., DEPARTMENT OF DEFENSE LAW OF WAR MANUAL § 16.1.2 (2015).

defensive cyber operations (DCO), and DODIN operations.[3] OCO refers to cyberspace activities intended to project power dodin (i.e. cause an effect) "in and through cyberspace."[4] DCO are defensive cyber activities taken in response to an adversary's actions (such as an attack, or imminent threat), while DODIN operations are those typically known as cyber security efforts that protect one's computer network and information from compromise.[5]

In addition to these activities, Joint Publication 3-12 defines three "layers" of cyberspace operations: physical network, logical network, and persona.[6] The physical network layer refers to the geographic location of the computers, servers, networking equipment, cables and wiring, and includes the hardware and software components.[7] The logical layer is a higher level of abstraction and refers to the application layer of internet communication, consisting of, for example, a website, database, email application, etc.[8] Each of these applications may serve, store and process data that physically resides in multiple locations simultaneously (striped or mirrored across many storage devices or networks).[9] Finally, the persona layer represents the digital identity of an individual or entity, such as a social media user account.[10] Further, as described in Joint Publication 3-12, there may be a one-to-one, many-to-one, or one-to-many relationship between an actual individual (or individuals) and a digital persona (or personas), which may include many components of the physical and logical network layers.[11]

B. Collateral Damage

The U.S. Department of Defense (DoD) defines collateral damage as the "unintentional or incidental injury or damage to persons or objects that would *not* be lawful military targets in the circumstances ruling at the time."[12]

---

3. CHAIRMAN OF THE JOINT CHIEFS OF STAFF, JOINT PUBLICATION 3-12 at vi (2013).

4. *Id.*

5. *Id.* Note that the terms computer network defense (CND), computer network attack (CNA) or computer network exploitation (CNE) are still employed in some contexts, though are deprecated. UNITED STATES ARMY, UNITED STATES ARMY TRAINING AND DOCTRINE COMMAND 19 (2010). In that context, CND refers to actions taken "to protect, monitor, analyze, detect and respond to unauthorized activity" within a computer network. CNA refers to actions taken "through the use of computer networks to disrupt, deny, degrade, or destroy information resident in computers and computer networks, or the computers and networks themselves," and CNE refers to "enabling operations and intelligence collection capabilities conducted through the use of computers." *See* U.S. DEP'T OF DEF., DOD DICTIONARY OF MILITARY AND ASSOCIATED TERMS 277 (2017).

6. JOINT PUBLICATION 3-12, *supra* note 3, at I-2.

7. *Id.*

8. *Id.* at I-3.

9. *Id.*

10. *Id.* at I-4.

11. *Id.*

12. CHAIRMAN OF THE JOINT CHIEFS OF STAFF, JOINT PUBLICATION 3-60 at GL-6 (2007) (emphasis added).

Similarly, the Program on Humanitarian Policy and Conflict Research at Harvard University defines collateral damage as "incidental loss of civilian life, injury to civilians and damage to civilian objects or other protected objects or a combination thereof, caused by an attack on a lawful target."[13] Essentially, these definitions amount to accidental harm to non-military targets, and they are narrow in their description of both harm (considering only physical or property damage), and the object of any potential harm—objects or persons that would not be lawful to target in the first instance. For example, consider a bomb that destroys a military facility, but which also damages an adjacent military command center and a civilian school. In that instance only damage to the school would be considered collateral damage. Ancillary damage to the command center is simply a side effect that is favorable to the attacker.

These definitions also suggest that accidental harm suffered by friendly forces (or to the attackers themselves) would not be considered collateral damage.[14] Other outcomes not included in the definition of "collateral damage" are harms suffered by the attacker as a result of any retaliation in any form such as diplomatic, informational, military, or economic (sometimes referred to as "DIME"). Note that in *intentional* attacks on civilian facilities or people, the attack and any subsequent harms would not be considered collateral damage, but would instead constitute a violation of the laws of war.[15]

C. Damage and Harm

A formal definition of harm (or damage) is necessary for a discussion of cyber collateral damage because absent any harm (cyber, or otherwise), there would be no collateral damage to evaluate. In attacks using conventional weapons, the damage caused by such weapons is often straightforward, though perhaps not easy, to estimate. Indeed, military doctrine describes a specific process for estimating the physical damage caused to property due to a kinetic weapon.[16] Moreover, there is no conceptual ambiguity about what constitutes "damage."

Conversely with cyber operations, there are both conceptual and practical challenges involved in evaluating damage. At the conceptual level, there is a lack of agreement about what constitutes "harm," specifically surrounding the question of whether mere breaches of the confidentiality, integrity or availability

13. PROGRAM ON HUMANITARIAN POLICY AND CONFLICT RESEARCH AT HARVARD UNIVERSITY, MANUAL ON INTERNATIONAL LAW APPLICABLE TO AIR AND MISSILE WARFARE 3 (2009).

14. NRC describes some of these effects as "blowback," for example, when an attack on an enemy also directly (though inadvertently) causes harm to U.S. firms. In addition, this considers situations where a successful attack on an enemy State's network infrastructure would also prevent a U.S. firm from doing business with another State that depends on supplies from the target State.

15. *See* DoD LAW OF WAR MANUAL, *supra* note 2 at § 5.3

16. *See* CJCSI 3160.01, *supra* note 1, at § D.

of data without any physical effects should constitute "damage" for the purposes of collateral damage estimations. At a practical level, the outcomes of cyber operations can be much more uncertain than physical operations, and evidence of whether any damage has occurred at all may be unavailable. This is so for a number of reasons. Harms that originate in code may be latent or transient. They may rely on the confluence of a number of different events to achieve their peak damage. Failures in technical systems may emerge for reasons that have nothing to do with code that is deliberately introduced. And victim States may have an incentive to keep secret the harm they have suffered so as not to project an image of vulnerability to the broader community.

Indeed, a critical observation, and one main purpose of this Article, is to demonstrate the ways in which evolving notions of harm in the cyber domain lack a comfortable place in the traditional context of collateral damage. For instance, consider a software vulnerability exploited by an adversary. The vulnerability is used to install a software program that causes the adversary's power station to overload and be physically destroyed. In this case, the method of committing the attack (i.e. using computer software to destroy the power station) should be irrelevant for the discussion of damage assessment. Whether caused by a conventional bomb, or cyber attack, the physical effects from this example are similar, as would be the assessment of any collateral damage. It is only when the outcomes are contained to computing systems that traditional procedures break down.

However, two further scenarios illustrate the difficulties involved. In the first, consider code that is deliberately introduced into a system causing it to cease operating (but producing no physical damage). In the course of this attack, there are transient effects on (but again, no physical damage to) another system that would not be a lawful target for a cyber operation. Has there been collateral damage? Consider again the same vulnerability that is exploited, but instead of overloading a power plant's operations, malware is placed on a computer which could—but has not yet—affected the power plant's operations. Has any harm occurred? If so, what is it, and what would be the appropriate boundary of the response if malware is implanted but not activated? One can imagine many other possible scenarios, such as malware that is installed simply to observe network traffic on an adversary's computer network that also unintentionally collects information on computers that are not lawful targets of the operations. What harm or damage has been caused by this form of surveillance? Or even more directly, consider that same software unintentionally deletes a corporate or governmental database (again, without any physical effects). Has any "damage" actually occurred, even if large economic losses result? The answers to these questions lie at the seams between U.S. legal doctrine and the law of armed conflict, both of which are struggling to keep pace with technology and the capabilities afforded by information technology.

Consider Figure 1 which characterizes an operation along two dimensions: the cause (means) of the operation, and the effect.

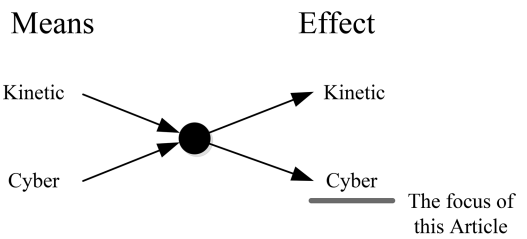

Figure 1
Cause and effect.

As shown, we consider that many operations can be conducted either with kinetic or cyber means (e.g. physically destroying a server). And similarly, they can cause either (or both) kinetic or cyber effects. However, for the purpose of this Article, it is the *consequence* of an operation that we focus on, rather than the *means* by which it occurs. That is, we are concerned with collateral cyber (not kinetic) effects, whether caused by kinetic or cyber means.

D. Process for Evaluating Military (Collateral) Damage as a Result of Operations

Next, we draw on multiple sources of U.S. legal and military doctrines, as well as other academic work, to consider the conditions under which a military operation (traditionally, a kinetic military operation) could cause collateral damage. We then leverage this analysis to consider when a *cyber* operation could produce collateral damage. Most substantively, we draw on the work of the Tallinn Manual, a document which reflects the combined effort of dozens of international legal scholars and former practitioners since 2009 to consider how the law of armed conflict applies to cyber operations.[17]

For the purpose of establishing a baseline understanding, we first consider that a cyber operation conducted by the United States must be grounded in a source of domestic legal authority.[18] That operation will then include three components: identifying a lawful target, invoking a lawful participant (the individual who carries out the operation), and, particularly in the context of cyber operations, differentiating between a military activity and intelligence collection. The components and their relationships are illustrated in Figure 2.

17. *See Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge University Press ed., 2013). Note that the Tallinn Manual narrowly focused on applying the law of armed conflict to cyberspace. However, this does not distract from the purpose of this exercise. We also fully recognize that the Tallinn Manual is not a treaty or binding interpretive text.

18. This distinction is critical because the particular authority by which an operation is conducted incorporates many critical factors beyond the scope of this Article, not the least of which concerns any wartime protections that the individual would enjoy if caught by an adversary.

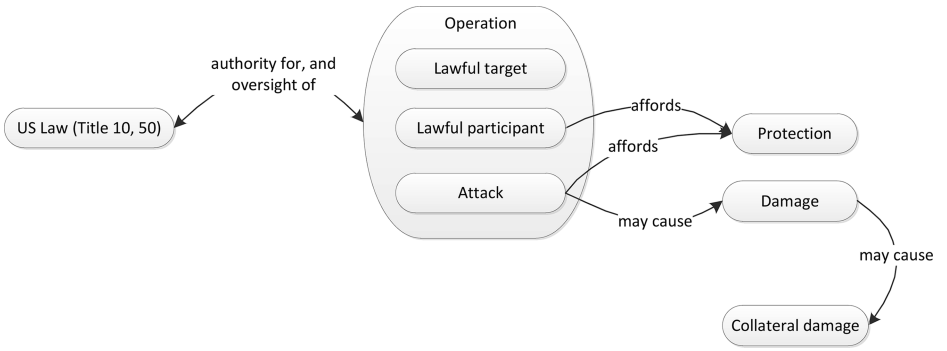

Figure 2
Collateral Damage Relationship diagram.

Authority: The first step concerns the rules of U.S. law that authorize military operations. For our purpose, there are two key provisions. First, Title 10 of the U.S. Code governs the functions and responsibilities of the U.S. armed forces, which grants authority for traditional military operations, including offensive operations.[19] Oversight for these operations is provided primarily by the House and Senate Armed Services Committees and internal executive branch processes. Second, U.S. Code Title 50 governs some U.S. conduct during times of war, and grants authority for intelligence activities, including CNE (espionage) operations.[20] Oversight for these operations is conducted by Congress, the executive branch, and the courts, and may require formal presidential findings. The distinction between military and intelligence operations has blurred for certain types of kinetic special operations and for certain types of cyber operations.[21]

In the specific case of cyber operations, the distinction is ambiguous because both military operations and intelligence gathering operations require the same initial steps—gaining access to an adversary's system, identifying the system's functions and the relationship among its parts, and keeping that access as persistent and stealthy as possible. This overlap raises questions about the legal authorities under which an operation is taking place (which has implications for the oversight to which it is subject),[22] and also can produce significant strategic effects.[23] Some authors suggest that because it is difficult for a victim to

19. *See* 10 U.S.C. § 101 et seq.

20. *See* 50 U.S.C. § 1 et seq.

21. *See e.g.* Andru E. Wall, *Demystifying the Title 10-Title 50 Debate: Distinguishing Military Operations, Intelligence Activities & Covert Action,* 3 HARV. NAT'L SEC. J. 85, 129 (2011) (concerning the operation to kill Osama Bin Laden).

22. *See* Gary D. Brown, *Spying & Fighting in Cyberspace: What is Which?*, 8 J. NAT'L SECURITY L. & POL'Y 621 (2016) (discussing at length the legal and strategic issues caused by overlapping military and intelligence authorities in cyber operations).

23. *See generally* BEN BUCHANAN, THE CYBERSECURITY DILEMMA: HACKING, TRUST, AND FEAR BETWEEN NATIONS (2017) (Analyzing why cyber operations that are intended to be only intelligence collection operations might be misinterpreted by the target and might cause inadvertent escalation).

determine whether a particular network intrusion is meant to be destructive or rather simply to collect intelligence, such intrusions may inadvertently exacerbate tensions between nations and escalate hostilities.[24]

Test of lawful target: The lawfulness of a proposed operation is central to collateral damage assessments while lawfulness, in turn, is determined by a wide range of factors governed by International Humanitarian Law (IHL), also known as the Law of Armed Conflict (LOAC).[25] Three broad categories of analysis are relevant to the lawfulness of a cyberattack. The first is whether LOAC applies—that is, whether there is an armed conflict or set of hostilities sufficient to trigger its applicability.[26] A second broad set of issues involves the question of who is a lawful target under IHL. This "distinction between combatants and non-combatants constitutes one of the two cardinal principles" of the Law of Armed Conflict.[27] Parties are targetable if they are, for example, members of an adversary's armed forces, but non-combatants can also be targeted if and for such time as they are directly participating in hostilities.[28] Both sets of issues are challenging on their own. They are also challenging to apply to the context of cyber operations, as there is at present little state practice and *opinio juris* to answer the question, for example, of when an independent hacker is taking direct part in cyber hostilities and therefore targetable under the Laws of Armed Conflict. Detailed discussions of both issues are, however, outside the scope of this paper.

What is most directly relevant is the third broad set of concerns around which LOAC revolves—namely, how a state can engage in hostilities once it has determined that it may lawfully use force against particular targets. If the second question—who is a lawful target—concerns itself with the fundamental LOAC principle of distinction, this third question concerns itself with the other fundamental LOAC principle of proportionality. The law of armed conflict does not demand that strikes take place with zero collateral damage.[29] Rather the Additional Protocols to the Geneva Convention bar "an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation

24. *See Id.*

25. *See generally*, GARY D. SOLIS, THE LAW OF ARMED CONFLICT: INTERNATIONAL HUMANITARIAN LAW IN WAR (2010).

26. We leave aside for the purposes of this article an extended discussion of distinctions between international armed conflicts and non-international armed conflicts, and complicated questions about when a non-international armed conflict arises. We also do not discuss the relationship between IHL and International Human Rights Law. For an extended discussion, *see, e.g.,* KENNETH WATKIN, FIGHTING AT THE LEGAL BOUNDARIES: CONTROLLING THE USE OF FORCE IN CONTEMPORARY CONFLICT (2016).

27. YORAM DINSTEIN, THE CONDUCT OF HOSTILITIES UNDER THE LAW OF INTERNATIONAL ARMED CONFLICT 33 (2010).

28. *See* INTERNATIONAL COMMITTEE FOR THE RED CROSS, INTERPRETIVE GUIDANCE ON THE NOTION OF DIRECT PARTICIPATION IN HOSTILITIES (2009).

29. Greg McNeal, *Targeted Killing and Accountability*, 102 GEO. L. J. 681, 749–50 (2014).

to the concrete and direct military advantage anticipated."[30] The rule is necessarily ill-defined and "no objective standards exist as to where this 'turning point' lies;"[31] determinations are by necessity fact-bound.

Applying these concepts to the context of cyber operations poses particular challenges. First—and a central challenge for this paper—is, what is considered "harm" or "damage" for the purposes of determining collateral damage, a topic we take up at length below. A second challenge involves the practical difficulty of actually estimating how much damage may result from a contemplated cyber operation. Determining with confidence what systems are connected requires an intimate understanding of the networks one intends to attack—a substantial intelligence undertaking. And a final challenge is how to weigh the anticipated military advantage against the expected collateral harms—a difficult task given the indeterminacy in each calculation.

Test of lawful participant: Next, we consider the tests necessary to determine whether an individual[32] who engages in an operation (cyber or otherwise) is acting lawfully under the laws of armed conflict. While this might be a concern in certain contexts, in most of the cyber operations described in this article the operation will clearly have been conducted by official government personnel acting in their capacity as state actors.

Cyber operations complicate traditional understandings of lawful participants because, in a manner unique to cyber operations, the antecedent steps for CNE (computer network exploitation) and CNA (computer network attack)—namely developing malware or other capabilities and deploying that malware to exploit specific systems or networks– are the same, and may be conducted by government or non-government individuals.

Test of an attack: Next, we examine when an operation may be considered an "attack." The two matters of key interest are the attacker's behavior and the consequences of the attack (Tallinn (2013), rule 30.7). Note that for the purpose of this test, the method of attack is not relevant. Article 49 of the Geneva Conventions (Geneva, 1949) defines an attack as an "[act] of violence against the adversary, whether in offense or in defense" (Protocol 1, 1977) and is considered accepted international law. This alone suggests that an operation would only be considered an "attack" if it inflicted "harm."[33] Indeed, drawing on this understanding, leading scholars of military strategy have contended that most of what is described as cyber war is not in fact "war," because "war" in the

30. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I) art. 51, June 8, 1977, 1125 U.N.T.S. 3. For non-international armed conflicts, see Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), arts. 1.2, 13.3, June 8, 1977, 1125 U.N.T.S. 609.

31. NILS MELZER, TARGETED KILLING IN INTERNATIONAL LAW 361 (2009).

32. Note that we consider only humans and not robots or algorithms in this Article.

33. For a list of exceptions, *see Tallinn Manual*, *supra* note 17, at Rules 30.2, 30.14–15, 30.18, 32.5, 36.

traditional Clausewitzian understanding must be violent, instrumental, and political in nature.[34] Instead, cyberattacks are best understood as one of a combination of espionage, sabotage, or subversion.[35]

Test of harm: By all accounts of military actions, the most widely accepted test for whether harm has occurred is whether an operation causes injury or death to persons, or damage or destruction to objects.[36] Therefore, harm would not include: inconvenience, irritation, stress, or fear, because they do not amount to "loss of life," "injury," or "damage."[37] For example, disabling civilian internet access—whether by kinetic or cyber means—may be considered a nuisance, but would not rise to the level of harm. In this regard, there is a threshold of inconvenience versus damage, torture or terror.

This final test is perhaps the most difficult to transpose directly from operations causing kinetic effects, to those causing cyber effects. Based on a traditional analysis of collateral damage (i.e., under relatively stable understandings of international law), collateral damage would occur only when physical "damage" has occurred, which, in turn, can only result from an "attack" (i.e., a hostile kinetic action). And only when *accidental damage to civilian property or persons* has occurred, can collateral damage result.

This observation is startling, because it implies only a very narrow set of conditions which could possibly lead to a formal recognition of cyber collateral damage. However, this finding is unsatisfying for several reasons, and leads to the rather surprising conclusion that collateral damage to data—no matter how significant—cannot be recognized as "real" collateral damage as long as there are no physical effects. This means that if a country's entire property ownership record, for example, or a stock market's daily trading ledger, was inadvertently deleted or (worse) manipulated, there would be no recourse under the traditional mechanisms used to guard against violations of the Law of Armed Conflict. This would, it is fair to say, be inconsistent with the expectations of reasonable people in a digital era. Given global interdependencies in telecommunications infrastructure it might also leave third parties without recourse, in the event that a conflict between two nations involving cyber capabilities causes collateral effects. But there are also a number of other reasons why collateral harm that affects data without generating physical effects could be considered collateral damage.

The first and most important reason to include collateral effects on data when calculating collateral damage is that doing so reflects the empirical and normative importance of data in today's world. Indeed, few doubt the catastrophic social impact that would derive from harm to certain kinds of data, even if that harm does not also result in physical effects. If, for example, a digital attack was

34. Thomas Rid, *Cyber War Will Not Take Place*, 35 J. STRATEGIC STUD. 5, 10 (Feb. 2012).
35. *Id.* at 15.
36. *Tallinn Manual*, *supra* note 17, at Rules 10, 12, 30.1, 30.4, 30.6, 35.4, 38.5, 38.6.
37. *Id.* at 133.

to inadvertently destroy or manipulate a nation's health records, its land ownership records, or its banking and securities records, nobody doubts that the social, political, and economic effects would be profound.

Prevailing interpretations of international law and the Law of Armed Conflict regarding cyber operations recognize that such effects would constitute a prohibited use of force or an armed attack justifying the right to self-defense. Thus, the Tallinn Manual begins by defining a use of force with respect to cyber operations in terms of its "scale and effects."[38] But while acts that "injure or kill persons or damage or destroy objects are unambiguously uses of force,"[39] operations that do not do so may also be considered a prohibited "use of force." According to the Tallinn Manual, determining whether a cyberattack constitutes a use of force depends on the attack's severity, immediacy of effect, directness, invasiveness, measurability of effects, the military character of the attack, state involvement, and presumptive legality.[40]

Furthermore, cyber operations of sufficient scale and effects may be considered armed attacks triggering the lawful right to self-defense.[41] Again, "any use of force that injures or kills persons or damages or destroys property would satisfy the scale and effects requirement."[42] On the other hand, "acts of cyber intelligence gathering and cyber theft, as well as cyber operations that involve brief or periodic interruption of non-essential cyber services, do not qualify as armed attacks."[43] The international experts who drafted the Tallinn Manual were divided about the intermediate cases, however. For instance, they split on the issue of whether an attack on a financial market that caused the crash of a major international stock exchange would constitute an armed attack.[44]

These views echo those of leading scholars. What's more, given the absence of treaties or state practice that clearly define prohibited uses of force and armed attacks, their opinions carry especially significant weight. Like the Tallinn Manual, these scholars accept the idea that—in principle—attacks on, for example, financial systems could cause sufficient enough damage to rise to the level of armed attacks.[45] While the number of attacks that would fall into that category at present might be small, as greater portions of society become "connected," the social impact of harm to data alone will grow. It is easy to

38. *Id.* at Rule 11.
39. *Id.* at Rule 11.8.
40. *Id.* at Rule 11.8–9.
41. *Id.* at Rule 13.2.
42. *Id.* at Rule 13.6.
43. *Id.*
44. *Id.* at Rule 13.9.
45. *See, e.g.*, Oona A. Hathaway et al., *The Law of Cyber-Attack,* 100 CALIF. L. REV. 817, 848 (2012); COMM. ON OFFENSIVE INFO. WARFARE, NAT'L RESEARCH COUNCIL, TECHNOLOGY, POLICY, LAW, AND ETHICS REGARDING U.S. ACQUISITION AND USE OF CYBERATTACK CAPABILITIES 221 (2009) ("cyberattacks on the controlling information technology for a nation's infrastructure that had a significant impact on the functioning of that infrastructure (whether or not it caused immediate large-scale death or destruction of property) would be an armed attack for Article 51 purposes").

foresee a point at which the scale and effects of cyberattacks against data are so profound that they constitute uses of force or armed attacks, notwithstanding the lack of physical damage.

If a cyberattack that damages data therefore could constitute an armed attack, it follows that collateral damage that "merely" harms data should also be considered in evaluating the collateral damage anticipated from cyber operations. A related point is that, by ensuring that collateral damage calculations take into account the harm to data, the operation planning process will be likely to take it into account as well. As we saw above, collateral damage calculations are integral to the proportionality evaluations that are a part of the law of armed conflict. Given the potentially catastrophic consequences that may derive from harm to data, this is an important development.

Integrating damage to data into an assessment of collateral damage will also add clarity and discipline to the process. Over time it will ensure that the government develops a rigorous methodology for making these determinations and deep expertise in doing so. It will also establish a set of institutional processes from across the government that can be brought to bear on these decisions.

If the collateral damage estimation methodology includes cyber operations that affect data the government will have to make difficult determinations about thresholds. But this is no different than the traditional collateral damage estimation process, which must first determine how much collateral damage will take place and then weigh that against the anticipated military advantage.

II. WHAT IS CYBER COLLATERAL DAMAGE?

As we have demonstrated, traditional definitions of collateral damage do not apply cleanly to cyber effects. Despite the difficulties involved in formally transposing rules and concepts from one domain to another, we intuitively feel—and experience bears out—that cyber operations can have significant, harmful, unintended consequences. And in cyberspace, unknown (and potentially unknowable) interdependencies, in which systems, networks, and code depend on each other in ways that might not be apparent, make the task of determining cyber collateral damage more difficult.[46]

Notwithstanding these challenges, for the purpose of this Article, we define cyber collateral damage as: *Unintended harm to a computer or information system that is not the target of a lawful cyber operation.*[47] *Where "harm" is defined as either a) the deletion, manipulation, or alteration of computer code*

46. To be fair, this challenge would exist whether initiated by a kinetic or cyber attack.

47. The recognition of unintended consequences raises a potential confounding issue regarding degree of causality. That is, how many causal links down the chain of effects is it practical in order to attribute a negative consequence to an actor? We do not attempt to resolve this issue here, but merely offer that simply one step removed is sufficient, if only because beyond this, there would likely exist too many confounding variables.

governing the operation of hardware or software that is not specifically intended by the party conducting a lawfully-authorized operation, or b) the compromise of the integrity or availability of a computer network or data, or exfiltration of data, that is not specifically intended by the party conducting a lawfully-authorized operation.[48,49]

This definition accomplishes two main objectives. First, it broadens the scope of collateral damage assessments to include harm inflicted upon data without physical effects. And second, it captures scenarios in which a particular activity is intended (e.g. seizing a server) even if the specific harm done includes damage that is both intended and unintended (where, for example, that server houses content that is both the legitimate target of a cyber operation and content that is not and could not be the target of a lawful cyber operation).

The first component of the definition—harm to data—is explained in detail above. This represents a departure from the way in which collateral damage is generally conceptualized, but is justified because of the critical importance that data plays in a huge range of social contexts, and because of the significant ramifications for public trust if certain particularly sensitive data sets are corrupted or destroyed.

The second component of the definition is meant to capture the impact that the shared nature of technical infrastructure can have on the collateral damage that results from different types of cyber operations. Thus, if a server hosts both lawful and unlawful content (e.g. child pornography), a law enforcement operation to block the unlawful content may also deny the legitimate content. This might be intended in that the law enforcement agency might block the unlawful content knowing that it will also block a significant amount of lawful content (an example of this phenomenon is discussed below). But in this situation the lawful content should be considered collateral damage and the government should be forced to weigh the degree of such damage against the anticipated benefit of the operation. Because the lawful content could not itself be the target of a lawful cyber operation, the definition attempts to distinguish between lawful targets and intended targets (which the server in this hypothetical example would be). In doing so it also is consistent with the traditional definition of collateral damage, which focuses on harm to targets that are not lawful targets.

Experiences in recent years have illustrated the kinds of unintended harm that can befall computer systems as a result of cyber operations in military and non-military (domestic law enforcement) contexts. While we explore only a few concrete examples below, we expect that with the ever-growing use of, and

48. We may consider a nuance regarding accidental vs inadvertent. A bomb dropped on a military formation that is blown off-course, destroying a civilian target is an accident. While, a bomb dropped on what is assessed to be a military formation but turns out to be a civilian target is inadvertent. However, for the purpose of this Article, both are considered collateral damage,

49. We do not specifically address the issue of how collateral harm to a computer system should be considered when the effect is temporary and fully reversible.

reliance on, cyber capabilities, such examples are only likely to spread and increase in number. As such, the examples described are intended to be illustrative, rather than exhaustive or definitive.

Again, for the purpose of identifying cyber collateral damage, we focus on the outcome of an operation, whether caused by kinetic or non-kinetic actions.

A. Military Examples

In 2003, during the early part of the Iraq war, the U.S. military physically destroyed communication systems in Iraq as part of a larger attack. In addition, however, it also disabled satellite and other communications equipment that provided service not only to Iraqi military forces, but also to civilians in Iraq and neighboring countries.[50] In a similar example, in 2008, a military cyber operation reportedly dismantled a web forum which was hosted on a server in Iraq and used by al-Qaida to plan operations against American troops. However, this operation also impacted the internet connectivity and IT systems of computers in Saudi Arabia, Germany, and Texas.[51] We posit that in both of these examples, the damage to civilian communications should reasonably be considered as cyber collateral damage.

There are also instances where the military has decided against operations because of their potential collateral effects. For example, during congressional testimony in 2016, a U.S. military official stated that the military would not disrupt internet access to areas controlled by ISIS because the consequences—denying internet connectivity to civilians who lived in the area and depended on that connectivity—would be unjustified.[52]

In addition, in 2003, in the lead up to the Iraq war, the Pentagon reportedly developed plans to engage in a cyber attack against Iraq's banking system. However, because it was estimated that the attack would generate such extreme consequences on regional and global financial stability, as well as on civilian infrastructure and IT systems, the plan was never executed.[53]

B. Domestic Examples

Like the military, U.S. law enforcement agencies conduct cyber operations. And while there are fewer specific examples of collateral damage in the non-military context, they do exist—as we show below—and will continue to surface.

---

50. John Markoff and Thom Shanker, *Halted '03 Iraq Plan Illustrates U.S. Fear of Cyberwar Risk*, N.Y. TIMES (Aug. 1, 2009), http://www.nytimes.com/2009/08/02/us/politics/02cyber.html.

51. Ellen Nakashima, *Dismantling of Saudi-CIA Web site illustrates need for clearer cyberwar policies*, WASHINGTON POST (Mar. 19, 2010) http://www.washingtonpost.com/wp-dyn/content/article/2010/03/18/AR2010031805464.html.

52. Peter Micek and Deji Olukoton, *U.S. military official: Internet shutdowns don't help during conflict,* ACCESSNOW (June 22, 2016) https://www.accessnow.org/military-official-internet-shutdowns-isis-conflict/.

53. Alex Goldman, *Cyber War Could Cause Global Collateral Damage*, INTERNETNEWS (August 04, 2009) http://www.internetnews.com/security/article.php/3833131/Cyber+War+Could+Cause+Global+Collateral+Damage.htm.

One of the clearest examples in which a government agency struggled with cyber collateral damage involved a Pennsylvania statute designed to block child pornography. Specifically, the statute enabled law enforcement officials to obtain a court order requiring ISPs to block child pornography. However, internet service providers maintained that no matter which blocking technique they used (DNS filtering, IP filtering, or URL filtering) the blocking resulted in an excessive impact on uninvolved legitimate web traffic. Ultimately, legitimate users who lost access as a result of the blocked traffic filed suit in federal District Court, claiming that their First Amendment rights had been infringed. Because the "burden on protected expression [was] substantial whereas there is no evidence that the Act has impacted child sexual abuse" the court upheld the plaintiffs' claims and struck down the statute on First Amendment grounds.[54]

Botnet takedowns offer another context in which anticipated collateral damage is important. Botnets are networks of hijacked computers that can be controlled remotely by cybercriminals for nefarious ends—the theft of banking credentials, commission of advertising fraud ("clickfraud"), or the conduct of denial of service attacks. Botnet takedowns often rely on coordinated public and private legal actions in which private companies obtain injunctive relief of various forms and court orders mandating that ISPs shut down traffic from command and control servers while law enforcement agencies seize servers that host malicious traffic.[55]

These actions have salutary effects—terminating the mechanism by which a wide range of cybercrimes take place–but they also can have unwanted collateral effects. For example, in 2014, Microsoft sought to disrupt a massive criminal botnet operation by confiscating 22 subdomains operated by an internet service provider—domains which it alleged operated and distributed malware. The confiscation of these subdomains, however, resulted in users being unable to access legitimate parts of those domains.[56] In addition, the takedown of the "No-IP" botnet, rendered unavailable a considerable amount of legitimate, in addition to malicious, web traffic.[57] To the extent that botnets are disrupted or, indeed, commandeered by government actors, takedowns that yield these collateral effects could result in Fourth Amendment claims that authorities had seized domain names in an unreasonable manner. Before engaging in botnet takedowns, government actors will likely need to be

54. Ctr. For Democracy & Tech. v. Pappert, 337 F. Supp. 2d 606, 656 (E.D. Pa. 2004).

55. *See, e.g.*, Zachary Goldman and Damon McCoy, *Deterring Financially Motivated Cyber Crime*, 8 J. NAT'L SECURITY L & POL'Y 595 (2016) (describing how botnets are used and the legal means by which they are disrupted).

56. Dan Goodin, *Millions of dynamic DNS users suffer after Microsoft seizes No-IP domains*, ARS TECHNICA (June 30, 2014), http://arstechnica.com/security/2014/06/millions-of-dymanic-dns-users-suffer-after-microsoft-seizes-no-ip-domains/.

57. Brian Krebs, *Microsoft Darkens 4MM Sites in Malware Fight, Krebs on Security*, KREBS ON SECURITY (Jul. 14, 2014), http://krebsonsecurity.com/2014/07/microsoft-darkens-4mm-sites-in-malware-fight/comment-page-1/.

confident that they can estimate the effects—both intended and potentially unintended—with confidence.

Finally, the use of private sector "hack back" techniques—which are currently prohibited by the Computer Fraud and Abuse Act (CFAA)[58] but are the subject of a lively legal reform debate—is a cyber method that calls out for improvements in estimating collateral damage. At present, the CFAA prohibits natural or legal persons from gaining unauthorized access to other peoples' computer networks, even if they are first the victim of an attack.[59] But as the cost of cybercrime continues to grow, some have called for modifying the CFAA to permit companies to retaliate against hackers in certain circumstances.[60] Such calls have not been heeded, however, primarily because of concerns that private sector retaliatory attacks might have unanticipated collateral effects: they might disrupt legitimate web traffic or cause unanticipated damage and inadvertently escalate situations in ways that undermine or thwart government cyber efforts. To the extent that private companies are able to estimate collateral damage with confidence it might be possible to design a regime that penalizes cybercrime perpetrators without inadvertently unleashing some of the evils associated with an unrestrained "hack back" tactic.

This section introduced some non-military contexts in which cyber collateral damage occurs and needs to be more effectively governed. The next section will suggest ways in which the military's traditional means of estimating collateral damage can be adapted for the digital age.

III. COLLATERAL DAMAGE ESTIMATION

A. For Military Cyber Operations

In this section, we examine the U.S. military's process for estimating collateral damage, and seek to adapt from it an analogous methodology for cyber operations. Unsurprisingly, the collateral damage estimation methodology (CDEM), is heavily focused on notions of conventional munitions, physics-based computer models, physical environments and surroundings, geospatial targeting, (physical) structural composition and damage, distance-based warhead blast (fragmentation) effects, and error calculations.[61]

The CDEM does not account for secondary effects, such as explosions from weapons or fuel depots. However, while the focus is clearly on common non-strike (i.e. civilian) entities, it does mention a number of nontraditional

58. *See* 18 U.S.C. § 1030.

59. *Id* at § 1030a.

60. *E.g.*, JUAN C. ZARATE, CYBER FINANCIAL WARS ON THE HORIZON: CONVERGENCE OF FINANCIAL AND CYBER WARFARE AND THE NEED FOR A 21ST CENTURY NATIONAL SECURITY RESPONSE 26 (2015), http://www.defenddemocracy.org/content/uploads/publications/Cyber_Financial_Wars.pdf.

61. CJCSI 3160.01, *supra* note 1, at § D. Further, the CDE makes clear that it does not account for nuclear, non-kinetic or nonlethal capabilities, surface-to-surface direct fire weapons, air-to-surface direct fire weapons smaller than 105mm. *Id.* at § D-4.

cyber-related objects such as computer networks, websites, IP addresses,[62] and bank accounts.[63] That being said, the core no-strike list category codes[64] do not identify any such information categories, nor is there any consideration for targeting, for example, a bank account or IP address.

The CDEM addresses five main questions:[65]

1. Can I positively identify the target?
2. Are there any civilian people or objects within the effects range of the target?
3. Can I reduce the collateral damage by using a different weapon or approach and still accomplish the mission?
4. What is my estimate of the collateral damage to persons and objects?
5. Do those estimates of collateral damage exceed the relative military benefits that would be achieved? [66]

The first two steps of the CDEM involve determining whether the targeted object (or person) is a lawful military target "in accordance with the [laws of war] and applicable [rules of engagement]." That is, the object should be geospatially located, and evaluated as to whether it serves military, dual, or pure civilian use. The third step involves a weaponeering exercise in order to determine whether an alternative weapon could be used to reduce the amount of damage (collateral and otherwise), while still satisfying the mission's goal. The fourth and fifth steps seek to evaluate the extent of collateral damage once a final weaponeering decision has been made, and to determine whether that amount of damage is appropriate relative to the anticipated military benefit.

Given that this methodology was developed for a kinetic environment, we next look to adapt the CDEM to cyber operations. In effect, we seek to develop a CDEM for cyber, or CDEM-C.

Step 1: Recall, the first step of the CDEM is concerned with identifying the physical location of the target. In conventional warfare, the act of identifying the target and targeting the munition can be straightforward. For example, a physical object can be located using longitude and latitude coordinates on a map. While computing equipment can also be identified using longitude/latitude coordinates, the data that make up an internet application or database can exist in many places simultaneously and may even be distributed across multiple physical systems. Therefore, what does it mean to *identify* a target in cyberspace?

62. It is unclear why an IP address would be considered a non-strike entity.
63. *Id.* footnote 6.
64. CJCSI 3160.01, *supra* note 1, at tbl.C-A-1.
65. *Id.* at § D-A-7.
66. This is the so-called Rule of Proportionality.

Recall that U.S. military doctrine considers three layers of cyberspace operations: the physical network, logical network, and the persona. This is a useful construct when understanding how targets may be identified in a cyberspace operation.

First, consider the persona layer. An operation may require establishing contact with an individual through her social media account in order to foster a relationship and exchange messages. In this case, the 'target' is simply that online user account, uniquely identified by the account name.

Next, consider targeting a website in the logical network layer. In this case the most appropriate measure for a target may be the IP address, because it is the website's IP address that enables a networked computer to send and receive messages with other computers.[67] In other cases, the MAC address of the device may be necessary to uniquely identify and target the object. In addition to the web site, the operation may also target other content such as data within a database server, or user credentials.

Finally, the physical network layer provides the most straightforward analogy to a kinetic operation because it concerns the geographic location of a physical device, cable, or pieces of IT equipment, and that device or object will unambiguously reside in land, air, or space. Therefore, consider an operation that requires infiltrating an office building in order to infect a computer, but the mission is agnostic to which computer, since they are all part of the same network segment. In this case, only the physical location of the office, and the computers within it are important. Similarly, consider an operation that requires covertly entering a house (or other facility), and installing malware on the computer of a particular individual. In this case, the individual (or the individual facility) is the ultimate target, but for the purpose of estimating collateral damage, only the physical location of the building and the individual's computer are relevant.

In addition, it is conceivable that an operation may involve targets that exist in across many of these layers. For example, a mission may first call for befriending an individual on her private social media account in order to learn personal information about the individual, such as a contact information at an office. The operation may then employ social engineering in order to trick the user into opening an email with a malicious attachment. Once executed, the malware could then seek out a specific computer or network segment in order to compromise the ultimate target which may be a physical supervisory control and data acquisition (SCADA) control system (e.g. a dam or power facility). In this case, targets within each of the persona layer (the social media account), the logical network layer (the control system computer), and the physical layer (the SCADA system) are involved in the operation.

---

67. Although a device does not need an IP address to listen to network communication, it does need one to interact on that network. Networks themselves—particularly networks carrying an adversary's sensitive national security information—might be isolated from the commercial Internet and might require bespoke access mechanisms.

The point of these examples is to demonstrate that one distinguishing feature of cyber operations, relative to kinetic ones, is that targeting may involve a combination of one or more of the persona, physical network or logical network layers, and that collateral damage estimation may be required for each of these layers.

Step 2: The second step of the CDEM is concerned with understanding and estimating the extent to which civilian or non-combatant assets could be affected. In our new context, we must consider effects both to information technology and physical systems. Of all the steps of the CDEM, one of the most difficult is accurately estimating the direct and collateral effects on non-combatant assets, here again, the challenge lies in the fact that information technology is implemented as a complex system with relationships and dependencies that can be nonobvious, poorly documented, and far-reaching. For example, in 2003, a single computer vulnerability and a series of cascading failures of the electrical grid caused power loss to almost 60 million people across the east coast of the United States and Canada.[68] In addition, consider any of the major computer worms or viruses that have affected modern computing systems, such as Slammer, Slapper, Blaster, etc. Each was designed as malicious code to affect computing systems, but their sheer impact across the internet illustrates the scope of the damage that can result from a single attack.

Therefore, the second step of the CDEM-C must account for (insofar as it is possible) the connectivity, reliance, and interdependence of computing systems, data, and services. As previously mentioned, we do not claim to solve the difficult problem of how to model and predict dependencies of software and hardware systems in this Article. However, we do suggest that, while difficult, this is a solvable problem. Despite the seemingly chaotic network of technology that underpins an organization, a city and a country, these are deterministic systems which operate according to specific instructions. They are not machines driven by random processing. Therefore, with enough information about the hardware, software, and third party relationships of a target, it should be possible to fully identify both direct and collateral (whether kinetic or cyber) effects resulting from the compromise of a computing system. Therefore, determining the anticipated collateral effects of a proposed cyber operation is a bounded (if difficult) challenge. If it is possible to estimate collateral damage then it is possible to plan lawful, legitimate, and effective cyber operations.

Step 3: The third step considers whether a different cyber weapon could be used that would achieve the same goal, but reduce the expected collateral impacts. This step could be accomplished in a number of ways, each of which depends on the source of collateral effects.

First, one may consider tailoring a computer worm or virus to target not just any computing system, but one with a narrower set of parameters. For example,

68. Kevin Poulsen, *Software Bug Contributed to Blackout*, SECURITYFOCUS (Feb. 11 2004), http://www.securityfocus.com/news/8016.

in the Stuxnet attack, the developers appeared to go to great lengths to tailor the impact of the code to the specific kind of control system used by Iranian reactors.[69] The code that constituted the cyber weapon reportedly targeted only the specific model of industrial control system that operated Iran's nuclear facilities and ceased to continue propagating itself after it had infected three machines. Richard Clarke, the cybersecurity coordinator during the Bush administration, considered these steps to be a deliberate attempt to reduce any effects to proximate IT systems.

In addition, one could alter the type of attack so as to reduce kinetic and cyber impacts to IT systems and data, such as attacking a different system that would cause an equally desired effect. Further, one could render a system neutral rather than destroying it, disable a particular capability, or sever a network connection of an adversary. In addition, one could modify the environment or circumstances in such a way as to reduce the impact, such as attacking during a different time of day.

Step 4: The fourth step of the CDEM-C is to re-evaluate the discussion from the second step in which the operator estimates the collateral impacts to civilian persons or objects. By this point, no new information should be necessary; rather, this step simply requires performing the same evaluation using new criteria.

Step 5: The final step addresses the question of whether, given the specific cyber weapon (potentially modified in steps 3 and 4), the anticipated effects are appropriate given the conditions at the time. Using the definition offered above, the commander would determine whether the anticipated collateral harms outweigh the objectives. If so, then the operation should be reassessed. Otherwise the operation would proceed with reasonable confidence.

Each step of the CDEM is summarized in Table 1, along with the analogous steps for the CDEM-C.

Table 1. Cyber CDEM

CDEM		Cyber CDEM
1.	Can I positively identify the target?	Can I positively identify the target's online persona, IP address, network, or computer location?
2.	Are there any civilian people or objects within the "effects range" of the target?	Are there any civilian data, or IT systems located on the same subnet as, or dependent upon connectivity with, the target? Or, are there any services or functions which rely on the targeted system?

69. Jon Lindsay, *Stuxnet and the Limits of Cyber Warfare*, 22 SECURITY STUDIES 365, 387 (2013).
70. We consider that this should apply equally to data resident in a cloud service provider.

CDEM		Cyber CDEM
3.	Can I reduce the collateral damage by using a different weapon or approach and still achieve my goal?	Can I reduce the collateral damage by exploiting a different vulnerability, adjusting the circumstances of attack, or launching a different operation and still achieve my goal?
4.	What is my new estimate of the collateral damage to persons and objects?	What is my new estimate of the range of collateral effects to data, computing, or IT systems?
5.	Given those estimates, do I still comply with the rule of Proportionality?	Given those estimates, do I still comply with the rule of Proportionality?

Now that we have developed a modest approach for estimating collateral damage for cyber effects, we next provide a similar approach in the context of domestic law enforcement.

B. For Law Enforcement Cyber Operations

The previous sections examined collateral damage from the perspective of cyber operations. But U.S. law enforcement agencies also conduct cyber operations and must grapple with similar considerations. Therefore, the scholarly and policy communities should also work to develop conceptual frameworks for evaluating collateral damage from proposed cyber operations in these contexts.

To some extent this work has begun. In at least one instance (described above), a federal court, weighing the collateral effects of a cyber operation, struck down the government's attempt to establish a content takedown regime for child pornography sites because the regime had overly broad collateral effects. There are other contexts in which the cybersecurity community would benefit from a refined conceptual framework for evaluating collateral damage. For instance, consider the previously described examples involving botnet takedowns and the use of hack back techniques by private companies that have been the victims of cyberattacks. In both instances, ongoing debates about the wisdom and efficacy of these cyber responses hinge, in part, on determining acceptable levels of cyber collateral damage. With a more refined framework for evaluating cyber collateral damage, lawmakers and government officials can make more effective decisions about expanding legal authorities to engage in these activities.[71]

While the CDEM-C discussed above emerged from the military context, the concepts and framework apply with equal force to the evaluation of domestic/

71. *See Cyber Crime: Modernizing Our Legal Framework for the Information Age*, Hearing Before the S. Comm. on the Judiciary, Subcommittee on Crime and Terrorism, 114th Cong. (2015) (statement of David M. Bitkower, Deputy Assistant Attorney General, Criminal Division, U.S. Dep't of Justice) (advocating for expanded legal authorities to combat botnets).

law enforcement cyber operations. But while the methodology for assessing collateral damage for domestic operations will have many of the same characteristics as in the military context, there will be a few notable differences for domestic law enforcement, which we describe below.

Step 1: First, the legitimacy of domestic cyber operations will depend on identifying an appropriate target—what is the action that law enforcement agencies or private companies are attempting to take, and against whom? Determining the appropriate target, however, differs significantly between the law enforcement and military contexts in terms of the process for doing so, and the institutions involved. At present, cyber interventions undertaken by law enforcement agencies are conducted pursuant to court orders. These can be search warrants (or search warrant-like instruments) or injunctions. But they typically involve courts concurring with an initial law enforcement assessment that the proposed target of a cyber operation is in fact involved in some kind of criminal activity and then licensing law enforcement interventions through an appropriate legal instrument. The main difference between law enforcement and military cyber operations then is the interposition of an external reviewer on law enforcement's proposed cyber activities in the form of a court that reviews and authorizes them ex ante. This external review should have two important effects. First, it will presumably engender the development of refined methods for estimating the collateral impact of cyber operations given the powerful ex ante effects of anticipated judicial scrutiny on government decision-making.[72] Second, it should broaden the base of legitimacy for those activities as the public becomes confident that law enforcement agencies have sought and received external approval from "detached and neutral magistrates" for their actions.

The second and third considerations—whether there are any uninvolved objects that will be affected by the contemplated action, and whether there are steps that one can take to select an intervention that will minimize the harm to uninvolved objects—should also be considered in the context of law enforcement cyber operations.

Step 2: Here, the need for warrants or court orders to engage in cyber operations could provide a vehicle to ensure that these questions have been considered as thoroughly as possible. And to the extent that a given operation is governed by the Fourth Amendment (because it is a search or a seizure) the Amendment's requirements that all such interventions be reasonable may provide a built-in framework for analyzing when anticipated collateral damage may be excessive in comparison to the legitimate law enforcement objective being pursued. A fundamental challenge, however, will remain: given that law enforcement (or private actors) may not always know in advance how networks are

72. *See generally* Ashley Deeks, *National Security Litigation, Executive Policy Changes, and Judicial Deference*, 82 FORDHAM L. REV. 828 (2013) (discussing the impact that anticipated judicial review has on executive branch legal decision making).

structured or what endpoint devices are connected to them, it might not always be possible to anticipate these collateral effects with confidence.

Step 3: The third step in the traditional CDEM is considering whether the same effects can be achieved with less collateral damage. A similar calculation is necessary in the CDEM-C context, but, because of the interdependencies identified above, may be harder to determine. In the law enforcement context, some "least restrictive means" requirements exist, generally in the investigative context.[73] In that instance, the harm that investigators must minimize is damage to privacy, so the regimes built around the Wiretap Act and other investigative regimes (e.g. the FBI's Domestic Operations Manual) will not apply directly to the CDEM-C framework (and in any event "least restrictive" is not quite the same as "least destructive"). But conceptually the idea is similar and should be incorporated into domestic cyber operations.

Steps 4 and 5: The final steps—estimating the collateral damage and determining whether the law enforcement action complies with applicable rules of proportionality—are also similar. The main difference is that there will often be an external body, namely a court, making these determinations when they decide whether to authorize the government's proposed cyber activities in the first instance. This judicial involvement and the application of the Fourth Amendment's legal framework may have an important disciplining effect on proposed cyber operations, and certainly should incentivize the development of more precise methodologies for determining the collateral effects of proposed cyber operations. It also, however, puts a premium on judicial expertise in cyber activities and the ability to evaluate the government's claims about collateral damage and proportionality.

Each of these steps is summarized in Table 2, along with the analogous step for the CDEM-C.

Table 2. Domestic law enforcement approach

CDEM		**Domestic LE Approach**
1.	Can I positively identify the target?	LE must identify target in legal process (warrant or other instrument).
2.	Are there any civilian people or objects within the "effects range" of the target?	LE to determine whether there are any civilian data, or IT systems located on the same subnet as, or dependent upon connectivity with, the target. Or, are there any services or functions which rely on the targeted system?
3.	Can I reduce the collateral damage by using a different weapon or approach and still achieve my goal?	LE to comply with *least restrictive means* test

73. *See e.g.*, Wiretap Act, 18 U.S.C. § 2510 et seq. (2008).

CDEM		Domestic LE Approach
4.	What is my new estimate of the collateral damage to persons and objects?	Operational planning and legal process should account for revisions to initial plans.
5.	Given those estimates, do I still comply with the rule of Proportionality?	Court will evaluate the anticipated benefit of the LE activity and grant the legal process, or require further operational refinements.

CONCLUSION

This Article has identified a critical and growing disconnect regarding how collateral damage is understood in the conventional military context and how it should be understood in the cyber domain. In the military context, it very narrowly relates only to damage resulting from a hostile action that causes physical or property damage to a civilian target. However, this suggests that common cyber effects, such as espionage, denial of service, or disruption of critical infrastructure, would fall short of being characterized as causing damage (collateral or otherwise), and therefore never be considered an attack. Intuitively, however, we recognize that unintended effects caused by cyber activities have the potential for causing effects that are much broader than traditional military doctrine would describe. This Article has begun to lay the conceptual groundwork for a richer understanding of this critical disconnect—one that reflects the pervasiveness of cyber and cyber-physical systems in our digital world.

There are two main challenges in estimating collateral damage in the cyber context. The first is determining whether one can estimate the collateral consequences of cyber operations with confidence, and the main obstacle to doing so is the pervasiveness of unknown interdependencies in cyberspace. Cyber operators need to have confidence in their ability to predict the impact of proposed cyber operations. Only when cyber operators can predict in advance what the unintended consequences of their operations may be, can they design meaningful approaches to mitigate unwanted and harmful effects. While it may be difficult to determine these interdependencies with confidence it is not impossible—cyber systems are products of human engineering and have only the properties that we give them. All that is left to do is to map those properties and see how they interact with each other.

The second, and perhaps most conceptually difficult, is deciding what harms we seek to guard against. As in kinetic strikes, cyber operations that cause unintended physical harms ought to be considered collateral damage and minimized. The more difficult case is one in which a cyber operation causes only unintended "harm" (alteration or deletion) to data with no physical effects. Is this to be considered collateral damage? Will that stretch the concept of collateral damage too far? Does it raise difficult questions of thresholds—how much manipulation of data will need to take place for unacceptable collateral

damage to result? While we have provided preliminary answers above, these and related questions will spark much-needed additional research.

With time, we predict that the military and law enforcement cyber operations communities will need to develop a rigorous framework for evaluating collateral damage. We hope that the insights offered in this Article help lay the groundwork for formally and appropriately integrating and assessing digital harms into military and law enforcement operations.

***