The End of Cyber-Anarchy? How to Build a New Digital Order

Published in: Foreign Affairs, 2022, Gale General OneFile

Ransomware attacks, election interference, corporate espionage, threats to the electric grid: based on the drumbeat of current headlines, there seems to be little hope of bringing a measure of order to the anarchy of cyberspace. The relentless bad news stories paint a picture of an ungoverned online world that is growing more dangerous by the day--with grim implications not just for cyberspace itself but also for economies, geopolitics, democratic societies, and basic questions of war and peace.

Given this distressing reality, any suggestion that it is possible to craft rules of the road in cyberspace tends to be met with skepticism: core attributes of cyberspace, the thinking goes, make it all but impossible to enforce any norms or even to know whether they are being violated in the first place. States that declare their support for cybernorms simultaneously conduct large-scale cyber-operations against their adversaries. In December 2015, for example, the UN General Assembly for the first time endorsed a set of 11 voluntary, nonbinding international cybernorms. Russia had helped craft these norms and later signed off on their publication. That same month, it conducted a cyberattack against Ukraine's power grid, leaving roughly 225,000 people without electricity for several hours, and was also ramping up its efforts to interfere in the 2016 U.S. presidential election. For skeptics, this served as yet further evidence that establishing norms for responsible state behavior in cyberspace is a pipe dream.

Yet that skepticism reveals a misunderstanding about how norms work and are strengthened over time. Violations, if not addressed, can weaken norms, but they do not render them irrelevant. Norms create expectations about behavior that make it possible to hold other states accountable. Norms also help legitimize official actions and help states recruit allies when they decide to respond to a violation. And norms don't appear suddenly or start working overnight. History shows that societies take time to learn how to respond to major disruptive technological changes and to put in place rules that make the world safer from new dangers. It took two decades after the United States dropped nuclear bombs on Japan for countries to reach agreement on the Limited Test Ban Treaty and the Nuclear Nonproliferation Treaty.

Although cybertechnology presents unique challenges, international norms to govern its use appear to be developing in the usual way: slowly but steadily, over the course of decades. As they take hold, such norms will be increasingly critical to reducing the risk that cybertechnology advances could pose to the international order, especially if Washington and its allies and partners reinforce those norms with other methods of deterrence. Although some analysts argue that deterrence does not work in cyberspace, that conclusion is simplistic: it works in different ways than in the nuclear domain. And alternative strategies have proved equally or more deficient. As targets continue to proliferate, the United States must pursue a strategy that combines deterrence and diplomacy to strengthen the guardrails in this new and dangerous world. The record of establishing norms in other areas offers a useful place to start--and should dispel the notion that this issue and this time are different.

A NEW FACT OF LIFE (AND WAR)

As cyberattacks become more costly, U.S. strategy to defend against them remains inadequate. A good strategy has to begin at home but simultaneously recognize the inseparability of cyberspace's domestic and international aspects--the domain of cyberspace is inherently transnational. Furthermore, cybersecurity involves a blurring of public and private vulnerabilities. The Internet is a network of networks, most of which are privately owned. Unlike nuclear or conventional weapons, the government does not control them. Accordingly, companies make their own tradeoffs between investing in security and maximizing short-term profit. Yet inadequate corporate defense can have huge external costs for national security: witness the recent Russian cyberattack on SolarWinds software, which allowed access to computers across the U.S. government and the private sector. And unlike with military security, the Pentagon plays only a partial role.

In the realm of global military conflict, computer networks have become a fifth domain, in addition to the traditional four of land, sea, air, and space, and the U.S. military recognized this with the creation of U.S. Cyber Command in 2010. Among the special characteristics of the new cyber-domain are the erosion of distance (oceans no longer provide protection), the speed of interaction (much faster than rockets in space), the low cost (which reduces barriers to entry), and the difficulty of attribution (which promotes deniability and slows responses). Still, skeptics sometimes describe cyberattacks as more of a nuisance than a major strateg problem. They argue that the cyber-domain is ideal for espionage and other forms of covert action and disruption but that it remains far less important than the traditional domains of warfare; no one has died because of a cyberattack. That, however, is becoming an increasingly difficult position to take. The 2017 WannaCry ransomware attack damaged the British National Health Service by leaving computers encrypted and unusable, forcing thousands of patients' appointments to be canceled, and hospitals and vaccine producers have been directly targeted by ransomware attacks and hackers during the covid-19 pandemic.

What's more, there remains much that even experts do not understand about how the use of cybertools could escalate to physical conflict. Consider, for example, the fact that the U.S. military depends heavily on civilian infrastructure and that cyber-penetrations could seriously degrade U.S. defensive capabilities in a crisis situation. And in economic terms, the scale and cost of cyber-incidents have been increasing. According to some estimates, the Russian-sponsored 2017 NotPetya attack on Ukraine, which wiped data from the computers of banks, power companies, gas stations, and government agencies, cost companies more than $10 billion in collateral damage. The number of targets is also expanding rapidly. With the rise of big data, artificial intelligence, advanced robotics, and the Internet of Things, experts estimate that the number of Internet connections will approach a trillion by 2030. The world has experienced cyberattacks since the 1980s, but the attack surface has expanded dramatically; it now includes everything from industrial control systems to automobiles to personal digital assistants.

It is clear that the threat is mounting. Less clear is how U.S. strategy can adapt to face it. Deterrence must be part of the approach, but cyber-deterrence will look different from the more traditional and familiar forms of nuclear deterrence that Washington has practiced for decades. A nuclear attack is a singular event, and the goal of nuclear deterrence is to prevent its occurrence. In contrast, cyberattacks are numerous and constant, and deterring them is more like deterring ordinary crime: the goal is to keep it within limits. Authorities deter crime not only by arresting and punishing people but also through the educational effect of laws and norms, by patrolling neighborhoods, and through community policing. Deterring crime does not require the threat of a mushroom cloud.

Still, punishment plays a large role in cyber-deterrence. The U.S. government has publicly stated that it will respond to cyberattacks with weapons of its choice and with force proportional to the harm inflicted on its interests. Despite a decade of warnings, thus far, a "cyber-Pearl Harbor" has not happened. Whether the United States treats a cyberattack as an armed attack depends on its consequences, but this makes it difficult to deter actions that are more ambiguous. Russia's disruption of the 2016 U.S. presidential election fell into such a gray area. And although some recent Chinese and Russian cyberattacks appear to have been conducted primarily for the purposes of espionage, the Biden administration has complained that their scale and duration moved them beyond normal spying. This is why deterrence in cyberspace requires not just the threat of punishment but also denial by defense (building systems resilient enough and hard enough to break into that would-be attackers won't bother to try) and entanglement (creating links to potential adversaries so that any attack they launch will likely harm their own interests, too). Each of these approaches has limits when used on its own. Entanglement has more of an effect when used against China, because of a high degree of economic interdependence, than it does against North Korea, with whom there is none. Denial by defense is effective in deterring nonstate actors and second-tier states but less likely to prevent attacks by more powerful and proficient actors. But the combination of a threat of punishment and an effective defense can influence these powers' calculations of costs and benefits.

In addition to improving the defense of networks inside the United States, in recent years, Washington has adopted doctrines that U.S. Cyber Command has dubbed "defend forward" and "persistent engagement"--simply put, small-scale acts of cyberoffense, such as the disruption, diversion, or takedown of a network. Some press accounts credit these practices with reducing Russian interference in the 2018 and 2020 U.S. elections. But entering and disrupting an adversary's network poses some danger of escalation and must be carefully managed.

SETTING SOME RULES

Despite its defensive and offensive capabilities, the United States remains highly vulnerable to cyberattacks and influence operations, owing to its free markets and open society. "I think it's a good idea to at least think about the old saw about [how] people who live in glass houses shouldn't throw rocks," remarked James Clapper, then the director of national intelligence, during 2015 congressional testimony on Washington's responses to cyberattacks. Clapper was stressing, rightly, that although Americans may be the best at throwing stones, they live in the glassiest of houses. That reality gives the United States a particular interest in the development of norms that reduce incentives to throw stones in cyberspace.

Negotiating cyber-arms-control treaties would be extremely difficult, because they would not be verifiable. But diplomacy on cyberspace is hardly impossible. In fact, international cooperation on developing cybernorms has been going on for more than two decades. In 1998, Russia first proposed a un treaty to ban electronic and information weapons. The United States rejected the idea, arguing that a treaty in this area would be unverifiable because whether a line of code is a weapon or not can depend on the intent of the user. Instead, the United States agreed that the un secretary-general should appoint a group of 15 (later expanded to 25) government experts to develop a set of rules of the road; they first met in 2004.

Six such groups have convened since then, and they have issued four reports, creating a broad framework of norms that was later endorsed by the UN General Assembly. The groups' work has strengthened the consensus that international law applies to the domain of cyberspace and is essential for maintaining peace and stability in it. In addition to grappling with complicated questions of international law, the report that was issued in 2015 introduced 11 voluntary, nonbinding norms, the most important ones being a mandate to provide states with assistance when requested and prohibitions against attacking civilian infrastructure, interfering with computer emergency response teams, which respond after big cyberattacks, and allowing one's territory to be used for wrongful acts.

The report was viewed as a breakthrough, but progress slowed in 2017 when the expert group failed to agree on international legal issues and did not produce a consensus report. At Russia's suggestion, the un supplemented the existing process by forming the Open-Ended Working Group, which is open to all states and involves consultations with nongovernmental actors: dozens of private companies, civil society organizations, academics, and technical experts. Early in 2021, this new group issued a broad, if somewhat anodyne, report that reaffirmed the 2015 norms, as well as the relevance of international law to cyberspace. Last June, the sixth expert group also completed its work and released a report that added important details to the 11 norms first introduced in 2015. China and Russia are still pressing for a treaty, but what is more likely to happen is the gradual evolution of these norms.

In addition to the un process, there have been many other forums for discussion about cybernorms, including the Global Commission on the Stability of Cyberspace. Initiated in 2017 by a Dutch think tank, with strong support from the Dutch government, the GCSC (of which I was a member) was co-chaired by Estonia, India, and the United States and included former government officials, experts from civil society, and academics from 16 countries. The GCSC proposed eight norms to address gaps in the existing un guidance. The most important were calls to protect the "public core" infrastructure of the Internet from attack and to prohibit interference with electoral systems. The CCSC also called on countries not to use cybertools to interfere with supply chains; not to introduce botnets into others' machines in order to control them without the host's knowledge; to create transparent processes that states can follow in judging whether to disclose flaws and vulnerabilities they discover in others' coding; to encourage states to promptly patch cybersecurity vulnerabilities when discovered and not hoard them for possible use in the future; to improve "cyber hygiene," including through law and regulations; and to discourage private vigilantism by making it illegal for private businesses to "hack back," that is, to launch counterattacks against hackers.

These efforts are less flashy (and less expensive) than the development of sophisticated cyberdefense systems, but they will play a crucial role in curbing malign activity online. Many further norms can be imagined and proposed for cyberspace, but the important question now is not whether more norms are needed but how they will be implemented and whether and when they will alter state behavior.

THE NEW PRIVATEERS

Norms are not effective until they become common state practice, and that can take time. It took many decades for norms against slavery to develop in Europe and the United States in the nineteenth century. The key question is why states ever let norms constrain their behavior. There are at least four main reasons: coordination, prudence, reputational costs, and domestic pressures, including public opinion and economic changes.

Common expectations inscribed in laws, norms, and principles help states coordinate their efforts. For example, although some states (including the United States) have not ratified the un Convention on the Law of the Sea, all states treat a 12-mile limit as customary international law when it comes to disputes about territorial waters. The benefits of coordination--and the risks posed by its absence--have been evident in cyberspace on the few occasions when targets have been hacked through abuse of the Internet's domain name system, which is sometimes called "the telephone book of the Internet" and is run by the nonprofit Internet Corporation for Assigned Names and Numbers, or ICANN. By corrupting the phone book, such attacks put the basic stability of the Internet at risk. Unless states refrain from interfering with the structure that makes it possible for private networks to connect, there is no Internet. And so, for the most part, states eschew these tactics.

Prudence results from the fear of creating unintended consequences in unpredictable systems and can develop into a norm of nonuse or limited use of certain weapons or a norm of limiting targets. Something like this happened with nuclear weapons when the superpowers came close to the brink of nuclear war in 1962, during the Cuban missile crisis. The Limited Test Ban Treaty followed a year later. A more distant but historical example of how prudence produced a norm against using certain tactics is the fate of privateering. In the eighteenth century, national navies routinely employed private individuals or private ships to augment their power at sea. But in the following century, states turned away from privateers because their extracurricular pillaging became too costly. As governments struggled to control privateers, attitudes changed, and new norms of prudence and restraint developed. One could imagine something similar occurring in the domain of cyberspace as governments discover that using proxies and private actors to carry out cyberattacks produces negative economic effects and increases the risk of escalation. A number of states have outlawed "hacking back."

Concerns about damage to a country's reputation and soft power can also produce voluntary restraint. Taboos develop over time and increase the costs of using or even possessing a weapon that can inflict massive damage. Take, for example, the Biological Weapons Convention, which came into force in 1975. Any country that wishes to develop biological weapons has to do so secretly and illegally and faces widespread international condemnation if evidence of its activities leaks, as the Iraqi leader Saddam Hussein discovered. It is hard to imagine the emergence of a similar blanket taboo against the use of cyberweapons. For one thing, it is difficult to determine whether any particular line of code is a weapon or not. A more likely taboo is one that would prohibit the use of cyberweapons against particular targets, such as hospitals or health-care systems. Such prohibitions would have the benefit of piggybacking on the existing taboo against using conventional weapons on civilians. During the covid-19 pandemic, public revulsion against ransomware attacks on hospitals has helped reinforce that taboo and suggested how it might apply to other areas in the realm of cyberspace. Something similar might evolve if hackers were to cause an increase in fatal accidents from the use of electric vehicles.

PEER PRESSURE

Some scholars have argued that norms have a natural life cycle. They often begin with "norm entrepreneurs": individuals, organizations, social groups, and official commissions that enjoy an outsize influence on public opinion. After a certain gestation period, some norms reach a tipping point, when cascades of acceptance translate into a widespread belief and leaders find that they would pay a steep price for rejecting it.

Embryonic norms can arise from changing social attitudes, or they can be imported. Take, for example, the spread of concern for universal human rights after 1945. Western countries took the lead in promoting the Universal Declaration of Human Rights in 1948, but many other states felt obliged to sign on because of public opinion and subsequently found themselves constrained by external pressure and by concern about their reputations. One might expect such constraints to be stronger in democracies than in authoritarian states. But the Helsinki process, a series of meetings between the Soviet Union and Western countries in the early 1970s, successfully included human rights in discussions about political and economic issues during the Cold War.

Economic change can also foster a demand for new norms that might promote efficiency and growth. Norms against privateering and slavery gathered support when these practices were economically in decline. A similar dynamic is at work today in the cyberrealm. Companies that find themselves disadvantaged by conflicting national laws relating to privacy and the location of data might press governments to develop common standards and norms. The cyber-insurance industry may put pressure on authorities to flesh out standards and norms, especially in regard to the technology embedded in the myriad household devices (thermostats, refrigerators, home alarm systems) that are now online: the so-called Internet of Things. As more and more devices become connected to the Internet, they will soon become targets for cyberattacks, and the impact on citizens' daily lives will increase and foster demand for domestic and international norms. Public concern will only accelerate if hacking becomes more than a nuisance and begins to cost lives. If fatalities increase, the

Silicon Valley norm of "build quickly and patch later" may gradually give way to norms and laws about liability that place more emphasis on security.

CYBER-RULES ARE MADE TO BE BROKEN

Even with international consensus that norms are needed, agreeing where to draw redlines and what to do when they're crossed is another matter. And the question becomes, even if authoritarian states sign up for normative conventions, how likely are they to adhere to them? In 2015, Chinese President Xi Jinping and U.S. President Barack Obama agreed not to use cyberespionage for commercial advantage, but private security companies reported that China adhered to this pledge for only a year or so before it returned to its old habit of hacking U.S. corporate and federal data, although that happened in the context of worsening economic relations marked by the rise of tariff wars. Does this mean the agreement failed? Rather than make it a yes or no question, critics argue that the focus (and any ensuing warning against such actions) should be on the amount of damage done, not the precise lines that were crossed or how the violations were carried out. An analogy is telling the hosts of a drunken party that if the noise gets too loud, you will call the police. The objective is not the impossible one of stopping the music but the more practical one of lowering the volume to a more tolerable level.

There are other times when the United States will need to draw principled lines and defend them. It should acknowledge that it will continue to carry out intrusions in cyberspace for purposes it deems legitimate. And it will need to state precisely the norms and limits that Washington will up-hold--and call out countries that violate them. When China or Russia crosses a line, the United States will have to respond with targeted retaliation. This could involve public sanctions and also private actions, such as freezing the bank accounts of some oligarchs or releasing embarrassing information about them. U.S. Cyber Command's practices of defend forward and persistent engagement can be useful here, although they would best be accompanied by a process of quiet communication.

Treaties regarding cyberspace may be unworkable, but it might be possible to set limits on certain types of behavior and negotiate rough rules of the road. During the Cold War, informal norms governed the treatment of each side's spies; expulsion, rather than execution, became the norm. In 1972, the Soviet Union and the United States negotiated the Incidents at Sea Agreement to limit naval behavior that might lead to escalation. Today, China, Russia, and the United States might negotiate limits on their behavior regarding the extent and type of cyber-espionage they carry out, as Xi and Obama did in 2015. Or they might agree to set limits on their interventions in one another's domestic political processes. Although such pledges would lack the precise language of formal treaties, the three countries could independently make unilateral statements about areas of self-restraint and establish a consultative process to contain conflict. Ideological differences would make a detailed agreement difficult, but even greater ideological differences did not prevent agreements that helped avoid escalation during the Cold War. Prudence can sometimes be more important than ideology.

This seems to have been the approach explored by the Biden administration at a June summit with Russian President Vladimir Putin in Geneva, where cyberspace played a larger role on the agenda than nuclear weapons. According to press accounts, U.S. President Joe Biden handed Putin a list of 16 areas of critical infrastructure, including chemicals, communications, energy, financial services, health care, and information technology, that should be, in Biden's words, "off limits to attack, period." After the summit, Biden disclosed that he had asked Putin how he would feel if Russian pipelines were taken out by ransomware. "I pointed out to him that we have significant cyber-capability, and he knows it," Biden remarked at a press conference. "He does not know exactly what it is, but it is significant. And if in fact they violate these basic norms, we will respond with cyber. He knows." Thus far, however, it is unclear to what extent Biden's words have been effective.

One problem with specifying what needed to be protected might be that it implied that other areas were fair game--and that ransomware attacks from criminals in Russia would continue no matter what. In the cyber-realm, nonstate actors serve as state proxies to varying degrees, and rules should require their identification and limitation. And because the rules of the road will never be perfect, they must be accompanied by a consultative process that establishes a framework for warning and negotiating. Such a process, together with strong deterrent threats, is unlikely to fully stop Chinese and Russian interference, but if it reduces its frequency or intensity, it could enhance the defense of U.S. democracy against such cyberattacks.

CHANGING BEHAVIOR

In cyberspace, one size does not fit all. There may be some norms related to coordination that can accommodate both authoritarian and democratic states. But others cannot, such as the "Internet freedom" agenda introduced by U.S. Secretary of State Hillary Clinton in 2010. It proclaimed a free and open Internet. One can think of norms organized in a set of concentric circles with what Europeans call "variable geometry" of obligations. Groups of democracies can set a higher standard for themselves by agreeing on norms related to privacy, surveillance, and free expression and enforcing them through special trade agreements that would give preference to those that meet the higher standards, along the lines suggested by the cybersecurity expert Robert Knake. Such agreements could remain open to other states--so long as they are willing and able to meet the higher standards.

Diplomacy among democracies on these issues will not be easy, but it will be an important part of U.S. strategy. As James Miller and Robert Butler, two former senior Pentagon officials, have argued, "If U.S. allies and partners support cyber norms, they are likely to be more willing to support imposing costs on violators, thus substantially improving the credibility, severity (through multilateral cost imposition), and sustainability of U.S. threats to impose costs in response to violations."

The Biden administration is wrestling with the fact that the domain of cyberspace has created important new opportunities and vulnerabilities in world politics. Reorganizing and reengineering at home must be at the heart of the resulting strategy, but it also needs a strong international component based on deterrence and diplomacy. The diplomatic component must include alliances among democracies, capacity building in developing countries, and improved international institutions. Such a strategy must also include developing norms with the long-term goal of protecting the old glass house of American democracy from the new stones of the Internet age.

~~~~~~~~

By Joseph S. Nye, Jr.

JOSEPH S. NYE, JR., is University Distinguished Service Professor Emeritus at and former Dean of the Harvard Kennedy School. He is the author of Do Morals Matter? Presidents and Foreign Policy From FDR to Trump.
~~~~~~~~