A HUMAN RIGHTS RESPONSE TO GOVERNMENT HACKING

accessnow

SEPTEMBER 2016

accessnow

Access Now (www.accessnow.org) defends and extends the digital rights of users at risk around the world. By combining innovative policy, global advocacy, and direct technical support, we fight for open and secure communications for all.

For more information, please contact Amie Stepanovich at **amie@accessnow.org (PGP Fingerprint: CBBE4CF3 84B5FCA7 3BAAF3D0 FF726BC2 1C1DA0C7)** or visit our website **www.accessnow.org**.

This paper is an Access Now product. The primary author is Amie Stepanovich, with assistance from Daniel Bedoya-Arroyo, Gustaf Bjorksten, Michael Carbone, Drew Mitnick, Donna Wentworth, and Nathan White.

Access Now would like to thank Jochai Ben-Avie, Nate Cardozo, Andrew Crocker, Wafa Ben Hassine, Susan Landau, Christopher Parsons, Nick Selby, Ton Siedsma, Mohammad El Taher, Jamie Tomasello, and others for their input and comments on earlier drafts. This paper does not necessarily reflect their views. We would also like to thank our staff for their input, in particular Raman Jit Singh Chima, Josh Levy, Brett Solomon, and the Policy Team for their review and oversight.

TABLE OF CONTENTS

A HUMAN RIGHTS RESPONSE TO GOVERNMENT HACKING

EXECUTIVE SUMMARY

When governments engage in hacking it creates significant risks for human rights. However, there has yet to be an international public conversation on the scope, impact, or human rights safeguards for government hacking.

This paper raises the question of how human rights apply in the context of government hacking targeted at non-government and private sector actors. This includes government hacking that is perpetrated directly by the state, through a contractor or independent employee at the government's request or through government pressure, or otherwise sponsored by a state entity.

We start with a brief global tour of the history of government hacking and provide examples of potential government hacking activities. Hacking is the manipulation of software, data, a computer system, network, or other electronic device without the permission of the person or organization responsible for that software application, data, computer system, network, or electronic device, and/or without the permission or knowledge of users of that or other software, data, computers, networks, or devices ultimately affected by the manipulation. **Ultimately, more information is necessary to determine the full scope and impact of government hacking.** However, we also posit that there are three broad categories of government hacking that encompass current activities, so-divided based upon the objective to be accomplished: messaging control, causing damage, and commission of surveillance or intelligence gathering.

We then discuss both intended and unintended risks and consequences posed by government hacking, focusing on the significant ways that government hacking interferes with human rights as embodied in international treaties and declarations including rights to privacy, free expression, and due process. As a normative matter, we note, government hacking is particularly invasive and should be proscribed. However, we take further note that such hacking is already occurring and likely will become increasingly prevalent. Based upon international law and the broad human rights impacts, **we conclude that there should be a presumptive prohibition on all government hacking.**

Finally, we analyze the three identified categories of government hacking to determine whether our established presumption may be lawfully rebutted. In the first two categories — messaging control and causing damage — we determine that this presumption cannot be overcome. However, we find that, with robust protections, it may be possible, though still not necessarily advisable, for a government to overcome the presumptive prohibition in the third category, government hacking for surveillance or intelligence gathering. We note that the circumstances under which it could be overcome are both limited and exceptional, and we identify ten strong safeguards, including vulnerability disclosure and oversight, that must both be implemented and complied with to meet that standard. Absent government compliance with all ten safeguards, the presumptive prohibition on hacking remains.

In conclusion we reiterate that the human rights analysis is only one piece of the puzzle for government hacking, and the high threat that it poses to other interests may (and probably should) necessitate additional limitations and prohibitions.

I. INTRODUCTION

When governments engage in hacking it creates significant risks for human rights, and as we become more connected than ever before, these risks are becoming more pronounced. We are surrounded by digital devices — from computers and smartphones, to connected "things" on our wrists, in our pockets, or installed throughout our homes, all the way to our electric grid and utilities — and the number of devices we interact with continues to grow. Hacking operations can target any or all of these in order to gather information about us, exercise control over different aspects of our lives, or cause physical harm. Despite this risk, and the growing number of governments that are either engaging in hacking themselves or through a third party, there has yet to be an international public conversation on the scope, impact, or necessary safeguards for government hacking.

This paper raises the question of how human rights apply in the context of non-wartime government hacking targeted at non-government and private sector actors, an increasingly common activity that includes actions taken by or on behalf of a state. We save for another day the question of the human rights implications of government hacking in times of war or between governments; other publications have attempted to provide guidance on rules and obligations for government actors in these contexts.[1] However, we do look to these publications and other reports of government activities for lessons on the consequences of government hacking during peacetime.

In addition to the threat to human rights posed by government hacking, there are other risks involved with the activity, including the risk of damage to the property of internet users and the financial stability of private entities. Activities taken in pursuit of government hacking can also undermine global digital security, and by extension, global security as a whole.[2] Government hacking can also have other direct unexpected and unintended consequences.

Accordingly, there should be a presumptive prohibition on government hacking. In a subset of limited, exceptional cases, a government may be able to overcome this presumption, but it would require the safeguards discussed below, with which the government must comply in full.

In this paper we will first lay out a brief background of known hacking operations around the world. Then we will discuss the different types of government hacking and identify the examples of potential government hacking operations. We will explain the potential impacts and risks of government hacking in different circumstances and the interaction between government hacking and human rights. Finally, drawing on international law and broadly accepted standards, this paper sets out Ten Human Rights Safeguards for Government Hacking in certain instances, all of which must be adhered to for the government to overcome the presumptive prohibition on the activity. These safeguards include mechanisms for transparency, robust oversight, including public oversight, and access to remedy.

[1] *See*, *e.g.*, Tallinn Manual on the International Law Applicable to Cyber Warfare (Michael N. Schmitt, 2013), available at https://issuu.com/nato_ccd_coe/docs/tallinnmanual [hereinafter "Tallinn Manual"]; see also U.N. Secretary-General, Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, U.N. Doc. A/70/174 (July 22, 2015), https://www.accessnow.org/cms/assets/uploads/archive/UN_cyberspace_report.pdf.
[2] *See*, *e.g.*, Steven M. Bellovin, Matt Blaze, Sandy Clark, & Susan Landau, *Lawful Hacking: Using Existing Vulnerabilities for Wiretapping on the Internet*, Northwestern Journal of Technology and Intellectual Property (2014), available at http://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1209&context=njtip (for a discussion on the policy issues and security ramifications of government hacking).

TERMS OF ART

CONTENT-REWRITING PROXY

A server (a computer system or an application) that acts as an intermediary for requests from clients seeking resources from other servers, and as the content of that resource is passed back through the proxy it is modified before being delivered to the user that requested it.

DOMAIN NAME SYSTEM

A "system for naming computers and network services that is organized into a hierarchy of domains. DNS naming is used in TCP/IP networks, such as the Internet, to locate computers and services through user-friendly names."[3]

MALWARE

A type of computer program designed to infect a user's device or system and alter it.

METADATA

Non-content information about a communication, such as its duration, the participants, and their location.

"SOCK PUPPET ARMY" MECHANISM

A manual or automated system using a large number of fake user accounts to repeatedly present a particular point of view in the hope of swaying public opinion to that point of view.

VULNERABILITIES

A technical design or implementation flaw in information technology products or systems that could potentially be used to exploit or penetrate a product or system (hardware or software, to include open-source software).[4]

ZERO DAY EXPLOIT

An exploit that utilizes a vulnerability that is unknown to the developer of a product or service. So-named because the vendor of the software in which the vulnerability exists has had zero days to mitigate it.

ZERO-KNOWLEDGE PROOF

A cryptographic method by which one party (the prover) can prove to another party (the verifier) that a given statement is true, without conveying any information apart from the fact that the statement is indeed true.[5]

[3] DNS Defined, Microsoft TechNet, https://technet.microsoft.com/enus/library/bb629410.aspx (last visited July 29, 2016).

[4] Commercial and Government Information Technology and Industrial Control Product or System Vulnerabilities Equities Process, available at https://www.eff.org/document/vulnerabilities-equities-process-redactions (last visited July 29, 2016) [hereinafter, "U.S. Vulnerabilities Equities Process"].

[5] *See*, *e.g.*, Matthew Green, *Zero Knowledge Proofs: An Illustrated Primer*, A Few Thoughts on Cryptographic Engineering, Nov. 27, 2014, http://blog.cryptographyengineering.com/2014/11/zero-knowledge-proofs-illustrated-primer.html.

II. BACKGROUND

Though we have more information about government hacking than ever before, we still know startlingly little about the nature of or extent that any government actively engages in hacking. Here we'll take a global tour to decode what information on government hacking is available.

We know, for example, that the United States government has exploited vulnerabilities in computer systems for decades. The National Security Agency's (NSA) Office of Tailored Access Operations has been active since the late-1990s.[6] Similarly, the Federal Bureau of Investigation (FBI) has engaged in hacking operations since at least the early 2000s.[7]

A report in Newsweek last year explained:

> *According to the U.S. Intelligence Community's 2015 "Worldwide Threat Assessment" report, Russia and China are the "most sophisticated nation-state actors" in the new generation of cyberwarfare, and Russian hackers lead in terms of sophistication, programming power and inventiveness.*[8]

APT1, a group with ties to the Chinese government, is believed to have engaged in hacking activities since at least 2006.[9] Russia is suspected to have been behind sophisticated attacks against Estonia's government websites in 2007, shutting down user access to many online services.[10] Australia has broadly authorized government hacking since 1999.[11] The law was amended in 2014 to expand the reach of the government to conduct hacking activity in bulk.[12] In 2016, the Australian Prime Minister explained that his government's hacking capabilities were "very considerable."[13]

The German intelligence agency — known as the BND — has reportedly been engaged in government hacking since at least 2009.[14] The German police admitted to hacking in 2011.[15] In the United Kingdom,

[6] Kim Zetter, *NSA Hacker Chief Explains How to Keep Him Out of Your System*, Wired, Jan. 28, 2016, https://www.wired.com/2016/01/nsahackerchiefexplainshowtokeephimoutofyoursystem/; *see also,* Disrupting Nation State Hackers, Enigma, https://www.usenix.org/conference/enigma2016/conferenceprogram/presentation/joyce (last visited July 29, 2016).

[7] Kim Zetter, *Everything We Know About How the FBI Hacks People*, Wired, May 15, 2016, https://www.wired.com/2016/05/history-fbis-hacking/; see also, Operational Technology Division, Federal Bureau of Investigation, https://www2.fbi.gov/hq/otd/otd.htm (last visited July 29, 2016).

[8] Owen Matthews, *Russia's Greatest Weapon May Be its Hackers*, Newsweek, May 7, 2015, http://www.newsweek.com/2015/05/15/russias-greatest-weapon-may-be-its-hackers-328864.html.

[9] APT1: Exposing One of China's Cyber Espionage Units (Mandiant, 2013), available at https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf.

[10] Ian Traynor, *Russia Accused of Unleashing Cyberwar to Disable Estonia*, the Guardian, May 16, 2007, https://www.theguardian.com/world/2007/may/17/topstories3.russia.

[11] Australian Security Intelligence Organisation Legislation Amendment Act 1999 (No. 161), Parliament of Australia, *available at* http://www.austlii.edu.au/au/legis/cth/num_act/asiolaa1999664/sch1.html (last visited August 2, 2016), amending the Australian Security Intelligence Organisation Act 1979 § 25A, Parliament of Australia, *available at* http://www.austlii.edu.au/au/legis/cth/consol_act/asioa1979472/s25a.html (last visited July 29, 2016).

[12] National Security Legislation Amendment Bill (No. 1) 2014, Parliament of Australia, http://parlinfo.aph.gov.au/parlInfo/search/display/display.w3p;query%3DId%3A%22legislation%2Fbillhome%2Fs969%22;rec=0 (last visited July 29, 2016); see also, Michael Bradley, Five Questions About Our New National Security Laws, ABC News, Oct. 1, 2014, http://www.abc.net.au/news/2014-10-01/bradley-five-national-security-questions/5783142.

[13] *Australia Admits Government Hack Attacks, Boosts Cyber Security*, PhysOrg, Apr. 21, 2016, http://phys.org/news/2016-04-australia-hack-boosts-cyber.html.

[14] Pierluigi Paganini, *Trojan & Co, the New Frontiers of Espionage*, Security Affairs, Nov. 13, 2011, http://securityaffairs.co/wordpress/166/cyber-crime/trojan-co-the-new-frontiers-of-espionage.html.

[15] *Id.*

the Home Office officially acknowledged the use of its authorities to engage in hacking activity, dubbed Equipment Interference, in a 2015 Draft Code of Conduct, finalized in 2016.[16] Earlier activity by the UK has been documented by the press.[17] A news story in 2016 alluded to Italy's use of malware on a mobile phone to bypass encryption protections.[18] In addition, France passed a law in 2016 that broadly authorizes government hacking.[19]

In 2015, an anonymous activist compromised the servers of Hacking Team, a private company established in 2003 that sells tools and services to facilitate hacking. Internet emails and other documents published by the activist revealed that the company had been contracting with repressive governments since at least 2004.[20] Clients of Hacking Team have included government agencies or agents in Egypt, Italy, Korea, Turkey, Mexico, India, and Colombia, among others. These operations and others like them are particularly offensive in that they occur in the absence of a legal framework for the activity (and likely in violation of domestic law) and interfere with the human rights of people who have committed no crime, including political opponents, journalists, and activists.[21]

In at least one instance in 2014, several nations cooperated in an international hacking operation known as Operation Onymous, purportedly to identify individuals suspected of engaging in criminal activity.[22] A hacking program named Warrior Pride, which is operated jointly by the "Five Eyes" countries — the U.S., UK, Canada, Australia, and New Zealand — was revealed in the documents made available by Edward Snowden.[23]

Even when we have information about government hacking processes and procedures, any information about how they are used and whether they are effective isn't transparent.

For example, in 2014 the United States reinvigorated its Vulnerabilities Equities Process, or VEP — a process to determine whether to disclose vulnerabilities so they can be patched — after revelations surfaced suggesting that the NSA had been aware of the Heartbleed vulnerability but kept it secret, leaving it open to be exploited.[24] The VEP, details of which were revealed in heavily redacted form as a result of a Freedom of Information Act lawsuit by the Electronic Frontier

[16] Joint Representations by Access, the Center for Democracy & Technology, the Electronic Frontier Foundation, and New America's Open Technology Institute on "Interception of communications and equipment interference: draft codes of practice", Mar. 20, 2015, *available at* https://www.accessnow.org/cms/assets/uploads/archive/Joint%20GCHQ%20Representation.pdf; *see also*, Equipment Interference Code of Practice (draft for public consultation), Home Office (2015), *available at* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/401863/Draft_Equipment_Interference_Code_of_Practice.pdf; Equipment Interference Code of Practice, Home Office (2016), *available at* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/496069/53693_CoP_Equipment_Interference_Accessible.pdf.

[17] Katie Collins, *Anonymous and LulzSec Targeted by GCHQ DDoS Attacks*, Wired, Feb. 5, 2014, http://www.wired.co.uk/news/archive/2014-02/05/gchq-ddos-attack-anonymous.

[18] Sebastian Rotella, *ISIS via WhatsApp: 'Blow Yourself Up, O Lion'*, ProPublica, July 11, 2016, https://www.propublica.org/article/isis-via-whatsapp-blow-yourself-up-o-lion.

[19] ION 2016-731, Law strengthening the fight against organised crime, terrorism and their financing, and improving the efficiency and guarantees of criminal procedure (June 3, 2016), *available at* https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=FE4569C44829C0F79D4BF17709AAA7EF.tpdila18v_1?cidTexte=JORFTEXT000032627231&categorieLien=id.

[20] Philip Willan, *Former Hacking Team Developer Reportedly in Contact with a Terrorist Group*, PC World, July 31, 2015, http://www.pcworld.com/article/2955592/former-hacking-team-developer-reportedly-in-contact-with-a-terrorist-group.html.

[21] *See, e.g.*, Bill Marczak, Claudio Guarnieri, John Scott-Railton, and Morgan Marquis-Boire, *Hacking Team and the Targeting of Ethiopian Journalists* (Citizen Lab 2014), available at https://citizenlab.org/2014/02/hacking-team-targeting-ethiopian-journalists/.

[22] Andy Greenberg, *Global Web Crackdown Arrests 17, Seizes Hundreds of Dark Net Domains*, Wired, Nov. 7, 2014, https://www.wired.com/2014/11/operation-onymous-dark-web-arrests/.

[23] Jacob Appelbaum, Aaron Gibson, Claudio Guarnieri, Andy Müller-Maguhn, Laura Poitras, Marcel Rosenbach, Leif Ryge, Hilmar Schumndt, and Michael Sontheimer, *The Digital Arms Race: NSA Preps America for Future Battle*, Der Dpiegel Online, Jan. 17, 2015, http://www.spiegel.de/international/world/new-snowden-docs-indicate-scope-of-nsa-preparations-for-cyber-battle-a-1013409.html.

[24] Michael Daniel, *Heartbleed: Understanding When We Disclose Cyber Vulnerabilities*, the White House, Apr. 28, 2014, http://www.whitehouse.gov/blog/2014/04/28/heartbleed-understanding-when-we-disclose-cyber-vulnerabilities.

Foundation, appears to be largely mandatory for newly discovered or purchased vulnerabilities that are not publicly known.[25] However, the FBI was evidently able to sidestep the VEP in 2016, when the agency revealed that when it used an exploit it apparently leased to get data from the iPhone of one of the perpetrators of the San Bernardino attack, the agency did not submit the vulnerability to the process.[26] The agency appeared to exploit a loophole in the VEP by purchasing the rights to use the exploit, but never actually learning how the vulnerability worked.[27]

Moreover, government hacking operations are expanding. The market for exploits continues to grow, even as governments and companies seek to build legitimate reporting mechanisms.[28] The United Kingdom is currently finalizing a new surveillance law that would explicitly authorize not only hacking, but hacking in "bulk," despite urgent objections raised by several organizations.[29] The U.S. Supreme Court recently approved controversial updates to the Federal Rules of Criminal Procedure, which remove limits on law enforcement hacking, arguably blessing U.S. hacking operations, including those that target computers en masse located around the world. Under the updated rule, a single warrant could be used to target not only criminals but also potentially millions of victims of botnet exploitation.[30] The changes will automatically go into effect unless the U.S. Congress acts to withdraw or amend them before December 2016. Meanwhile, Kazakhstan recently mandated the installation of software on user devices to provide direct access to communications and services.[31] And, while Hacking Team has suffered a few small setbacks, there are several companies competing to take over as the premier hacking tool supplier to repressive nations around the world.[32]

III. WHAT IS GOVERNMENT HACKING?

The term "hacking" has had a number of different connotations throughout the history of its use. For the purposes of this paper, hacking means the manipulation of software, data, a computer system, network, or other electronic device without the permission of the person or organization responsible

[25] Vulnerabilities Equities Process, *supra* fn 4; see also Complaint for Injunctive Relief for Violation of the Freedom of Information Act, 5 U.S.C. § 552, Electronic Frontier Foundation v. National Security Agency, Office of the Director of National Intelligence (2014) (No. 3:14-cv-03010), *available at* https://www.eff.org/document/eff-v-nsa-odni-complaint.

[26] Russell Brandom, *the FBI Bought an iPhone Hack, But Not the Right to Tell Anyone How it Works*, the Verge, Apr. 27, 2016, http://www.theverge.com/2016/4/27/11518754/fbi-apple-iphone-hack-vulnerability-disclosure-vep.

[27] *See* Ellen Nakashima, *Comey Defends FBI's Purchase of iPhone Hacking Tool,* Washington Post (May 11, 2016), https://www.washingtonpost.com/world/nationalsecurity/comeydefendsfbispurchaseofiphonehackingtool/2016/05/11/ce7eae54161611e6924d838753295f9a_story.html.

[28] *See* Sebastian Anthony, *The First Rule of Zero-Days is No One Talks About Zero-Days (So We'll Explain)*, Ars Technica, Oct. 20, 2015, http://arstechnica.com/security/2015/10/the-rise-of-the-zero-day-market/; Multistakeholder Process: Cybersecurity Vulnerabilities, National Telecommunications & Information Administration (Apr. 8, 2016), https://www.ntia.doc.gov/other-publication/2016/multistakeholder-process-cybersecurity-vulnerabilities; Joseph Cox, *As Hackers Continue to Target Porn Sites, Pornhub Launches Bug Bounty Program*, Motherboard, May 10, 2016, https://motherboard.vice.com/read/pornhub-bug-bounty.

[29] Investigatory Powers Bill: Written Evidence Submitted by Access Now, Parliament of the United Kingdom, *available at* http://www.publications.parliament.uk/pa/cm201516/cmpublic/investigatorypowers/Memo/IPB72.htm (last visited July 29, 2016).

[30] *See* Brett Solomon, *This Arcane Rule Change Would Give U.S. Law Enforcement New Power to Hack People Worldwide*, Slate, May 11, 2016, http://www.slate.com/blogs/future_tense/2016/05/11/the_rule_41_change_would_give_u_s_law_enforcement_power_to_hack_people_worldwide.html.

[31] *See* Karl Bode, *Kazakhstan Decides to Break the Internet, Wage All Out War on Encryption*, TechDirt, Dec. 9, 2015, https://www.techdirt.com/articles/20151204/07412332986/kazakhstan-decides-to-break-internet-wage-all-out-war-encryption.shtml.

[32] *See*, Dan Goodin, Massive Leak Reveals Hacking Team's Most Private Moments in Messy Detail, Ars Technica, July 6, 2015, http://arstechnica.com/security/2015/07/massive-leak-reveals-hacking-teams-most-private-moments-in-messy-detail/.

for that software application, data, computer system, network, or electronic device, and/or without the permission or knowledge of users of that or other software, data, computers, networks, or devices ultimately affected by the manipulation.

It would be extremely difficult to list and discuss all of the different types of government hacking because the term encompasses a broad array of activities. For any specific goal, there are countless means and methods of achieving it. Therefore, instead of listing specific activities, this paper divides government hacking into three categories based on the broad goal to be achieved. These categories are:

1. **Messaging control -** control the message seen or heard, particularly by a specific target audience.

2. **Causing damage -** causing some degree of harm to one of any number of target entities.

3. **Commission of surveillance or intelligence gathering -** compromise the target in order to get information, particularly on an on-going basis.

The list below indicates only some of the possible scenarios that fall within these categories. It is meant to be illustrative and not exhaustive.

1. Messaging Control

a. **Preventing Message Dissemination** - Altering the network, or devices in the network, to prevent information from reaching an audience.
b. **Manipulation of the Domain Name System** - Creating copies of websites with a distinct message. This activity, dubbed a "fake domain attack," can be conducted through hijacking of the real website domain to point to an alternate page, using a similarly named domain and manipulating the tags to give it a preferential placement in search results, spoofing the domain name system to return a false result, or through other means.[33]
c. **Rewriting Content** - Introducing content-rewriting proxies on the network at key points to alter content in transit.
d. **Flooding Communications Channels** - Deployment of tools like an automated "sock puppet army" mechanism to repeat messages in forums, polls, or other places where conversation occurs on the internet, conveying a single point of view. This method makes it seem as if a large number of people support a specific idea.[34]
e. **Website Defacement** - Intentionally changing the content or presentation of another's website, typically through the exploitation of a vulnerability to "trick" the database access code into revealing administration passwords, giving the attacker access to modify or update the website.

2. Causing Damage

a. **Modifying Physical Systems or Devices Internally** - A rare technique utilized by exerting control, including through the insertion of malware, to change the operation of hardware or software,

[33] *See, e.g.*, Michael Carbone, *One of These Things is Not Like the Other: Report on Fake Domains Attacks on Civil Society Released*, Access Now, Aug. 1, 2013, https://www.accessnow.org/one-of-these-things-is-not-like-the-other-report-on-fake-domains-attacks-on/.
[34] *See, e.g.*, Jordan Robertson, Michael Riley, and Andrew Willis, *How to Hack an Election*, Bloomberg, Mar. 31, 2016, http://www.bloomberg.com/features/2016-how-to-hack-an-election/.

often in a way that inflicts physical harm upon the system, up to and including destruction of the system or device. The goal could be to influence entire systems (e.g., a national power grid) or specific devices.

b. **Modifying Physical Systems or Devices Externally** - Similar to the above, exerting control to change the operation of a system or device, but in order to achieve a result external to the system. For example, by modifying a weapons system to miss its intended target.[35]

c. **Modification of Data** - Accessing a file system or database in order to add, delete, or modify data. This could include data that implicates an individual in a crime or other unsavory activity (or deletes data about an individual's involvement therein).

d. **Denial of Service** - Activity to prevent a target from being able to carry out some particular activity, including by forcing it to engage exhaustively in some other activity.

3. Commission of Surveillance or Intelligence Gathering

a. **Endpoint or Host Compromise** - Implemented directly with zero-day or known exploits, via trojans or malware, or by crafting malicious resources and tricking the target to visit that resource. Can be used to steal stored data or facilitate on-going surveillance.

b. **Monitoring Communications Channels** - Gaining access to channels that people use to communicate with one another in order to eavesdrop on the content of messages or on communications data identifying the participants in a communication or other metadata. This may include redirecting traffic to ensure that it is routed through a control vector. This strategy may also be used to uncover the identities and locations of targets.[36]

c. **Cracking Encryption** - Compromising the confidentiality, integrity, authentication, non-repudiation, and zero-knowledge proof properties of encryption systems to enable access to content or assist with additional hacking. This activity could also belong under "messaging control" if it attacks the authentication properties of encryption, which can call into question the identity of the people or companies in communication with one another.

All government hacking substantially interferes with human rights. While in many ways this interference may be similar to more traditional government activity, the nature of hacking creates new threats to human rights that are greater in both scale and scope. Hacking can provide access to private information, both be it stored or in transit, or even, such as with a keystroke logger, while it is being created or drafted. Exploits and malware used in operations can act unpredictably, damaging hardware or software or infecting non-targets and compromising their information. Even when a particular hack is narrowly designed, it can have unexpected and unforeseen impact.

There are also other interests threatened by government hacking. All of the harms discussed below can extend to users beyond the specific target. When released into the wild, some hacking tools can proliferate in either form or function, spreading broadly to other devices and networks. These tools are nearly impossible to control, and can impact individuals or groups who are in contact with a target, as

[35] *See, e.g., For God and My President: State Surveillance in Uganda* (Privacy International, 2015), *available at* https://privacyinternational.org/sites/default/files/Uganda_Report.pdf ("Intrusion technologies are capable of collecting, modifying and extracting data communicated and stored on a device. To do this, malware must be installed on the device. Once installed, it embeds itself in all system functions, collecting and transmitting data to the operator as the infected device operates normally from the user's perspective.")

[36] *See, e.g.*, Dan Goodin, *Tunisia Plants Country-Wide Keystroke Logger on Facebook* (Google and Yahoo! too), the Register, Jan. 25, 2011, http://www.theregister.co.uk/2011/01/25/tunisia_facebook_password_slurping/.

well as those who are totally unrelated. Users of shared computers or systems, like those in a library or office, are at increased risk of incidental infection. The Stuxnet case is perhaps the best-known instance of a developer losing control of a piece of malware. Stuxnet was a worm likely created by the Israeli and U.S. governments to infect Iranian nuclear facilities. However, the worm spread far beyond what was originally intended, and Stuxnet was eventually found on more than 100,000 machines belonging to people all around the world.[37]

IV. HARMS OF GOVERNMENT HACKING

Government hacking specifically interferes with human rights. Along with the International Covenant on Civil and Political Rights, the European Declaration of Human Rights, and other international documents, the Universal Declaration on Human Rights ("UDHR") sets out the human rights possessed by individuals. The UDHR was adopted by the UN General Assembly in 1948 and has received broad support from governments around the globe. Below we identify some of the rights that are identified in the UDHR which are most clearly implicated by government hacking:

Article 10: "Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal, in the determination of his rights and obligations and of any criminal charge against him."

Due process rights are central to ensuring fairness and the protection of other human rights. However, because hacking often takes place remotely and surreptitiously, there is no inherent notice to the target of the activity, which makes judicial challenges difficult to pursue. Therefore it is important for governments to ensure that notice is provided to those who are impacted by government hacking activities. Furthermore, the tools used for government hacking are complicated and hard to understand for individuals without specific technology training, which may include the judges or government officials authorizing or overseeing hacking (in cases where it is authorized and overseen). The tools can therefore often be approved in error.

Article 12: "No one shall be subjected to arbitrary interference with his privacy, family, home[,] or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to protection of the law against such interference or attacks."

All government hacking that facilitates access to "Protected Information" interferes with the right to privacy. Protected Information is "information that includes, reflects, arises from, or is about a person's communications and that is not readily available and easily accessible to the general public." This includes private information and public information that is aggregated or analyzed in a way that elucidates on non-public information. While only one of the categories of government hacking we identify occurs specifically to facilitate surveillance, nearly all government hacking facilitates or provides access to the Protected Information of a target. As such, the right to privacy is perhaps the right most directly interfered with by government hacking.

[37] *See* Kim Zetter, *How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History*, Wired, July 11, 2011, https://www.wired.com/2011/07/howdigitaldetectivesdecipheredstuxnet/.

Article 17: "(1) Everyone has the right to own property alone as well as in association with others. (2) No one shall be arbitrarily deprived of his property."

In some cases, government hacking may intentionally seek to exercise domain over private property in some way, including in order to do some harm to that property. Some examples of this are identified in category three, discussed above. However, even when it is not the goal, government hacking often has direct deleterious effects on user devices and networks. As noted technologist Steven Bellovin has said, "when you hack a system, you don't actually know what's going to happen."[38] This is true any time you interfere with a system's operation, even when it is authorized. In March 2016, a system update by Apple for iPads left many devices totally non-functional — what is known as "bricking" the device.[39] This meant that users lost all of their information that wasn't backed up elsewhere. Apple had most definitely tested the update thoroughly and had specifically developed the software that was being updated, but they missed something, and that had disastrous consequences for many users. Hacking operations are even more unpredictable.[40] For example, during an operation to install an exploit that would further its surveillance abilities, the U.S. NSA allegedly ended up blacking out the internet across Syria.[41]

Article 18: "Everyone has the right to freedom of thought, conscience[,] and religion..."

Article 19: "Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive[,] and impart information and ideas through any media and regardless of frontiers."

Article 20: "(1) Everyone has the right to freedom of peaceful assembly and association. (2) No one may be compelled to belong to an association."

Government hacking can have both direct and indirect impacts on rights to thought, expression, and association. Government hacking in category two, as identified above, can directly quell public speech and dissent through hiding or obscuring ideas that the government wants to repress. Like other forms of surveillance, widespread government hacking may also chill speech more broadly, particularly speech of activists, writers, and journalists.[42] Government hacking can also limit or totally block publication of or access to information or online forums, either specifically by shutting down websites, blocking content, deleting data, or bricking a device used for access.

[38] Hacking America, New America, *available at* https://www.newamerica.org/oti/events/hacking-america/ (last visited July 29, 2016).
[39] See Shaun Nichols, *iPad Bricked by iOS 9.3? Don't Worry, We'll Get Through This Together*, the Register, Mar. 24, 2016, http://www.theregister.co.uk/2016/03/24/ipad_reader_stories/.
[40] *See, e.g.*, Mark Raymod, Greg Nojeim, and Alan Brill, *Private Sector Hack-Backs and the Law of Unintended Consequences*, Center for Democracy & Technology (Dec. 15, 2015), https://cdt.org/insight/private-sector-hack-backs-and-the-law-of-unintended-consequences/.
[41] *See, e.g.*, Spencer Ackerman, *Snowden: NSA Accidentally Caused Syria's Internet Blackout in 2012*, the Guardian, Aug. 13, 2014, https://www.theguardian.com/world/2014/aug/13/snowden-nsa-syria-internet-outage-civil-war; *see also* Peter Micek, *US Must Remedy NSA's 2012 Syrian Internet Shutdown*, Access Now, Aug. 15, 2014, https://www.accessnow.org/us-must-remedy-nsas-2012-syrian-internet-shutdown1/.
[42] *See* Elizabeth Stoycheff, *Under Surveillance: Examining Facebook's Spiral of Silence Effects in the Wake of NSA Internet Monitoring*, Journalism & Mass Communication Quarterly (2016), http://jmq.sagepub.com/content/early/2016/02/25/1077699016630255.full.pdf?ijkey=1jxrYu4cQPtA6&keytype=ref&siteid=spjmq; *Chilling Effects: NSA Surveillance Drives U.S.* Writers to Self Censor, PEN America, *available at* https://pen.org/chillingeffects (last visited May 27,2016).

Threats to human rights are not the only threats of government hacking. Government hacking should also be weighted against the following types of harm, which are outside the purview of this report:

Financial Harms

Several of the above activities can have a negative financial impact on the target, including through a direct modification of debts or assets, causing expenditure on recovery or legal remedy, loss of business or customers, or the cost of time and energy (including lost resources) in finding a new forum or platform.

Property Harms

Some hacking causes direct harm to devices or software, which can limit or cut off operability. In cases where data are stored on a device that is "bricked" — or rendered inoperable — by hacking, that data could be permanently lost. Replacement of devices and efforts to recover data may also be expensive, adding to the financial harm.

Reputational Harms

Certain types of hacking can harm the image of a target, either with a specific audience or the general public. Reputational harm could occur due to several reasons, including through the impression that someone said or did something that they did not or the presumption that a target was uniquely unable to resist the attack and therefore engages in inadequate security practices.

Digital Security Harms

The need to create and maintain offensive capabilities necessary for effective hacking operations can undermine global digital security. It can contribute to under-reporting of vulnerabilities, and therefore less patching of those security weaknesses. The potential for vulnerabilities to be inserted indiscriminately in software updates or directly into hardware or software,[43] or introduced into internet infrastructure, can also undermine user trust in the internet, which can have a major impact on global communication and the digital economy. The stockpiling of vulnerabilities, by not only black market actors but also governments, has increased the market prices for those vulnerabilities, making it hard for bug bounty programs to compete.[44]

Causal Harms

Government hacking also causes incidental harms, not intended or anticipated as part of the hack but nevertheless directly caused by it. There are several types of incidental harm. For example, certain types of hacks could leave both the target and others open to further attacks by creating new vulnerabilities that could be exploited by other actors, both to get to the original target or, potentially, the target's connections. Alternatively, as we saw with Stuxnet, the discovery of certain types of

[43] *See, e.g.*, Kimberlee Morrison, *Tor is Vulnerable to Malware and Government Surveillance*, SocialTimes, Nov. 10, 2014, http://www.adweek.com/socialtimes/tor-malware-government-surveillance/207544.

[44] *See, e.g.,* Katie Moussouris, *The Wolves of Vuln Street The First System Dynamics Model of the 0day Market,* HackerOne, Apr. 14, 2015, https://hackerone.com/blog/thewolvesofvulnstreet; Pierluigi Paganini, *Zeroday market, the governments are the main buyers,* Security Affairs, May 21, 2013, http://securityaffairs.co/wordpress/14561/malware/zerodaymarketgovernmentsmainbuyers.html.

hacking strategies can lead to copycat efforts, which can further negatively impact individuals.

All of these harms should be considered before any government hacking operations are authorized. From a normative perspective, the serious interference of government hacking with human rights partnered with the significant risk of additional harm should strongly caution against government hacking and highly suggests that such activity should be proscribed. However, as discussed, government hacking is already occurring around the world, and at an increasing rate. Accordingly, it is important to discuss its legal status in regard to human rights. In the next section we focus on the application of human rights law and policy to the three identified categories of government hacking.

V. GOVERNMENT HACKING AND HUMAN RIGHTS

As explained above, there are at least three categories of government hacking based on the desired goal, and countless ways to achieve those goals. In addition, there may also be a broad array of rationales invoked by governments to justify their use of hacking rather than other means to accomplish a desired outcome. For example, in many cases governments may claim to employ hacking because it can be the easiest or most efficient means to achieve a desired outcome or because it is covert. In other cases, a government may argue that hacking is necessary to bypass encryption protections that prevent access to certain types of information related to an investigation. When that is done on a targeted basis, it may be a less intrusive way to conduct certain types of surveillance without broadly undermining the integrity of the entire internet. However, even in cases where governments may cite the benefits of government-sponsored hacking, the activity can (and will) result in other harms, as set out above.

In order to fully understand the implications of government hacking, **it is necessary that more information about the nature and extent of current hacking operations around the world be made public.** The public requires more transparency regarding how governments decide to employ hacking and how and when hacking activity has had unanticipated impacts.

As noted above, all forms of government-sponsored hacking interfere with human rights. In 2011, Frank La Rue, then the United Nations Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, issued a report that stated, in part, "When a cyber-attack can be attributed to the State, it clearly constitutes...a violation of its obligation to respect the right to freedom of opinion and expression."[45] International human rights instruments, like the International Covenant on Civil and Political Rights, the European Declaration of Human Rights, and the Universal Declaration of Human Rights, guarantee these rights, as well as rights to privacy, association, and due process, amongst others.

The significant interference with human rights caused by government hacking necessitates a presumptive prohibition on the activity. However, while in most cases government hacking is irreconcilable with human rights, in others it may be possible for governments to overcome this presumption. Below we analyze the human rights impact of the three categories of government

[45] Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Frank La Rue, Report of the Special Rapporteur to the Human Rights Council on key trends and challenges to the right of all individuals to seek, receive and impart information and ideas of all kinds through the Internet, *UN Doc. A/HRC/17/27*, May 16, 2011, ¶ 52, *available at* http://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf.

hacking identified above and discuss what safeguards are necessary to protect human rights.

1. Messaging Control

Hacking for messaging control is a direct affront to the human rights to freedom of expression, opinion, religion, and expression. The European Court of Human Rights (ECtHR) has held that the human rights protections in the European Convention for the Protection of Human Rights and Fundamental Freedoms "basically [prohibit] a Government from restricting a person from receiving information that others wish or may be willing to impart to him."[46]

According to well-established international law, any measure by government to restrict these rights must be provided for by law, serve a legitimate aim, and must be necessary and proportionate to achieving that aim.[47] This means that these measures must be the least restrictive means for accomplishing a government's legitimate priority.[48]

Messaging control through hacking cannot meet this standard. Government hacking in this context can limit or prevent an individual, group, or entire population from accessing and disseminating information, or can alter that information to change its content without notice to either the senders or the recipients of the communications. In fact, the entire purpose of this type of hacking requires that parties are unaware of the government's intervention. Government attempts to control the dissemination of information in this manner are tantamount to censorship and represent the most nefarious type of prior restraint.

Perhaps the most well-known form of message control traditionally has been the development and use of propaganda. In 1983, the United Nations Office for the High Commissioner for Human Rights called for a total prohibition on "propaganda for war and any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence."[49] The rules for other forms of propaganda are less well defined.[50] However, it is very clear that "propaganda, when it is pervasive, massive[,] and systematic, is detrimental to the freedom of the media," along with other human rights.[51]

Too frequently, messaging control conducted through government hacking is pervasive, massive, and systematic. Employing hacking to modify, hide, or delete a message can have widespread effects, influencing national or international dialogues. Even when targeted at a single person, covert government hacking to dictate a message can stifle dissent, manipulate individual thoughts, and create global chilling effects on speech. Finally, government hacking for messaging control, if permitted, has the potential to set off a global messaging war which could substantially undermine cross-border communication.[52]

[46] Toby Mendel, *Freedom of Information as an Internationally Protected Human Right*, Privacy International, *available at* https://www.article19.org/data/files/pdfs/publications/foiasaninternationalright.pdf.

[47] UN Human Rights Committee (HRC), General Comment No. 34, Article 19, Freedoms of Opinion and Expression, Sept. 12, 2011, CCPR/C/GC/34, available at, http://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf.

[48] *See*, *supra* fn 44.

[49] UN Office of the High Commissioner for Human Rights (OHCHR), General Comment No. 11 (CCPR), Prohibition of Propaganda for War and Inciting National, Racial or Religious Hatred, May 10, 1999, *available at* http://www.ohchr.org/Documents/Issues/Opinion/CCPRGeneralCommentNo11.pdf.

[50] *See, e.g., Propaganda and Freedom of the Media,* Organization for Security and CoOperation in Europe, The Representative on Freedom of the Media (2015), *available at* http://www.osce.org/fom/203926?download=true.

[51] *Id.*

[52] *See, e.g.*, id. ("dangers of propaganda become a useful excuse for governments to restrict or even ban all hostile messages, actual and potential, coming from abroad.").

For these reasons, government hacking in this category is inconsistent with human rights law and policy, and should be fully prohibited by law.

2. Causing Damage

Government hacking to cause damage is perhaps the most invasive form of government hacking. While, as we explain above, hacking is unpredictable and can often result in unintentional harm to systems or devices, this type of government hacking sets out to instigate that outcome.

Government hacking that falls under this umbrella is often designed specifically to deprive a person of their property in some way. This implicates due process protections, which require a fair trial overseen by a competent judicial authority, qualified legal representation, and the ability to appeal. It also directly conflicts with the right recognized in most countries for individuals to own private property. When the damage a government seeks to carry out also implicates human life or well-being, the threat to human rights is exceptionally grave. Government hacking to do damage also implicates other human rights, such as freedom of expression and association, since these rights are frequently exercised using devices that such hacking could render inoperable.

The legal concept that might be most closely applicable in analyzing government hacking in this category is the doctrine of eminent domain, where a government may claim control over private property. Eminent domain has been held as consistent with human rights when "the interference [serves] a public interest, is proportionate, and [is] authorized by law."[53] In addition, such taking must be compensated.[54]

In the context of hacking, that standard has so far been impossible to meet. The individual interests implicated are significant. A person could be deprived of the ability to communicate, could be implicated in a crime, or, in extreme cases, could have his or her life put in jeopardy. Additionally, because hacking tools have unexpected behavior, these risks could apply very broadly. Hacking that implicates internationally distributed or used hardware or software, or that impacts internet infrastructure, would undoubtedly have a global impact. Finally, it may be impossible to compensate for the potential losses, and the emotional impacts thereof, that would likely result from government hacking in this category, which could destroy, alter, or render permanently inaccessible property such as priceless documentation or information, belonging to targets and non-targets alike. Yet the public interest served through such hacking is theoretical at best. No concrete, compelling case has ever been made public that would require this type of hacking in order to serve an established public interest.

Some day technology may evolve in a way that alters this analysis. If a case is eventually made for hacking in the public interest, it must first be authorized by law, and made subject to robust public debate, to weigh the substantial threat the activity poses to human rights. The law will have to be limited to allowing the least intrusive possible means to strictly achieve the societal needs identified, which must be significant, and incorporate additional human rights protections. However, unless or

[53] *Taking Property for the Public Good: Eminent Domain Laws From Around the World* (N.Y. Law School International Review 2012) at 18, *available at* http://www.nyls.edu/documents/center_for_international_law/the_international_review_newsletter/cil_newsletter_springsummer2012.pdf.
[54] *Id.*

until such a case can be made, government hacking to cause damage must be explicitly prohibited.

3. Commission of Surveillance or Intelligence Gathering

The final category of government hacking is hacking for the purpose of surveillance or intelligence gathering. Government surveillance directly interferes with the human right to privacy. As the International Principles on the Application of Human Rights to Communications Surveillance ("the Principles") say, "privacy is a fundamental human right, and is central to the maintenance of democratic societies. It is essential to human dignity and it reinforces other rights, such as freedom of expression and information, and freedom of association, and is recognized under international human rights law."[55]

Any restrictions on rights to privacy and expression are subject to a "permissible limitations" test.[56] Pursuant to UN Human Rights Committee General Comment Number 34, such "permissible" restrictions must be provided by law; strictly serve a legitimate aim (respect of the rights and reputation of others, protection of national security or of public order, or of public morals or health); and meet a high standard of legality, proportionality, and necessity.

Government hacking in the context of surveillance is often more invasive than other forms of surveillance, and activities taken in pursuit thereof could grant nearly unfettered access to some of a person's most personal information, limited only by the imagination of the hacker and the design of the exploit. Traditionally, the incidents of government surveillance increase as the ability to conduct surveillance gets cheaper and easier.[57] Government hacking may greatly reduce the cost of surveillance and lowers certain barriers to surveillance because it can take place remotely.

Because of these significant considerations, government hacking for surveillance, like the other two categories, should be subject to a presumptive prohibition. However, a close analysis of human rights law and standards for government surveillance demonstrates that there may be instances when the government could overcome this presumption. To do so requires significant safeguards, both promulgated and adhered to. These safeguards must apply equally to government hacking that is perpetrated directly by the state, conducted through a contractor or independent employee at the government's request, compelled by the government, or takes place with state sponsorship.

The Principles provide a framework for the application of the standards in the United Nations Human Rights Committee's General Comment 34 and other international law.[58] From the Principles we can derive Ten Human Rights Safeguards for Government Hacking. While several of the Principles apply directly to the issue of government hacking, in some cases they should be applied even more stringently due to hacking's increased interference with human rights.

While the Principles, by their name, were drafted in the context of communications surveillance — broadly defined to include "the monitoring, intercepting, collecting, obtaining, analysing, using, preserving, retaining,

[55] International Principles on the Application of Human Rights to Communications Surveillance (May 2014), https://necessaryandproportionate.org/text [hereinafter "Necessary and Proportionate Principles"]; *see also, inter alia,* Universal Declaration of Human Rights, art. 12; UN Convention on Migrant Workers, art. 14, UN Convention of the Protection of the Child, art. 16.
[56] Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, U.N. Doc. A/HRC/23/40 ¶¶ 28, 29 (Apr. 17, 2013) (by Frank La Rue).
[57] *See, e.g.*, Kevin S. Bankston and Ashkan Soltani, *Tiny Constables and the Cost of Surveillance: Making Cents Out of United States v. Jones*, Yale L.J. (2014), *available at* http://www.yalelawjournal.org/forum/tiny-constables-and-the-cost-of-surveillance-making-cents-out-of-united-states-v-jones.
[58] Necessary and Proportionate Principles, *supra* fn. 54.

interfering with, accessing or similar actions taken with regard to information that includes, reflects, arises from or is about a person's communications in the past, present, or future" — they apply to all Protected Information.[59] As explained above, that is "information that includes, reflects, arises from, or is about a person's communications and that is not readily available and easily accessible to the general public," which would comprise not just communications but other information, such as location.

Finally, it is not enough to superficially incorporate these safeguards into a country's law or policy if governments do not comply with them in all circumstances. This means that it may also be necessary to legislate penalties for the failure to meaningfully adhere to these safeguards.

Legality / Legitimate Aim

Government hacking operations must be foreseeable by those who may be impacted by them. Therefore, authorization for government hacking must be specifically provided for by law, clearly written, and publicly available. The law should prohibit government hacking except in specific, limited circumstances. The information sought through government hacking should be defined with particularity in advance. Hacking should never be performed with either discriminatory purpose or effect.

Safeguard 1: Government hacking must be provided for by law, which is both clearly written and publicly available and which specifies the narrow circumstances in which it could be authorized. Government hacking must never occur with either a discriminatory purpose or effect.

Necessity / Adequacy / Proportionality

Hacking operations cannot be justified unless they are the least invasive legal means to get specified Protected Information. An application for hacking should specify: (1) the circumstances that make hacking necessary, (2) exactly what tool or means that the government plans to use to complete the operation, and (3) where (on what device) the government plans to use them. Applications must be time-limited and the use of hacking activity should never occur in perpetuity or without a set end date. The tool or means must be designed to return only the limited categories of information that are considered necessary about specified limited targets and if, in the course of the operation, any extraneous Protected Information is collected, it should be immediately purged. Bulk or indiscriminate hacking, or hacking that implicates the infrastructure of the internet, should not be authorized.[60] If there are categories of information the government is seeking in an operation that could be acquired through government hacking but may also be acquired by other means, then the government should pursue those other means, which may be done in tandem with the hacking operation.

Safeguard 2: Government actors must be able to clearly explain why hacking is the least invasive means for getting Protected Information in any case where it is to be authorized and must connect that necessity back to one of the statutory purposes provided. The necessity should be demonstrated for every type of Protected Information that is sought, which must be identified, and every user (and device) targeted. Indiscriminate, or mass, hacking must be prohibited.

Safeguard 3: Government hacking operations must never occur in perpetuity. Authorizations for government hacking must include a plan for concluding the operation. Government hacking

[59] *Id.*

[60] *See, e.g.*, Tallinn Manual, *supra* fn 1 at Rules 43, 49, and 51.

operations must be narrowly designed to return only specific types of authorized information from specific targets and to not affect non-target users or broad categories of users. Protected Information returned outside of that for which hacking was necessary should be purged immediately.

Competent Judicial Authority / Due Process

Applications to conduct hacking operations must be sufficiently detailed and approved by an independent judicial authority who has been adequately educated, to the extent possible, on the potential technological ramifications of the tool or the means being used, and any risks of unintended consequences. Courts must be adequately equipped to supervise these operations, which may be more technologically complex than other forms of surveillance previously authorized. The approval of the application should in all possible cases happen in an adversarial process, with parties able to argue on both sides of the issue, but in any case should include, at a minimum, an independent technical expert who can review the government's claims and tools and provide any additional information that is necessary for the judicial authority to understand the application and the risks that it poses.[61] Because government hacking operations may, to some extent, deny users of their property, the rights to due process require that all hacking operations, even in the case of an emergency, must be authorized by a judicial authority according to these safeguards.

Safeguard 4: Applications for government hacking must be sufficiently detailed and approved by a competent judicial authority who is legally and practically independent from the entity requesting the authorization and who has access to sufficient technical expertise to understand the full nature of the application and any likely collateral damage that may result. Hacking should never occur prior to authorization.

User Notification / Transparency / Public Oversight

Most hacking operations lack the inherent notice available with physical searches of devices. Furthermore, even when notice is required in national law, many countries withhold that notice indefinitely. However, the increased risk of harm to the device from the search makes notice much more important.

Specific applications for government hacking should be filed publicly and include details about the activity authorized in order to facilitate public conversation about the use of the hacking tools and activities, though it may not be necessary to disclose the exact nature of the tool or technique used. In an active investigation, government should provide as much detail as possible. This notice may be delayed, but such delay must be specifically limited and cannot continue in perpetuity, even if charges are never filed. Additionally, government actors who engage in hacking operations should monitor the effects of the hacking tools to the extent possible and publicly report any unexpected or unwarranted activity that occurs as a result of their use.

Safeguard 5: Government hacking must always provide actual notice to the target of the operation and, when practicable, also to all owners of devices or networks directly impacted by the tool or technique.

[61] *See*, *e.g.*, Amie Stepanovich, *The USA FREEDOM Act of 2015: What's In it?*, Access Now, Apr. 29, 2015, https://www.accessnow.org/the-usa-freedom-act-of-2015-whats-in-it/ ("However, the bill would establish "friends of the court," who could be called upon to provide expertise on the impact of surveillance on privacy, the technical implications of new methods and programs, and other specialized areas of knowledge.").

Safeguard 6: Agencies conducting government hacking should publish at least annually reports that indicate the extent of government hacking operations, including at a minimum the users impacted, the devices impacted, the length of the operations, and any unexpected consequences of the operation.

Integrity of Communications and Systems

Private entities should never be compelled to assist governments in operations to hack into their own products and services in ways that undermine user security. This includes compulsion, either explicit or otherwise, to adopt tools or technical standards to make it easier for governments to conduct hacking operations.

Safeguard 7: Government hacking operations must never compel private entities to engage in activity that impacts their own products and services with the intention of undermining digital security.

Safeguards for International Cooperation / Safeguards Against Illegitimate Access and Right to Effective Remedy

If, in pursuit of a hacking operation, Protected Information is yielded outside the scope of the authorization, the reason for the excess information should be studied and justification should be provided to the competent judicial authority, including measures that will be taken to ensure that the tool or technique used will not return unauthorized information in the future. Where avoidable and in line with these safeguards, including safeguards on necessity, extraterritorial government hacking should never occur unless lawfully authorized under principles of dual criminality.

Because unpatched vulnerabilities needlessly perpetuate global risks to users, vulnerabilities discovered or received by a government should be promptly disclosed to the developer. Delay in the disclosure of a vulnerability should be time-limited and extraordinary, only permitted where immediate disclosure would directly undermine the rights of users. Routine public reports should identify the exact number of vulnerabilities that are withheld along with the justification for the withholding.

Safeguard 8: If a government hacking operation exceeds the scope of its authorization, the agency in charge of the authorization should report back to the judicial authority the extent and reason.

Safeguard 9: Extraterritorial government hacking should not occur absent authorization under principles of dual criminality.

Safeguard 10: Agencies conducting government hacking should not stock vulnerabilities and, instead, should disclose vulnerabilities either discovered or purchased unless circumstances weigh heavily against disclosure. Governments should release reports at least annually on the acquisition and disclosure of vulnerabilities.

VI. CONCLUSION

Government hacking poses a great risk to human rights, and as a normative matter, should be proscribed. Most types of government-sponsored hacking are inherently inconsistent with human rights protections.

However, governments are currently engaged in hacking operations, and it is occuring without a robust public conversation on its risks and without transparency, rules, or oversight. The risks posed by government hacking are amplified when it is conducted in the dark or without human rights protections for users.

Under international law, government hacking substantially interferes with human rights and should be presumptively prohibited. In the limited cases where a government can overcome that presumption, soley for the purposes of surveillance or intelligence-gathering, there are Ten Human Rights Safeguards for Government Hacking that must be in place, including a clear, transparent framework that includes mechanisms for robust oversight, including public oversight, and access to remedy. Those safeguards must be complied with in every instance, and there must be accountability mechanisms in place and remedy for individuals impacted by the activity.

While the Ten Human Rights Safeguards for Government Hacking may ameliorate the human rights threats of government hacking, they do not address all potential harms that could be caused by the activity. It is important to consider all of the interests and costs of government hacking prior to implementing a law to authorize its use, and safeguards even beyond the expansive ones identified here may be necessary. Absent the full implementation of the Ten Human Rights Safeguards, the presumptive prohibition remains in all instances of government hacking.

Appendix: TEN HUMAN RIGHTS SAFEGUARDS FOR GOVERNMENT HACKING

There should be a presumptive prohibition on all government hacking. In any instance where government hacking is for purposes of surveillance or intelligence-gathering, the following ten safeguards must all be in place and actually complied with in order for a government to successfully rebut that presumption. Government hacking for the purposes of messaging control or causing damage cannot overcome this presumption.

1. Government hacking must be provided for by law, which is both clearly written and publicly available and which specifies the narrow circumstances in which it could be authorized. Government hacking must never occur with either a discriminatory purpose or effect;
2. Government actors must be able to clearly explain why hacking is the least invasive means for getting Protected Information in any case where it is to be authorized and must connect that necessity back to one of the statutory purposes provided. The necessity should be demonstrated for every type of Protected Information that is sought, which must be identified, and every user (and device) targeted. Indiscriminate, or mass, hacking must be prohibited;
3. Government hacking operations must never occur in perpetuity. Authorizations for government hacking must include a plan for concluding the operation. Government hacking operations must be narrowly designed to return only specific types of authorized information from specific targets and to not affect non-target users or broad categories of users. Protected Information returned outside of that for which hacking was necessary should be purged immediately;
4. Applications for government hacking must be sufficiently detailed and approved by a competent judicial authority who is legally and practically independent from the entity requesting the authorization and who has access to sufficient technical expertise to understand the full nature of the application and any likely collateral damage that may result. Hacking should never occur prior to authorization;
5. Government hacking must always provide actual notice to the target of the operation and, when practicable, also to all owners of devices or networks directly impacted by the tool or technique;
6. Agencies conducting government hacking should publish at least annually reports that indicate the extent of government hacking operations, including at a minimum the users impacted, the devices impacted, the length of the operations, and any unexpected consequences of the operation;
7. Government hacking operations must never compel private entities to engage in activity that impacts their own products and services with the intention of undermining digital security;
8. If a government hacking operation exceeds the scope of its authorization, the agency in charge of the authorization should report back to the judicial authority the extent and reason;
9. Extraterritorial government hacking should not occur absent authorization under principles of dual criminality;
10. Agencies conducting government hacking should not stock vulnerabilities and, instead, should disclose vulnerabilities either discovered or purchased unless circumstances weigh heavily against disclosure. Governments should release reports at least annually on the acquisition and disclosure of vulnerabilities.

In addition to these safeguards, which represent only what is necessary from a human rights perspective, the judicial authority authorizing hacking activity must consider the entire range of potential harm that could be caused by the operation, particularly the potential harm to cybersecurity as well as incidental harms that could be caused to other users or generally to any segment of the population.

Access Now (www.accessnow.org) defends and extends the digital rights of users at risk around the world. By combining innovative policy, global advocacy, and direct technical support, we fight for open and secure communications for all.

For more information, please contact Amie Stepanovich at **amie@accessnow.org (PGP Fingerprint: CBBE4CF3 84B5FCA7 3BAAF3D0 FF726BC2 1C1DA0C7)** or visit our website **www.accessnow.org**.

accessnow

@accessnow
https://www.accessnow.org