Coming in from the Cold: A Safe Harbor from the CFAA and the DMCA §1201 for Security Researchers

COMING IN FROM THE COLD

A SAFE HARBOR FROM THE CFAA AND THE DMCA §1201 FOR SECURITY RESEARCHERS

DANIEL ETCOVITCH
THYLA VAN DER MERWE

JUNE 2018

ASSEMBLY PUBLICATION SERIES

ABOUT ASSEMBLY

Assembly, at the Berkman Klein Center & MIT Media Lab, gathers developers, managers, and tech industry professionals for fifteen weeks to explore hard problems with running code through collaboration, iteration, and community feedback.

The program has three major components: a course on internet policy that Assemblers take part with other students from Harvard's graduate schools, a design thinking sprint, and a twelve-week development period. All three components of the program emphasize the importance of idea and expertise cross-pollination in order to find novel solutions to these complex problems.

Each Assembly cohort comes together around a defined challenge. In January 2017, the program – then known as the Berkman Klein Assembly – launched its pilot with a sixteen-person cohort focused on finding solutions for the future of digital security. This paper was one of six projects that emerged from the program, and it's unique in that it is a collaboration between a cohort member with expertise in security, Thyla van der Merwe, and a Harvard Law School student, Daniel Etcovitch, who met during the Internet & Society class co-taught by Jonathan Zittrain and Joi Ito.

The paper was completed shortly after the 2017 Assembly program concluded. We nevertheless have included it among the projects produced by the program and as the first publication in the Assembly publication series, which will serve as a place to highlight the collaborative, and often interdisciplinary publications that came out of the Assembly program.

— Assembly Project Team, May 2018

FOREWORD FROM THE AUTHORS

The idea for this paper was born during the Berkman Klein Center's Assembly program in 2017, taking advantage of its interdisciplinary approach and convening of experts. During the classroom component of the program, which was co-taught by HLS Professor Jonathan Zittrain and MIT Media Lab Director Joi Ito, we, the co-authors, found a common interest in legal barriers to security research. The class brought together the legal and policy approaches to technology, and it sparked the initial discussions that led to this collaboration.

Through our discussions, we began to engage with the reality that some security researchers – particularly academics – were concerned about potential legal liability under computer crime laws. The concern is so great that some researchers decide not to pursue research that benefits the public. That led us to develop the ideas presented in this paper for reforming the Computer Fraud and Abuse Act and Section 1201 of the Digital Millennium Copyright Act.

The process was remarkably explorative and interdisciplinary. We engaged with other members of the Assembly cohort, with security researchers all over the world, and with policy minds from across the spectrum. Where one of us looked for legal or judicial solutions, the other would find technological ones. Where technology had governance gaps, we tried to find law and policy plugs.

This work encapsulates the goals of the Assembly program, the bringing together of diverse experts to envision interesting and interdisciplinary solutions, and we are incredibly grateful to have benefitted from the environment it provided.

— Daniel Etcovitch and Thyla van der Merwe, May 2018

Coming in from the Cold: A Safe Harbor from the CFAA and the DMCA §1201 for Security Researchers

Daniel Etcovitch and Thyla van der Merwe[1]

CONTENTS

[1] Harvard Law School, Royal Holloway, University of London. Thank you to Kenny Patterson for guidance and feedback, to Flavio Garcia, Matthew Green, and Rob Carolina for helpful insights, and to all of the staff of the Assembly program at the Berkman Klein Center for Internet & Society for their support.

ABSTRACT

In our paper, we propose a statutory safe harbor from the CFAA and DMCA §1201 for security research activities. Based on a responsible disclosure model in which a researcher and vendor engage in a carefully constructed communication process and vulnerability classification system, our solution would enable security researchers to have a greater degree of control over the vulnerability research publication timeline, allowing for publication regardless of whether or not the vendor in question has effectuated a patch. Any researcher would be guaranteed safety from legal consequences if they comply with the proposed safe harbor process.

1. INTRODUCTION

Security vulnerability research – the study of software and hardware systems in order to identify vulnerabilities or other system characteristics – plays a crucial role in maintaining the safety and sustainability of the technology ecosystem. A critical part of the vulnerability research process is disclosure: security researchers inform the public of the discovered risk and potentially inform the vendor who created the technology product in order to give them an opportunity to mitigate. As shown in the study by Arora et al., the act of the disclosure encourages vendors to patch flawed products.[2] In addition, disclosure allows end-users to make informed decisions regarding the use and purchase of products, along with remedying security flaws so as to reduce their exposure to undue risk. This makes for more satisfied users and allows the market to simultaneously reward security and punish a lack thereof.

The act of informing the vendor prior to public release of a vulnerability, with the intention of allowing for a remediation period, is often termed responsible disclosure. Unfortunately, the security research industry is facing forces that not only makes research more difficult to undertake, but also creates negative incentives to engage in security research in the first place. In addition to technological or economic barriers, security researchers face legal barriers. In particular, parts of U.S. law have created potential liability for good-faith security research, that is, research that actively avoids causing harm to the public, and is used to improve the security of the system in question - curtailing potential socially beneficial activity. As we will examine in more detail, security researchers have been brought to court, charged with criminal offenses, and been stopped by court order from presenting their research at conferences, even though courts have rarely, if ever, decided that the conduct was actually illegal. This is the prototypical example of a legal regime creating a chilling effect: the threat of legal sanction generating fear and uncertainty that pushes individuals away from undertaking a particular activity. According to a recent report, 60% of respondent security researchers cite the threat of legal action as a deterrent to engaging with vendors when it comes to vulnerability disclosure.[3] The importance of security research will only continue to grow as users increasingly rely on the types of devices that are susceptible to security vulnerabilities.

---

[2] Aishish Arora, Ramayya Krishnan, Rahul Telang & Yubao Yang, *An Empirical Analysis of Software Vendors' Patch Release Behavior: Impact of Vulnerability Disclosure,* Information Systems Research, 2010 (finding that vulnerability disclosure accelerated patch releases by vendors).

[3] *Vulnerability Disclosure Attitudes and Actions – A Research Report form the NTIA Awareness and Adoption Group*, NTIA (2016), available at https://www.ntia.doc.gov/files/ntia/publications/2016_ntia_a_a_vulnerability_disclosure_insights_report.pdf.

As stated in the recent 2016 National Telecommunications and Information Administration (NTIA) Awareness Working Group[4] report discussing the attitudes of security researchers and vendors towards disclosure, "removing legal barriers, whether through changes in law or clear vulnerability handling policies that indemnify researchers, can also help" to encourage responsible disclosure. In this paper, we focus on the first suggestion – proposing a change in the law – and aim to provide a legislative solution focusing on the U.S. legal context, which would serve to eliminate chilling effects and create more certainty for security researchers. We focus primarily on academic security researchers, whose careers are dictated by stringent conference publication deadlines. By 'academic researcher' we mean anyone affiliated with an academic institution and publishing under the name of said institution. It is well understood within the academic community that publication is an essential component of a successful academic career. The maxim "publish or perish" is a strong reminder that the progression of science is reliant upon the expression and sharing of ideas, data, and evidence. To not do this is to obstruct science, and perhaps even perform poorly as a scientist. Hence, we focus on academic researchers – those most in need of publication and the ability to resist unreasonable curtailing of publication.[5]

In our paper, we propose a safe harbor for Security Research Activities (as defined below) from two U.S. statutes that have generated these chilling effects, the Computer Fraud and Abuse Act (CFAA) and §1201 of the Digital Millennium Copyright Act (DMCA). The safe harbor is in effect as long as the security researcher complies with a particular implementation of responsible disclosure. It involves the researcher disclosing the discovered vulnerability to the vendor first, waiting a mutually negotiated amount of time in order to give the vendor an opportunity to develop and deploy a mitigation, and then exercising her right to publicly disclose the vulnerability, whether as a published academic paper or otherwise, without the fear of a lawsuit or criminal charges. Of course, not all security research would fall under the jurisdiction of the above statutes. Research that covers security standards as opposed to subjecting specific vendor products to analysis might not constitute a violation of the CFAA or the DMCA §1201, for reasons discussed below, such as the fact that standards are often promulgated openly, giving explicit permission to outsiders to engage in security research. However, it is often the case that vendors confuse this type of research with research that explores particular vendor implementations of these standards, i.e., vendor products, and so will pursue those researchers threatening legal action. As a result, researchers engaged in standards research may sometimes suffer the same chilling effects as those engaged in products research.

We note that this suggestion is by no means new. A recently proposed statute aimed at improving the cybersecurity of Internet of Things devices purchased by the federal government includes a provision explicitly exempting researchers undertaking "good-faith research" done on government-purchased IoT devices from liability under the two statutes.[6] However, that exemption would be very narrow; we propose a far wider scope of exemption for good faith research. In documents published well before such legislation was introduced, the suggestion of protection from legal action has been prescribed in

---

[4] *See id.*

[5] Mike Smith, *Want to be a successful academic? It's all about getting published,* Times Higher Education World University Rankings (2017), available at https://www.timeshighereducation.com/blog/want-be-successful-academic-its-all-about-getting-published#survey-answer; in vulnerability research, top-tier publication is not generally journal-based, as is the case in many other disciplines.

[6] Internet of Things Cybersecurity Improvement Act of 2017, 115th Cong. § 1 (2017).

some vulnerability disclosure standards and guidelines, such the Forum of Incident Response and Security Teams, Inc. (FIRST.Org) guidelines,[7] and is present in certain vendor disclosure policies, such as those of Netflix[8] and Facebook[9] which state that legal action will not be pursued if researchers meet certain requirements. However, upon close examination, these "requirements" effectively dispossess the researcher of any bargaining power when it comes deciding the publication schedule, a potentially dangerous situation for researchers whose careers rely heavily on the ability to publish, and who are often subject to demanding conference deadlines. Our solution aims to not only offer legal protection to researchers, but also to shift the balance of power when it comes to negotiating publication timelines. As will be evident in the sections to follow, the disclosure landscape is, for lack of a better term, messy, and although many vulnerability disclosure guidelines and standards exist, many stakeholders are still unclear of how to approach the process of disclosure. Our solution also attempts to remedy this by offering a default, standardized mechanism for vulnerability disclosure. Of course, it may be the case that release of a vulnerability does not constitute socially beneficial activity. This may be true of a safety critical device, such as a pacemaker. We address this issue when considering counter arguments to our solution in Section 5.

There is also existing discussion of the legal and statutory dimensions of this issue. Regarding the DMCA, there is commentary on how it, and intellectual property system as a whole, should not be impacting security research.[10] Judicial and prosecutorial interpretations of the CFAA have been criticized as well, particularly by advocacy groups.[11] There is even at least one previous suggestion, by Cassandra Kirsch, of a safe harbor for security research, although her exploration of the topic is limited.[12] In a related area, some scholars have suggested reform of the Electronic Communications Privacy Act to facilitate data-sharing between security researchers.[13]

This paper focuses on a detailed, carefully scoped safe harbor for security research, including presenting the particular set of best practices researchers and vendors should comply with, and is organized as follows: Part 2 establishes the relevant definitions. Part 3 provides the necessary background, including a survey of the relevant statutes and court cases, an examination of the existing methods of vulnerability disclosure, as well as how and when they are used by security researchers. Part 4 states our proposed reform in detail, including specific statutory language, a system for classifying vulnerabilities and their associated timelines, and a mechanism for dispute resolution when the parties disagree on the vulnerability classification. Finally, Part 5 addresses both legal and technical counterarguments.

---

[7] *Guidelines and Practices for Multi-Party Vulnerability Coordination*, Forum of Incident Response and Security Teams, Inc. (FIRST.Org) (2016).
[8] Responsible Vulnerability Disclosure, Netflix (2017), available at https://help.netflix.com/en/node/6657.
[9] *White Hat*, Facebook (2017), available at https://www.facebook.com/whitehat.
[10] Deirdre Mulligan & Aaron Perzanowski. *The Magnificence of the Disaster: Reconstructing the Sony BMG Rootkit Incident*. 22 Berkeley Tech. L. J. 1157 (2007).
[11] Orin S. Kerr, *Vagueness Challenges to the Computer Fraud and Abuse Act*. 94 Minn. L. Rev. 1561 (2010).
[12] *See* Cassandra Kirsch, *The Grey Hat Hacker: Reconciling Cyberspace Reality and the Law*, 41 N. Ky. L. Rev. 383, 400 (2014).
[13] Aaron Burstein, *Amending the ECPA to Enable a Culture of Cybersecurity Research*. 22.1 Harv. J. L. & Tech. 167 (2008).

2. DEFINITIONS

In the following, we draw from relevant standards and guidelines to define terms and phrases that will be used throughout this work. We note that the majority of these definitions are sourced from ISO/IEC 29147,[14] and modified accordingly.

Advisory. An announcement or bulletin that serves to inform about a vulnerability in an online service or product. (ISO/IEC 29147:2014)

Coordinator. An optional participant that can assist finders and vendors throughout the disclosure process. (ISO/IEC 29147:2014)

A common function of a coordinator is to help finders locate and contact vendors. Other functions include coordinating vulnerabilities affecting multiple vendors, verifying vulnerabilities, and publishing advisories.

Finder. An individual or organization that identifies a potential vulnerability in a product or online service. (ISO/IEC 29147:2014)

In this paper we take Finder to refer to an external academic security researcher. It is true that a Finder may take on one of many forms, including users, coordinators, vendors, government agencies or specialist security companies, but we limit the scope of our work to individuals most affected by the right to publish: academic researchers who are engaged in Security Research Activities (see Definition below). While our proposed solution could be applicable to, or could be adjusted for, other types of Finders, we limited the scope of our inquiry to what we believe is the clearest case of beneficial vulnerability research.

Mitigation. An action that reduces the likelihood of a vulnerability being exploited and/or the impact of exploitation. (FIRST.Org guidelines)[15]

Online service. A service which is implemented via hardware, software or a combination of the two, provided over a communication line or network. (ISO/IEC 29147:2014)

Remediation. A patch, fix, upgrade, configuration change or document amendment that either removes or mitigates a vulnerability. In the case of hardware products and devices, this could also be in the form of a device recall. (Adapted from ISO/IEC 29147:2014.)

Safety-critical device. A component or system whose failure could result in serious injury to, or loss of life of, individuals.

Security Researcher. An individual who engages in systematic investigation, study or experimentation in order to contribute to general knowledge relating to cybersecurity by establishing, discovering,

[14] *ISO/IEC 29147: Information technology – Security techniques – Vulnerability disclosure*, International Organisation for Standardisation (2014).
[15] *Guidelines and Practices for Multi-Party Vulnerability Coordination*, Forum of Incident Response and Security Teams, Inc. (FIRST.Org) (2016).

developing, elucidating or confirming information about, or the underlying mechanisms relating to computer security, or other related matters to be studied. (Adapted from Public Health Service Act.)[16]

In this work, our use of the term 'security researcher', or even simply 'researcher', covers academic researchers who seek to engage, and are engaged, in good-faith Security Research Activities. We note that researchers may operate in teams. For the purposes of this work, each individual on the team is considered a researcher as per the definition given above.

Security Research Activities. Pursuits which involve accessing software, hardware or an online service solely for purposes of good-faith testing, investigation and/or correction of a security vulnerability, where such an activity is carried out in a controlled environment designed to avoid any harm to individuals or the public, and where the information derived from the activity is used promote the security or safety of the system in question, and its users, and/or any dependent systems. (Adapted from DMCA §1201[17].)

As is evident from the definition, for the purposes of this work, research activities exclude any pursuits which may be carried out by malicious actors wishing to cause harm to users.

User. An individual or organization that directly operates software or hardware products, or makes use of an online service. (ISO/IEC 29147:2014)

Vendor. An individual or organization that develops a product or service and/or is responsible for maintaining it. (ISO/IEC 29147:2014)

Vulnerability. A weakness in software, hardware, or online service that can be exploited. (FIRST.Org guidelines)

Vulnerability Report. All information reasonably necessary for a vendor to confirm the existence of a vulnerability and its possible exploit, potentially including a detailed textual description of the vulnerability, exploit code and a Proof-of-Concept (PoC).

3. BACKGROUND

I. Legal Primer

The disclosure of discovered vulnerabilities to the relevant vendor is not clearly mandated anywhere in U.S. law. Concurrently, vendors do have recourse when individuals exploit their system. The computer crime laws that they would use often implicate security researchers and their work. There are two major U.S. federal statutes under which security researchers might be liable in the normal course of their work: the Computer Fraud and Abuse Act (CFAA) and §1201 of the Digital Millennium Copyright Act (DMCA). Each act makes illegal a particular subset of activities: types of hacking (broadly defined) in the case of the CFAA and circumvention of technological protection measures that protect copyrightable material in the case of the DMCA §1201. Both focus on individuals gaining unauthorized access to protected technological devices. Each law only leaves a small amount of

[16] 42 C.F.R § 52.2 (2009).
[17] 17 U.S.C. § 1201(j)(1).

potential leeway for security researchers to escape liability. State laws are occasionally implicated, though mostly in the context of data breach notification rules, but are not the source of chilling effects in the security research field. The two federal statutes named above criminalize the investigation involved in security research, as opposed to any use or publication of the results. While it will become evident that security researchers are rarely successfully convicted of criminal charges, nor do they often lose civil suits based on these statutes, the uncertainty about potential outcomes restricts their ability to publish research due to restraining orders, forces them to undertake arduous and expensive legal battles, and puts up other barriers that create chilling effects and disincentives for undertaking and publishing security research on vulnerabilities. Each statute will be examined in turn, focusing on how it operates particularly in the context of security researchers, identifying relevant cases, and pointing out particular problems and uncertainties.

a. Computer Fraud and Abuse Act (CFAA)

The CFAA is the most important piece of U.S. legislation used to combat computer crime. Anybody who "intentionally accesses a computer without authorization or exceeds authorized access," (or runs code or a program to do the same) in order to accomplish one of several actions, is liable under the statute.[18] These actions include obtaining financial information, information from any U.S government department, and more.[19] The provision that security researchers generally fear is that it is illegal if an individual "intentionally accesses a computer without authorization or exceeds authorized access" and obtains "information from any protected computer."[20] That language is particularly vague, and has the potential to capture a large amount of research behavior. To fully understand the ambiguity and the areas in which researchers have had difficulty, it is worth examining the statutory definitions of some of the relevant terms in that provision along with the relevant cases below.

A "computer" under the CFAA is defined broadly, as an "electronic, magnetic, optical, electrochemical or other high speed data processing device performing logical, arithmetic, or storage functions"[21] and explicitly includes data storage facilities related to devices. A "protected computer" is a computer used by a financial institution, the government, or is used in interstate commerce[22]. As we will see in a relevant court case, judges have read the definition quite broadly. Devices including Boston's public transit card, the Charlie Card, have fallen under the CFAA's jurisdiction. The definition of "authorized access" has also been debated.[23] Some courts have considered any violation of a site's terms of service to be a user exceeding authorized access. The ACLU has recently challenged that interpretation in federal court on the grounds that it makes actions such as making multiple accounts or recording publicly available information illegal, which halts helpful research such as audit testing to find

---

[18] 18 U.S.C § 1030(a)(2). Even under a broad reading of "exceeds authorized access," which many courts have opted for, the statute would not include liability for researchers investigating certain promulgated standards, because the standards are released publicly and authorize access by researchers.
[19] *See id.* (liability is also triggered for exceeding authorized access on a computer system with intent to defraud and obtain something of value, intent to defraud traffics in passwords, and intent to extort).
[20] 18 U.S.C. § 1030(a)(2)(C).
[21] It is worth noting, in terms of considering the broadness of that definition, that your brain is an electrochemical processing device that performs those functions.
[22] 18 U.S.C. §§ 1030(e)(1)-(2).
[23] Orin S. Kerr, *Cybercrime's Scope: Interpreting "Access" and "Authorization" in Computer Misuse Statutes*, 78 N.Y.U. L. Rev. 1596 (2003).

discrimination.[24] This regime means that any action that a website or vendor (whether of software or hardware) decides is not permitted is deemed illegal with potential criminal penalties. The FBI and the Secret Service are the government entities that have been given the mandate to investigate potential violations of the CFAA,[25] which is an ominous reality that highlights the cybersecurity fears that underpinned the passing of the CFAA.

There are several relevant cases where security researchers have had difficulty with the CFAA. Notable examples include a case involving a group of MIT students and the Massachusetts Bay Transportation Authority (MBTA), and a case involving the researcher Weev and AT&T. In 2008, a group of MIT students found a vulnerability in the Charlie Card – the MBTA's public transit card. They planned to present their findings at DEFCON. However, after hearing about their findings and engaging in an aggressive exchange with the students, the MBTA filed for a temporary restraining order, claiming that the students violated the CFAA in reverse engineering and performing attacks on the Charlie Card. In communication between the parties before the hearing, the students did provide the MBTA with some, but not all, of the discovered information about the vulnerability with hopes that disclosure would settle the matter. While we do not know exactly what information was disclosed (and what was not), the students did only disclose that information after they announced their presentation and after the MBTA contacted them. After that disclosure, the MBTA claimed that the provided information was insufficient, demanded more information and, in their complaint to the court, claimed that responsible disclosure from the outset was an industry norm and suggested it would have resolved the issue. Partially due to the fact that there were no default rules for how responsible disclosure was conducted, meaning that the MBTA could plausibly claim that the disclosure procedures undertaken by the students were insufficient, their argument was convincing to the judge. While this has no legally binding impact, it was part of what made the MBTA's complaint narratively convincing to the judge and likely played a role in his decision to grant the temporary restraining order. Despite vigorous legal fighting by the students – represented by the Electronic Frontier Foundation (EFF) – and much press and academic attention, the restraining order was granted.[26] As a result, the students were not able to present their findings at DEFCON. Later, the court did refuse to extend the restraining order after a subsequent hearing and no CFAA charges were brought against the students. This did not stop the students' research from reaching an audience – the attention around the lawsuit made it so that the slides the students had already submitted to DEFCON, which included limited information about their findings, were widely read. So while the students were brought to court and underwent a difficult set of circumstances, the MBTA's real security-oriented goal was not fully accomplished.

The Aurenheimer case in 2014 provides another useful case study.[27] Aurenheimer, also known as Weev, was able to use a publicly accessible AT&T website and some computer science knowledge to access personal information of customers who had recently purchased iPads. AT&T had simply used reverse chronological identification numbers in their URL, which made it quite easy to access identifying information online. In the government's brief of the case, fear-inducing language was used intentionally to describe how Weev had spoofed his computer to operate like an iPad (a fairly common technique) and painted his actions as highly covert and illicit, when in fact he had not broken into any system at

[24] Complaint, *Sandvig v. Lynch*, No. 1:16-cv-01368 (D.D.C., June 29, 2016).
[25] 18 U.S.C § 1030(d).
[26] Temporary Restraining Order, *MBTA v. Anderson*, Civil Action No. 08-11364-GAO (D.Mass., 2009).
[27] *United States v. Aurenheimer*, 748 F.3d 525 (3rd Cir., 2014).

all. Weev was arrested, successfully convicted under the CFAA, and sentenced to 41 months in prison. An appeals court later overturned the ruling on technical grounds, not ever fully deciding if the CFAA did or did not apply to Weev's activity. Internet advocates were heavily involved in this fight, seeing it as a threat to security research and the open Internet. An amicus brief by "The Mozilla Foundation, Computer Scientists, and Privacy Experts," which included high profile members of the security research community, claimed that such a wide interpretation of the CFAA would criminalize commonplace research techniques.[28] A simple solution, such as an exemption for security research that researchers could rely on, would stop research activities from being caught in the CFAA's web and would stop a situation like this from ever occurring.

b. §1201 of the Digital Millennium Copyright Act (DMCA)

The DMCA was passed in 1998 as an effort at copyright reform in the Internet era, with §1201 in particular being focused on protecting digital rights management (DRM), the technical mechanism by which content owners protect their digital content. Based on this statute, it is illegal to circumvent a technological measure that protects a copyrightable work.[29] This regime was likely designed to apply to issues such as music on Compact Disks. §1201 makes it illegal to break the DRM on the CDs purchased at a retail location, regardless of whether you plan to use the songs ripped simply for personal use. The statute being designed to illegalize any breaking of DRM, regardless of intent or downstream use, covers a massive amount of potential activity. As an analogy, breaking any lock is illegal as long as whatever is on the other side of that door is copyrightable. There is an exemption for good faith security testing in §1201(j), but it is only for testing with the authorization of the owner and if it does not violate the CFAA. The factors the statute uses for deciding whether it is permissible security testing are (1) whether the information obtained is used solely for the benefit of the security of vendor's system or is shared directly with the vendor and (2) whether the method infringes any other statute. Researchers do not *only* use the information for the benefit of the vendor, they aim to publish their research and benefit the security of the larger ecosystem. Much security research might qualify under the second half of the first factor, if it is shared directly with the vendor, but that is only a factor and is not an automatic exemption.

There is one other safety valve in §1201: the Librarian of Congress has the mandate to issue exemptions from the DMCA for particular groups or activities every three years.[30] In 2015, after significant lobbying from civil society groups, the Librarian passed an exemption for security research, but it is limited in three major ways. First, it only applies to automotive vehicles, medical devices designed for implant, and to "devices primarily designed for use by individual consumers (including voting machines)"[31]. What qualifies as a device designed for use by individual consumers remains unclear in the jurisprudence, which has worried advocates and legal scholars. That uncertainty is a large part of

---

[28] Amicus Brief of Mozilla Foundation, Computer Scientists, Security and Privacy Experts in Support of Defendant-Appelant and Reversal, *United States v. Aurenheimer*, 748 F.3d 525 (3d Cir., 2014).
[29] 17 U.S.C. § 1201(a)(1)(A).
[30] 17 U.S.C. § 1201(a)(1)(C).
[31] *Id.*, at (j). A recent Copyright Office report has recommended expanding the exception in (j), but many groups are not convinced the recommendations fix the flaws described herein. *See* U.S. Copyright Office, *Section 1201 of Title 17: A Report of the Register of Copyrights* (June 2017), available at https://www.copyright.gov/policy/1201/section-1201-full-report.pdf; Mitch Stoltz, *Copyright Office Proposes Modest Fixes to DMCA 1201, Leaves Fundamental Flaws Untouched*, Electronic Frontier Foundation (June 2017).

the concern regarding the potential narrowness of this exemption. However the language is construed, it still leaves a significant amount of potential security research outside of the scope of the exemption. Second, the exemption is limited by a requirement that any information gathered be used only for the purpose of protecting the system, which again might eliminate the opportunity for researchers to write up and publish their research. This includes language suggesting information gathered can only be shared with the owner of the device. Third, and potentially most importantly, the entire exemption is conditional on the activity not violating the CFAA. If the activity violates the CFAA, it is not protected by the exemption. As discussed above, this captures a wide scope of activity. The Copyright Office, in recommending this limited exception, went out of its way to proclaim that it explicitly "favors responsible disclosure of security flaws."[32] In justifying the limited scope of this exception, the Copyright Office discussed concerns with a large exception risking wide disclosures causing security problems for health and safety, cited opposition to a broader exception by several government agencies including the FDA and EPA, and emphasized that they wanted to "allow room for other interested agencies to weigh in on this national debate."[33]

DMCA §1201 has been used in several cases involving security research. Some notable cases that we touch on below involve Adobe and the arrest of Dmitry Skylarov, and Ed Felten's potential interaction with Sony. Dmitry Skylarov, an employee at a Russian software company, developed a program that converted Adobe e-books into PDFs. This software was legal where he developed it. He traveled to the United States for DEFCON in 2001 and was arrested on the basis of §1201. He was eventually allowed to return home, but federal prosecutors brought criminal charges against the company he worked for, though the company was acquitted after an 18-month legal battle. There is no shortage of examples of researchers who have been scared off by §1201 charges. The EFF cites a long list of security researchers who changed their behavior as the result of the charges against Sklyarov.[34]

Our other example involves Professor Edward Felten, recently Chief Technologist at the FTC. In 2005, Felten found that Sony CDs were installing rootkits on computers that created a serious vulnerability. While Felten wanted to inform the public, he decided not to do so immediately. By his own account, the prosecutions under DMCA §1201 had become a source of fear for security researchers and that fear stopped him from disclosing the vulnerability.[35] Felten and his team sat on their discovery for a period of time, during which another researcher found the vulnerability and made it public, after which they followed suit. Luckily nobody sued either of them and Sony eventually settled with the FTC, but it is troubling that a high-profile researcher who had made a discovery with serious impact decided to withhold that research both from the public and from Sony due to fear of the law. This example shows that even in cases where the vendors having or resorting to legal recourse is less likely, there might be chilling effects on risk-averse researchers. The EFF filed a lawsuit challenging the 2015 security exemption discussed above, claiming it was too narrow. One of the named plaintiffs is Matthew Green, a security researcher at Johns Hopkins University. The complaint states that Green

---

[32] *See id.*, at 318.
[33] *See* U.S. Copyright Office, *Section 1201 Rulemaking: Sixth Triennial Proceeding to Determine Exemptions to the Prohibition on Circumvention*, 317 (Oct. 2015).
[34] *Unintended Consequences: Fifteen Years Under the DMCA*, Electronic Frontier Foundation (March 2013), available at https://www.eff.org/pages/unintended-consequences-fifteen-years-under-dmca.
[35] Edward Felten, *The Chilling Effects of the DMCA*, Slate.com (2013), available at http://www.slate.com/articles/technology/future_tense/2013/03/dmca_chilling_effects_how_copyright_law_hurts_security_research.html.

has declined to investigate certain products because of fear of §1201, a reality in security research that endangers the security of all devices, according to the complaint.[36] Both of these cases and their fallout demonstrate the chilling effects that DMCA §1201 has created, pointing to the need for not only a safe harbor for security researchers, but also to the need for a solution that explicitly provides more certainty to those wishing to engage in security research.

c. Potential Consequences

In order to assess the impact of this and consider potential solutions, it is important to consider how it actually impacts researchers. Testimonials from policy-conscious researchers like Felten and Green is relevant, but there is an open question as to whether researchers are even aware of the issue. The answer to that question is mixed. Some researchers are likely not aware of the detailed legal landscape, and most are likely not aware of the specifics. Many, though, are aware of the specter of potential liability and have heard ghost stories about security researchers facing criminal charges. That, in and of itself, chills potential research into some products or products by certain companies, as Green claims in his testimony. A researcher does not need to be aware of the exact bounds of the law (which are broad and uncertain for both the CFAA and DMCA §1201) in order to be concerned. In fact, a more vague, less defined fear might chill research even further than possession of the appropriate legal knowledge.

For the CFAA in particular, the uncertainty regarding its potential application to security research causes several problems. First, and most obvious, it has a chilling effect on potential security research. If a researcher knows a company or organization is potentially aggressive and has troubled relationships with security researchers, they might shy away from conducting research that could be highly helpful to the security and privacy of users. Some might also avoid the field of security research all together. The system, as it exists, incentivizes litigious, intimidating behavior. Second, knowing that a company can criminalize your activity by simply adding a line to their terms of service gives vendors a disproportionate amount of power in deciding the bounds of the law, and control over the field of security research.

The burden of potential CFAA claims also creates policy issues for researchers. While there are not many cases of security researchers actually being convicted under the CFAA, there is no shortage of examples where their lives were upended or they were forced to fight expensive and emotionally difficult legal battles. The MIT students had to surrender their documents and were prevented from presenting their paper. Weev was arrested and sentenced to multiple years in prison before his conviction was overturned. Stefan Puffer, as a contractor for his local county, ran a security test with a county official and a reporter in the room, in order to show that he could compromise a county computer. The county official allegedly told him that it was not a big deal at first, but once the reporter ran a story about the weak security of the county's computers the county called the FBI. They went to Puffer's home at 6:00 AM and confiscated his electronic devices. A jury would eventually acquit him, but only after getting all the way to a trial. That means the case survived a motion to dismiss, meaning the judge thought the charges were a cognizable claim, and survived a motion for summary judgment, meaning the judge thought there were material issues of fact to be decided. Puffer went through the full process, including a discovery process that would have meant disclosing his personal emails and

[36] Complaint, *Green v. United States*, No. 16-cv-01492 (D.D.C., 2016).

more, in order to be quickly acquitted at trial. Both Moulten and Puffer were charged under the CFAA. The hardship, the specter of criminal charges, and the potentially large cost of even the early stages of litigation are intimidating to researchers. If one believes that security researchers do useful work that has social and systemic benefit, this presents a serious issue. The FBI knocking on your door is likely enough make the average person more conservative in their daily activity. Even though as it progresses to trial, courts are likely to acquit a security researcher, that does not remove the fear of investigation or the cost, both emotional and monetary, of the early stages of litigation. Civil society groups who represent security researchers in scenarios like this, as the EFF did in MBTA v. Anderson, can be a huge help, but they cannot be everywhere at once. If we believe security research is valuable, this system is causing social harm.

In terms of DMCA §1201, the consequences are similar. The legislation captures a wide range of activity, potentially including research activity that is socially beneficial. Security researchers are contributing to the security and health of the technology landscape by conducting activity that might make them liable under DMCA §1201, from investigating CDs planting software on your devices to investigating the security of wireless car keys. The exemption process that happens every three years was designed as a safeguard, but not only is the process of applying for an exception arduous and overly legalistic, which creates a barrier for security researchers, it also has only produced limited exceptions. The 2015 exemption includes automotive vehicles, medical devices designed for implantation, and consumer devices designed for personal use. That does not cover the full scope of products where security research might be valuable. The legislation still captures potentially valuable research and its broad application in the past and present creates real chilling effects, as demonstrated by the experiences of Edward Felten, Matthew Green and many others.

In terms of the actual legal boundary, the security research industry is often relying on the discretion of the manufacturers to not bring claims against good faith security researchers, or relying on prosecutorial discretion in criminal cases. When the laws are overbroad and include the research activity, it will be up to prosecutors whether to pursue the case. While we may want to put faith in the prosecutor's office, there is evidence, as shown above, that not only do many prosecutors not have expertise in the relevant technical issues, but also that they are willing to prosecute in security research scenarios. Also, even if the vast majority of prosecutors reliably avoid going after researchers based on policy concerns, it likely only takes a very small number of incidents to create chilling effects. As researchers hear anecdotal stories of prosecution, they are less likely to engage in research relevant to the products or manufacturers implicated. If laws are criminalizing and/or chilling socially beneficial conduct, it is time to reevaluate them.

With that legal regime as a backdrop, we next discuss current vulnerability disclosure methods and practices.

II. Available Disclosure Methods

The disclosure methods available to finders, loosely speaking, include *full disclosure* and *responsible disclosure*. In the case of full disclosure, finders publicly report on discovered vulnerabilities without informing the affected vendor beforehand. The process of responsible disclosure advocates informing the relevant vendor, or set of vendors, prior to releasing any vulnerability details. In its most basic form, responsible disclosure usually allows for a vendor to develop a remediation, or to investigate

other satisfactory mitigations via the provision of a specified time window between vendor notification and public release. It is possible that finders and vendors agree on the duration of this time period, and work together to remedy the identified problem. This is known as coordinated vulnerability disclosure, and we consider this to be a form of responsible disclosure.

Prior to this work, no legislative framework governing vulnerability disclosure has been adopted, and these practices only exist as part of guidelines and recommended standards.[37] In the sections to follow, we provide a description of vulnerability disclosure practices, and discuss the perceived advantages, disadvantages, and stakeholder motivations associated with each method.

a. Full Disclosure

From the finder's perspective, full disclosure is perhaps the easier practice to execute; upon finding a vulnerability, the finder publishes vulnerability details without informing the affected vendor, or set of vendors. Some security experts argue that this practice prompts vendors to take speedy action in supplying a remediation or countermeasure to the problem. However, it is argued that this practice leaves systems, and indeed users, entirely vulnerable for the time period taken by the vendor to find and announce an appropriate remediation or countermeasure (unless of course users have the option of not using the affected product or service, and indeed follow this course of action). This practice informs, and potentially empowers, all classes of attacker: prior to publication, the vulnerability may only have been known, and potentially exploited, by attackers with the means and resources to have found it early on. In the time between disclosure and application of the appropriate remediation, the vulnerability can be exploited by all attackers who come to know of the finder's publication. Although this disclosure method seemingly absolves the finder of practical responsibility in that no communication with the vendor is required, it arguably does not result in the best outcome for system users, and also places a potentially taxing and rushed patch action on vendors, resulting in remediation that may not be optimal. The situation is particularly dire in the event that vendors do not become aware of the vulnerability, and malicious actors do.

b. Responsible Disclosure

This method of disclosure involves communication between the finder and the affected vendor: upon discovery of a vulnerability, the finder informs the vendor of the vulnerability details and affords the vendor time to devise an initial solution prior to publicizing the vulnerability details. The solution could be a complete mitigation or a stopgap that protects users of the system until such time as a complete mitigation is developed. Early standards, such as those published by the Internet Engineering Task Force (IETF)[38] in 2002, advocated offering the vendor a fixed period of thirty days to develop a remediation and/or the appropriate mitigations, with the suggestion that this period be extended if the scope of the problem is particularly large, and if the vendor acts in good faith to resolve the vulnerability.

---

[37] Steve Chistey & Chris Wysopal, *Responsible Vulnerability Disclosure Process – Internet Draft*, Internet Engineering Task Force (2002).
[38] *Id.*

Over time, this fixed grace period offered by the finder seemingly evaporated as the practice of coordinated vulnerability disclosure (CVD), introduced by Microsoft in 2010[39] became more popular, especially amongst vendors. As stated in Microsoft's initial announcement, under this disclosure principle, the finder "allows the vendor the opportunity to diagnose and offer fully tested updates, workarounds, or other corrective measures before any party discloses detailed vulnerability or exploit information to the public". Under this principle, the vendor provides the finder with regular updates regarding the development of the remediation, and it is possible that the vendor and the finder may work together to remedy the situation. As the timing of public disclosure rests largely on the time taken to develop a remediation, this form of disclosure generally favors vendors – a researcher wishing to publish may be forced to delay publication until the vendor has developed a solution or workaround. And indeed, more recent standards such as ISO/IEC 29147[40] by the International Organization for Standardization (ISO) are heavily focused on vendors, thus ignoring the needs of finders.

We note that there appears to be some ambiguity regarding what constitutes responsible disclosure. More recent standards imply that responsible disclosure is synonymous with coordinated vulnerability disclosure. In this work, we consider responsible disclosure to be any form of disclosure in which a finder informs a vendor of a discovered vulnerability prior to publicly releasing any of the vulnerability details, and additionally provides the vendor with a reasonable amount of time to develop a mitigation. We consider coordinated vulnerability disclosure to be a form of responsible disclosure in which the finder does not specify a fixed period for remediation development but instead agrees to wait for the development of the appropriate fix before publishing any details.

In some instances, a finder may not be able to establish a relationship directly with the vendor. In such cases, the finder may make use of what is known as a coordinator. This is an individual or organization that has the ability to act on behalf of one of the stakeholders, and potentially help to verify stakeholder claims and resolve conflicts (see Section 2). Coordinators are commonly well-known trusted third parties (TTPs). For instance, the CERT Coordination Center (CERT/CC) of the computer emergency response team for the Software Engineering Institute, offers coordination assistance to finders if the vulnerability affects multiple vendors, or if vendors are hard-to-reach or unresponsive. This organization also offers its services in the event that there is a disagreement or dispute between finders and vendors. Notably, once a vulnerability is received, this organization will allow vendors a period of forty-five days before publicizing vulnerability details. However, as stated in the CERT/CC vulnerability disclosure policy,[41] the organization is willing to be flexible in circumstances where the threat implied by the vulnerability is of a critical (or indeed trivial) nature, or in which changes to established standards may be required. The organization is also willing to negotiate alternate publication schedules with affected vendors, therefore engaging in CVD as initially described by Microsoft, but this is not specified as the default behavior.

---

[39] Matt Thomlinson, *Announcing Coordinated Vulnerability Disclosure*, Microsoft Corporation (2010).
[40] *ISO/IEC 29147: Information technology – Security techniques – Vulnerability disclosure*, International Organisation for Standardisation (2014).
[41] Vulnerability Disclosure Policy, CERT/CC (2017), available at http://www.cert.org/vulnerability-analysis/vul-disclosure.cfm.

At a high level, the responsible disclosure process, with or without a coordinator, involves the following key steps:[42]

1) Discovery. An individual, group of individuals, or organization finds a vulnerability and (possibly) a means of exploiting this vulnerability.

2) Notification. The finder informs the vendor or coordinator of the vulnerability. Initially, the finder may alert the vendor or coordinator of the existence of the vulnerability. If responsive, the vendor should provide an acknowledgement of the alert by the finder, as well as a means of communicating the vulnerability details securely. The finder should provide a detailed report of the vulnerability. In the case of basic responsible disclosure, the finder may also signal the duration of the remediation period.

3) Validation. The vendor verifies the finder's claims and confirms the existence of the vulnerability.

4) Resolution. The vendor develops and tests a fix or workaround. In the case of coordinated vulnerability disclosure, the vendor regularly updates the finder regarding the status of the fix, and the finder and vendor may work together closely in an attempt to develop a fix or workaround. In the case of basic responsible disclosure, the vendor should develop the fix or workaround within the remediation period stipulated, request an extension, or admits that the problem cannot be fixed.

5) Publication. The vendor, coordinator (if applicable), and the finder publicly release details of the vulnerability, along with its resolution. We note that some guidelines, such as ISO/IEC 29147, recommend publishing an advance advisory in the event that there is evidence that a vulnerability is being actively exploited in the wild. In other words, limited details of the vulnerability, as well as interim mitigations, are released so as to protect users from malicious attackers who may be exploiting the vulnerability.

The issue of vulnerability disclosure in the multiple vendor case has garnered more attention in recent years, with organizations such as the FIRST.Org, developing guidelines for disclosure in this model.[43] Our solution is presented in terms of one finder engaging with one vendor but we describe mechanisms for dealing with the multiple vendor case, and address issues that arise due to this complex situation throughout this work.

As evidenced in the previous section, disclosure is treated the same way by U.S. law regardless of disclosure method and can result in unfavorable outcomes for finders. The responsible disclosure methods and principles described above exist in guidelines and as best practices only; nowhere are these mandated by U.S. law, and hence in no way can they help to protect finders who choose to disclose responsibly. This reality creates an unusual set of incentives. The potential for suit under the CFAA or DMCA §1201 pushes researchers not to engage in any of these disclosure methods, thereby

---

[42] Steve Chistey & Chris Wysopal, *Responsible Vulnerability Disclosure Process – Internet Draft*, Internet Engineering Task Force (2002); *ISO/IEC 29147: Information technology – Security techniques – Vulnerability disclosure*, International Organisation for Standardisation (2014).

[43] *Guidelines and Practices for Multi-Party Vulnerability Coordination*, Forum of Incident Response and Security Teams, Inc. (FIRST.Org) (2016).

not disclosing at all. Researchers are driven away from responsible disclosure because vendors have complete leverage over them; a vendor can decide to never remedy the vulnerability and can still threaten legal action. The system as it exists is structured to harm altruistic finders, rather than to help them and the system at large.

III. Existing Disclosure Practices

After considering the development of vulnerability disclosure philosophies, as well as the formation of the vulnerability disclosure ecosystem, we now turn to vulnerability disclosure in practice and report on actual disclosure events, covering what we would consider to be successful deployment of responsible disclosure, as well as unsuccessful and/or negligent attempts.

We draw on the NTIA study that aimed to investigate the current, real-world levels of adoption of established disclosure practices and what barriers to adoption may exist. The report confirms that whilst security researchers generally engage in some form of coordinated disclosure, with 92% of survey respondents stating as much, many of them (60% of respondents), cite the threat of legal action as a deterrent to engaging with vendors when it comes to vulnerability disclosure. The report also found that when researchers have opted to pursue full disclosure, it has been largely out of frustration with regards to the lack of communication by vendors. Hence, we observe that although potentially promising in theory, CVD seems to fail in practice owing to a lack of actual coordination, and the chilling effects created by the threat of legal action. On the point of legal action, we note that the report states that researchers are discouraged from working with a vendor to disclose vulnerability. This appears to leave room for researchers to pursue full disclosure, however, in this case, they may also be subject to legal action, as is evidenced by the MBTA v. Anderson case described in Section 3.I. Nevertheless, the possibility of landing in legal hot water does clearly prey on the minds of security researchers.

The study also found that timelines are important to the academic research community. This is unsurprising given that a researcher's livelihood depends on the ability to publish. Nine in ten researchers expressed a desire for a remediation deadline to be in place. Interestingly, based on evidence provided by the study, many researchers appear to be willing to be semi-flexible regarding this deadline, as long as a) the deadline is finally set, and b) the vendor effectively communicates the issues surrounding the severity of the vulnerability and the resulting remediation. In other words, researchers seem to be reasonable when it comes to timelines as long the remediation timeline fits the severity of the vulnerability. As we will see in the sections to follow, our proposed solution takes the need for a deadline, as well as the willingness to be flexible, into account.

The report found that vendors with mature disclosure practices tended to develop these practices via an internal methodology, largely ignoring practices followed by their peers and those suggested in international standards. This points to a large discrepancy amongst vendors when it comes to vulnerability handling (including the receipt, processing and finder communication intentions surrounding the vulnerability), and again highlights the difference between theory and practice – even

though standards and guidelines exist, these may very well be ignored in practice. We direct the reader to the report[44] for further details concerning the number or respondents and survey methodology.

The vendor handling discrepancy is evident across various vulnerability disclosure policies in the technology industry. Netflix[45] and Facebook,[46] for instance, both offer a secure means of reporting a vulnerability, and both organizations claim that they will not initiate a lawsuit or bring legal action against the finder if the finder meets various conditions. One of these conditions, as given in both policies, is that the finder allow for a reasonable mitigation development period before publicly releasing the vulnerability details. Although introducing a safe harbor in theory, this effectively means that the publication deadline is vendor-controlled, and neither policy quantifies a 'reasonable' amount of time. In the case of Facebook, finders may additionally be offered a financial reward in the form of a bounty for their discovery. Netflix credits finders by naming them on the Netflix Security Researcher list (if the discovered vulnerability falls within the recognition scope). Neither policy explicitly mentions continuous communication with the finder, or hints at coordination when resolving the problem caused by the vulnerability – these are two of the fundamental principles of CVD.

The Google vulnerability disclosure[47] appears to be entirely in the form of a rewards program and there is no explicit mention of the provision of safe harbor, although it is hinted at in Google's security philosophy, with 60 days being cited as a reasonable amount of time to offer vendors for remediation purposes (prior to publically releasing vulnerability details). In the 2010 blog post[48] by Google's security team which lists this figure, it is also stated that in cases where Google has been be unable to meet reasonable publication timelines, the organization has been content to let publication proceed, indicating that the publication times are finder-controlled. However, we recognize that this sentiment may now be dated and that decisions of this nature are solely at the discretion of Google.

Interestingly, but perhaps unsurprisingly, companies that are in the business of producing hardware (and associated firmware) products have somewhat different vulnerability disclosure policies and handling processes. Qualcomm, for instance, runs a rewards program, which is by invitation only.[49] Their program specification lists devices which are considered to be within the research scope and an agreement not bringing legal action is not mentioned. The NXP disclosure policy[50] on the other hand addresses all finders and clearly defines the establishment of secure communications channel for the receipt of vulnerabilities. Again, a vulnerability scope is defined, and the policy clearly states that NXP is committed to working with finders to establish a publication timeline. However, the policy draws attention to the fact that hardware products may require longer remediation periods as it may be difficult (and perhaps impossible) to push updates to products in the field, or to recall products. We

---

[44] *Vulnerability Disclosure Attitudes and Actions – A Research Report form the NTIA Awareness and Adoption Group*, NTIA (2016), available at https://www.ntia.doc.gov/files/ntia/publications/2016_ntia_a_a_vulnerability_disclosure_insights_report.pdf.

[45] Responsible Vulnerability Disclosure, Netflix (2017), available at https://help.netflix.com/en/node/6657.

[46] Whitehat Information, Facebook (2018), available at https://www.facebook.com/whitehat.

[47] *Google Security Reward Programs*, Google, available at https://www.google.com/about/appsecurity/programs-home/.

[48] Chris Evans et. al, *Rebooting Responsible Disclosure: a focus on protecting end users*, Google (2010), available at https://security.googleblog.com/2010/07/rebooting-responsible-disclosure-focus.html.

[49] *Qualcomm Vulnerability Rewards Program*, Qualcomm, available at https://hackerone.com/qualcomm.

[50] *Product Security Incident Response Team*, NXP Semiconductors, available at http://www.nxp.com/about/about-nxp/corporate-responsibility/product-security-incident-response-team:PSIRT.

address the differences between hardware and software in the context of vulnerability disclosure, where we also discuss the mechanism of bug bounties.

In order to further provide a flavor of vulnerability disclosure in practice, in relation to the philosophies and models discussed in the previous section, we now consider the Megamos Crypto vulnerability disclosure incident.[51] Although the case does not implicate U.S. law, we discuss it as an example of the chilling effects created by the interaction of the relevant legal regimes and security vulnerability research, and to highlight the disruption caused to researchers as a result of the community-wide vague understanding of responsible disclosure.

In 2013, three European-based security researchers wanted to publish the details of an attack on the Megamos Crypto system, an automatic immobilizer system intended to guard against the theft of vehicles employing this system. Volkswagen used this product in some of its products. The algorithm used in the system was a proprietary cryptographic algorithm developed by Thales, and Thales in turn authorized the chip manufacturing company EM to embed the algorithm into microprocessor chips. In order to conduct their analysis, the researchers needed access to internal workings of the algorithm. They obtained these details by reverse engineering a third-party product, Tango Programmer, which could create keys for immobilizers using the Megamos Crypto system. The researchers found several flaws in the system, and in November of 2012 they approached EM with a description of the vulnerabilities, without knowledge that EM had been licensed by Thales to use the algorithm.

The researchers aimed to publish their findings at a top-tier security conference, USENIX Security Symposium, in August of 2013, but just prior to the conference Volkswagen learned of the work and filed a lawsuit in the High Court of England against all three members of the research team, and against their academic institutions. Publication of their paper was delayed because of this lawsuit, and the paper finally officially appeared in the USENIX Security Symposium in 2015, with a crucial part of the algorithm redacted.[52]

Regarding the responsible disclosure of the weaknesses found in the Megamos Crypto system, an excellent treatment of the case by Carolina and Paterson[53] points out that the researchers informed the chip manufacturer, EM, nine months prior to the desired publication date, without much in terms of a response from the company. The court then granted Volkswagen an emergency temporary injunction citing that the researchers did not disclose responsibly, which seems at odds with what is understood by 'responsible disclosure' within the security community. Although the multiple vendor issue potentially had a role to play in this case, there is obvious tension between what is understood about responsible disclosure by researchers, vendors and the courts.

In practice the vulnerability disclosure landscape is somewhat chaotic and unstructured; although relatively clear guidelines and standards exist, there appear to be many disparate approaches to dealing with vulnerability disclosure in the real world, many of which don't give the finder much influence over

---

[51] *Volkswagen Aktiengesellschaft v. Garcia, et al.*, 2013 EWHC 1832 (Ch) (June 25, 2013).
[52] Roel Verdult, Flavio D. Garcia & Baris Ege, Dismantling Megamos Crypto: Wirelessly Lockpicking a Vehicle Immobilizer. Supplement to the 22nd USENIX Security Symposium (USENIX Security 13) (2015), available at https://www.usenix.org/node/193261.
[53] Robert Carolina & Kenneth G. Paterson, *Megamos Crypto, Responsible Disclosure, and the Chilling Effect of Volkswagen Aktiengesellschaft vs Garcia, et al*, Royal Holloway, University of London (2013), available at http://www.isg.rhul.ac.uk/~kp/Carolina-Paterson-Megamos-comment-20130828.pdf.

the publication deadline, and some which result in legal action against the finder. As pointed out in the NTIA report, not all security researchers and vendors fully embrace the principles of state-of-the-art disclosure practices such as CVD, and the lack of trust between the two groups hurts not only researchers and vendors but also, importantly, end-users of technology. Our proposed solution aims to provide a unified and standard approach to vulnerability disclosure which would reduce the chilling effects on security research created by the existing legal regime, and would distribute the balance of power more evenly in the act of establishing publication schedules.

4. PROPOSED REFORM

In order to address the structural and incentive problems discussed above, we propose a legislative solution aimed at not only eliminating legal liability for good faith security research, but also at creating and disseminating a norm of responsible disclosure throughout the industry. The proposed reform aims to eliminate liability for the investigation activities that the two statutes discussed above criminalize, acting before any use or publication of the findings even become relevant. We believe that the statutory reform that we propose addresses the oft-cited shortcomings of responsible disclosure under the current system: vendors not mitigating vulnerabilities, continued liability for security researchers, and concerns about the safety of the ecosystem during the responsible disclosure process.

The proposed reform has two steps. First, we propose a statutory safe harbor that exempts security researchers who comply with a series of best practices from liability under the CFAA and under DMCA §1201. Second, we propose a communication and classification system under which the two parties, the researcher and the vendor, share information about the vulnerability and negotiate a classification for the vulnerability that will determine the relevant amount of time before the researcher is permitted to publish the vulnerability, regardless of whether the vendor has effectuated a mitigation. The safe harbor will naturally apply if this communication process is followed; the two parties are free to contract around the communication process, but will always be able to rely on the safe harbor by reverting to the established communication and classification protocol.

The following section explains the contours of the proposed legislative change, the communication and classification system, and a negotiation system for when the parties do not agree on the classification of the vulnerability. Our solution focuses on the case in which publication is deemed to be possible, and of benefit to the security community and end-users. We discuss arguments for when publication may not be possible, nor of benefit to systems, and the extent to which we agree with these arguments, in Section 5.

I. Statutory Reform

a. The Statutory Safe Harbor: Language and Conditions

We propose a statutory safe harbor that as a default rule, which would exempt security researchers from liability under the CFAA and DMCA §1201[54] for conduct that qualifies as Security Research

[54] This safe harbor would be in addition to, not a replacement for, the narrow exception in 1201(j). 17 U.S.C. § 1201(j).

Activities under the definition in Part 2. The statutory language could be executed as a standalone statute referring back to both other statutes or as an amendment to each of the two statutes, as is done in the recently proposed IOTA bill.[55] We are not the first to suggest such a safe harbor, but our proposal aims to finally provide a comprehensive guide to the specifics of such a statutory reform, and address some of the weaknesses of previous proposals.[56] An important detail of our proposal is that for each classification of vulnerability there is a pre-determined time period after which the researcher is permitted to disclose the vulnerability to the public, regardless of whether the vendor has patched the vulnerability or not, and the researcher would still retain their immunity from suit under the two statutes. This is not a coordinated vulnerability disclosure regime; the vendor does not control the public disclosure timeline. This addresses the chilling effects still present under existing voluntary responsible disclosure regimes and gives vendors a credible incentive to patch vulnerabilities in a reasonable amount of time.

The proposed safe harbor would exempt researchers only from the statutes discussed above, and only for the scope of their research activities. The key condition for obtaining this research safe harbor is engaging in the specific implementation of responsible disclosure that we will describe in more detail below. To qualify, the researcher would be required to disclose the crucial details of the vulnerability to the vendor, engage with a negotiation protocol over the classification of the vulnerability, and then refrain from disclosing the vulnerability or its existence to the public until the time period associated with the classification has elapsed. The proposal would act as a default rule, one that parties could mutually agree to circumvent via contract.[57] This proposed responsible disclosure regime operates as a baseline from which researchers can operate, in order to finally afford the research community adequate bargaining power, without the threat of legal action, when engaging with vendors. At the same time, it still allows vendors the opportunity to attempt to negotiate. The timeline attached to the vulnerability classification acting as a default would make it more difficult for vendors to justify withholding patches or requesting unreasonably extended timelines because the researchers could rely on the proposed safe harbors and the associated timeline. If the researcher is willing to allow a longer period by delaying publication, she could negotiate for other concessions. If a vendor is willing to provide monetary remuneration for disclosed vulnerabilities, as some do in existing bug bounty programs,[58] the vendor could request a more flexible or restrictive timeline which researchers could choose to accept, or which the researcher could decline in reliance on the safe harbor.

This system maintains the freedom of contract between the parties while providing more equal bargaining power and ensuring that the default outcome has the systemic benefits of mitigating chilling effects and encouraging expeditious patching of vulnerabilities. Those nuances differentiate our reform from other safe harbor proposals. Kirsch's proposed safe harbor suggestion would exempt researchers from suit so long as they disclose a discovered vulnerability to the vendor within 24-48 hours of discovery, and then waits some short "reasonable" period of time before disclosing it to the public,

---

[55] Internet of Things Cybersecurity Improvement Act of 2017, 115th Cong. § 1 (2017).

[56] *See* Cassandra Kirsch, *The Grey Hat Hacker: Reconciling Cyberspace Reality and the Law*, 41 N. Ky. L. Rev. 383, 400 (2014). For a more casual suggestion, *see* Ed Felten in *The Chilling Effects of the DMCA* (2013), available at http://www.slate.com/articles/technology/future_tense/2013/03/dmca_chilling_effects_how_copyright_law_hurts_security_research.html.

[57] For a more detailed discussion of how default rules operate in the legal context, as well as an exploration of potential issues with them, *see* David Charny, *Hypothetical Bargains: The Normative Structure of Contract Interpretations*, 89 Mich. L. Rev. 1815 (1990).

[58] *See* our discussion of bug bounty programs in Section 5, *supra*.

regardless of vendor action.[59] To address the concern that vendors will not implement mitigations, Kirsch proposes that the Federal Trade Commission should bring consumer protection actions.[60] This proposal takes a mostly full disclosure approach to building a safe harbor for security research, while we are proposing a responsible disclosure approach. We believe a responsible disclosure approach to be more prudent because it gives vendors an opportunity to patch or mitigate the vulnerability before it is made public, shielding users from some potential exploitation of the vulnerability for the period before disclosure. We also believe that limiting the scope of government entities involved, and not relying on agency enforcement, is likely to create more consistency across multiple administrations as well as more predictability. While we by no means believe our proposal to be the final, authoritative version of any statutory safe harbor for security research, we believe this proposed reform to be the most comprehensive so far. The legislative text, relying on the definitions above in Part 2, would read:

> A researcher shall not be sued, held liable or charged under the Computer Fraud and Abuse Act or §1201 of the Digital Millennium Copyright Act for their Security Research Activities if the researcher, in good faith,
>
> 1) Within two days of confirmed discovery, discloses the discovered vulnerability to the vendor in the form of a Vulnerability Alert followed by a Vulnerability Report
>
> 2) Includes in the Vulnerability Report all information reasonably necessary to the vendor's effort to find and patch the vulnerability
>
> 3) Includes in the Vulnerability Report a Preliminary Vulnerability Classification
>
> 4) If necessary, engages in the Classification Dispute Process
>
> 5) Complies with the Communication Process
>
> 6) Does not publish information relevant to the discovered vulnerability until the time period associated with the Final Classification has elapsed

In addition, if a researcher complies with the conditions above and the Communications Process below, but the vendor does not comply with their portions of the Communication Process, the vendor would be estopped from asserting claims under the CFAA and DMCA §1201 against the researcher for their research activities. Vendors would be encouraged to operate a secure channel to intake Vulnerability Reports that is clearly accessible to potential security researchers, which is crucial to ensuring that more recent entrants into the security research field are not disadvantaged by not having existing contact with vendors. Vendors would also be required to not disclose the vulnerability to the public until the timeline has elapsed, unless required by law such as due to a judicial ruling or subpoena, in order to not eliminate the incentive for the researcher to responsibly disclose due to vendors pre-empting their research publication. The requirement that all relevant submissions and communications be made in good faith merits highlighting and explaining. Any such statutory reform should include punishments for actors who submit Vulnerability Alerts, Reports, or any other such document without real reason to do so. The proposed reform imagines a statutory penalty for such activity, including

---

[59] Kirsch, *infra* note 56, at 400.
[60] *Id.*, at 401.

disqualification from the safe harbor, with the aim of preventing actors from spamming and overloading the intake systems of vendors who are seeking information about real vulnerabilities.[61]

b. The Communications Process

The Communications Process, the protocol by which the two parties communicate about the vulnerability, is also crucial. We describe this process in terms of one finder and one vendor and defer discussion of multiple vendors to the next section. The Communications Process includes not only disclosure of the vulnerability, but also discussion and negotiation regarding the classification of the vulnerability. Details considering our proposed classification scheme and a protocol for contested classification will be discussed in Part II of this Section.

The proposed Communications Process is as follows:

1) Researcher sends Vulnerability Alert to vendor
2) Vendor sends automated invite to a reasonably secure channel
3) Researcher sends basic information about vulnerability in order to establish credibility
4) Vendor screens received message, deciding whether to respond
5) Vendor responds, requesting further details
6) Researcher expeditiously sends Vulnerability Report including all reasonably relevant details
7) Both parties blindly commit Vulnerability Classifications
8) Vendor may contest Classification, including sending further information about the vulnerability or system to Researcher in order to convince them of a new Classification. See Contested Classification section below
9) If the parties do not agree on Classification, they engage in the Contested Classification Protocol[62]
10) If the parties do agree, the relevant timeline associated with the Classification becomes binding

There are several important components to this process that bear explicit explanation. First, steps five through eight should be executed over a maximum period of one week, with sufficient time being provided for both the vendor and the finder to respond satisfactorily. Second, it was a deliberate design decision that the researcher does not transmit all relevant information about the vulnerability at the very beginning of the communication process. Researchers seek to communicate with vendors who take the vulnerability seriously, and desire to maintain some leverage in the engagement. At the same time, there is no detriment to the vendor in the researcher not providing the information at the start,

---

[61] This analogizes to §512(f) of the DMCA, which has been used by courts to assess damages on parties who misrepresent their ownership of a particular copyright. *See* 17 U.S.C. § 512(f); *Automattic v. Steiner*, 82 F. Supp.3d 1011, 1026 (2015). Part of the underlying reasoning for that provision punishing those who submit unfounded takedown notices is to avoid overly burdening the parties processing the notices, as Automattic argued in the cited case.

[62] *See* Section 4.III.

because if the vendor responds expeditiously they will obtain the information soon thereafter. Third, vendors would not be required to answer Vulnerability Alerts if they believe in good faith that they are not legitimate.

We propose that the vendor should be required to send an automated receipt notification, so that researchers have confirmation that their vulnerability report is being reviewed. Requiring vendors to actively respond to every alert would be an unreasonable burden and would go far beyond any analogy to the status quo. Only requiring vendors to actively respond to alerts that they believe are legitimate analogizes to vendors only responding to emails or other communications that they believe are the product of bona fide vulnerability research, which mirrors the status quo.

Finally, we included the possibility that some parties will simply agree on the Classification, which would streamline the process significantly. For that reason, the vendor is encouraged to provide the researcher with further information in order to attempt to convince the researcher to amend her classification. As will be explained in the Contested Classification Protocol below, both parties benefit from agreement on a Classification, rather than both sides facing potentially suboptimal classification. The protocol also accounts for the possibility that researchers will engage intermediaries to prepare and communicate the Vulnerability Report. Coordinators, organizations that aid security researchers in identifying vendors, communicating with vendors, and disclosing vulnerabilities, are common actors in current disclosure practices.[63] The proposed reform places the burden of providing a sufficient Vulnerability Report on the researcher, but the researcher is absolutely able to engage a coordinator to assist in the preparation and communication. The coordinator would be bound by the same timeline as the researcher, barring them from disclosing the vulnerability to the public until the timeline associated with the relevant classification has elapsed.

c. Downstream Disclosure in Complex Supply Chains

Downstream disclosure presents another issue. With the complexity and inter-related nature of modern technology products, many discovered vulnerabilities will impact many parties who have integrated the relevant product into their larger infrastructure. Downstream disclosure, in this context, refers to notifying secondary vendors who rely on the product or service of the primary vendor, the organization responsible for maintenance and security of the product or service. We propose that the researcher's responsibility in the proposed reform is only to contact and disclose to the primary vendor, the direct manufacturer of the product containing the vulnerability, who is the party ultimately be responsible for patching that vulnerability. In our proposal, the primary vendor would be responsible for notifying downstream vendors of the existence of the vulnerability once it is confirmed. The extent and details of the primary vendor's downstream responsibilities would likely be dictated by contract in most cases, as they have likely sold or licensed their product or service to the secondary vendor, and those contracts would not interfere with the proposed solution here. If the primary vendor has agreed to notify secondary vendors regarding discovered vulnerabilities as part of their contract, or if they have agreed not to notify them but to push patches without notification, that agreement would be respected. In the absence of such an agreement, and if the secondary vendors would need to implement a patch on their own, the responsibility of the researcher is still to notify the primary vendor when possible. In fact, we

[63] *ISO/IEC 29147: Information technology – Security techniques – Vulnerability disclosure*, International Organisation for Standardisation (2014).

believe that the specifics of how information and mitigations would be dealt with in complex supply chains should and will be resolved by contract dynamics. Rather than imposing a brand new set of rules on a complicated ecosystem, we instead propose encouraging primary vendors to include (and secondary vendors to demand) terms in the relevant distribution contracts and terms of service that clearly lay out the parameters for downstream disclosure. The statutory reform described above has several relevant junctures that could be used as trigger points for requiring disclosure to secondary vendors, such as a vendor deciding to respond to a Vulnerability Alert in Step (5) of the Communications Process. If the relevant contracts include provisions requiring primary vendors to notify downstream vendors at that juncture, and require them to provide certain information or resources in order to effectuate a mitigation, that would manage the complexity of supply chains without imposing additional rules that risk being over or under-inclusive.

In terms of the Communications Process outlined above, we note that the finder will most likely only need to communicate with one vendor, the primary vendor. If the primary vendor is not available to the finder, the finder has the option of using a coordinator, in which case the Communications Process would run between the coordinator and the primary vendor. The case where the primary vendor is unavailable is addressed in Section 4.I.d, below.

As with the rest of the proposed reform, the researcher and vendor could circumvent the default of the primary vendor's responsibility to inform downstream vendors via contract. The researcher could form an agreement with the primary vendor that includes them taking on the burden of notifying secondary vendors in exchange for a more favorable publication timeline or for some other concession. Despite the existence of such a safe harbor, some vendors might still pursue a lawsuit in order to enjoin the researcher from disclosing the vulnerability to the public or for some other person. In such a case, the researcher would certainly have the safe harbor as a defense, and if they had complied with the conditions of the safe harbor, the vendor's suit would fail.

d. Post-Disclosure Issues

There are several other issues that might arise after attempted disclosure to the vendor. Two will be addressed here: nonresponsive or non-existent vendors and the chilling effects of potential litigation addressed via allocation of attorney's fees in potential lawsuits. Each will be addressed in turn.

First, in the case of a vendor that no longer exists or does not respond to a Vulnerability Alert, we propose that the researcher's responsibility is to undertake a good faith effort and exhaust all reasonable means. In the case where the vendor no longer exists, the researcher's responsibility is to undertake a good faith effort to find the party that would be most impacted by any publication of the discovered vulnerability, based on the logic that this would be the party most likely to bring suit under the CFAA or DMCA § 1201 should responsible disclosure not be attempted. If the vendor does exist and maintains a communication channel that they claim they monitor and respond to, attempting to use that channel should be sufficient under the standard. Otherwise, what qualifies as a good faith effort would be up to courts to decide, but we imagine it including browsing the vendor's website, sending an email to the support email address, or finding the most relevant line of contact available. If the researcher immediately contacts a coordinator, that coordinator has the same burden on behalf of the researcher. If the researcher meets the search and notification burden and still does not receive any response, they would still be eligible for the safe harbor and would be able to publish the vulnerability

without potential liability, based on the assumption that if the vendor does not exist or does not undertake enough care to respond to Vulnerability Alerts, disclosing the vulnerability to the public would serve to benefit the ecosystem at large including consumer safety.

Second, regarding potential lawsuits, the safe harbor protection alone does not resolve the fact that the burdens of even the early stages of a lawsuit may be enough to create chilling effects. In order to aid in resolving this issue, we propose that if a vendor initiates a lawsuit against a researcher who eventually prevails using the safe harbor as a defense, the vendor would be required to pay the researcher's attorneys fees. While this does not completely alleviate the burden on the researcher, as they are still forced to undertake the time and strain of the early stages of a lawsuit, it alleviates some of the burden and possibly serves as a disincentive when vendors are deciding, ex ante, whether to file a lawsuit: as losers of the lawsuit, vendors would be forced to bear that cost, potentially making vendors more risk averse in filing suits against researchers on the margin. The loser paying the winner's attorneys fees is not the default in the United States, but the convention does exist in some areas of U.S. law. Over "150 federal statutes authorize attorney fee shifting," and most do so in order to encourage or enable litigation in the public interest.[64] We believe the important policy rationales for granting attorneys fees to the winners are present in this use case: it would deter baseless litigation that would create social detriment, the statute is designed to create public benefit, and the parties in potential litigation have vastly unequal access to resources.

II. Vulnerability Classification

As is evidenced in Section 3, the method of negotiating a vulnerability publication deadline varies across guidelines as well as across vendor policies. In the case of Coordinated Vulnerability Disclosure (CVD), for instance, the finder often has very little influence on the eventual publication date, as the process requires the development of a remediation by the vendor before any vulnerability details can be made public. This affords the vendor significant influence over the publication deadline, which may, in some cases, affect the finder's livelihood. In the case of vulnerabilities being exploited in the wild at the time of discovery, some disclosure guidelines, such as the ISO/IEC 29147[65] standard, advise coordinated release of a limited advisory detailing the existence of the vulnerability, and short-term countermeasures that project users; full vulnerability details are not released. Although this may allow for finders to claim discovery, it may still impede publication desires. It is potentially possible, of course, that the finder and/or coordinator controls the publication deadline. This is true in the case of the CERT Coordination Center (CERT/CC)[66], and is implied by the Netflix disclosure policy. It is most certainly true in the Full Disclosure (FD) model, as in the IETF[67] recommendations, which allow for the finder to specify a fixed publication date, in turn granting the vendor a specified window in which

[64] Robert R. Percival & Geoffrey P. Miller, *The Role of Attorney Fee Shifting in Public Interest Litigation*, 47 L. & Contemp. Probs. 233 (1984).
[65] *ISO/IEC 29147: Information technology – Security techniques – Vulnerability disclosure*, International Organisation for Standardisation (2014).
[66] The CERT/CC allows for a fixed window of 45 days before disclosing vulnerability details, however, the organization will negotiate different timelines if need be.
[67] Steve Chistey & Chris Wysopal, *Responsible Vulnerability Disclosure Process – Internet Draft*, Internet Engineering Task Force (2002).

to patch the vulnerability before releasing vulnerability details. However, in these cases, the finder may be subject to legal action and thus in reality, arguably has very little control of the publication date.

Very rarely does the negotiation process include a formal classification of the vulnerability in question. As laid out in the previous section, our proposed legislative framework for vulnerability disclosure includes the vendor and the finder agreeing on a vulnerability classification. This classification will result in a severity score which in turn will dictate the publication timeline. In order to introduce a sense of "fairness" into the timeline negotiation process, we propose the use of a *contributive score* - arriving at this score entails the finder and the vendor both contributing inputs to the severity score. The aim of such a scoring process is to provide both parties with the opportunity to have influence on the publication deadline, taking their respective incentives and needs into account. A process such as this also standardizes the negotiation of publication timelines, normalizing the disparate approaches currently employed in practice.

a. Classification Protocol

We now present our proposal for the Classification Protocol, allowing for both the finder and the vendor to contribute to the severity score. We base our protocol on the Common Vulnerability Scoring System (CVSS) version 3.0,[68] developed by FIRST.Org. This is the scoring system used in the Common Vulnerabilities and Exposures dictionary,[69] an extensive collection of standardized cybersecurity vulnerability names and descriptions.[70] CVSS is one of the more popular means by which vulnerabilities can be classified. We address whether or not it is the most suitable scheme in Section V, and appreciate that classification schemes may vary over time, potentially causing update difficulties if a particular scheme has been codified into law. We therefore propose that our statutory reform be flexible enough to allow for any classification scheme permitting finder-vendor negotiation to be used. At this instance, we proposed that the classification scheme employed be the one outlined below.

The CVSS scoring system is composed of three scoring metric groups, namely the Base, Temporal, and Environmental groups. We explain each of these groups below:

> **Base Metric Group.** This group covers characteristics that are intrinsic to a vulnerability and immutable over time. The Base group is divided into two sets of metrics, namely the Exploitability metrics and the Impact metrics. The Exploitability metrics reflect the ease and technical means by which the vulnerability can be exploited. They relate to the component that is vulnerable and exhibits a weakness that can be exploited. The Impact metrics reflect on the direct consequence of the vulnerability being exploited, and covers the component(s) affected by the exploit. The weak component may be the affected component, however, this might not always be the case. For instance, a weakness in a software application may impact a hardware device or network resource.

---

[68] *Common Vulnerability Scoring System v3.0: Specification Document*, Forum of Incident Response and Security Teams, Inc. (FIRST.Org), available at https://www.first.org/cvss/specification-document.
[69] *Common Vulnerabilities and Exposures – The Standard for Information Security Vulnerability Names*, MITRE, available at https://cve.mitre.org/.
[70] We note that router manufacturer Cisco uses this system to score vulnerabilities received from external finders as part of its vulnerability disclosure policy. *See* Cisco, *Security Vulnerability Policy*, available at http://www.cisco.com/c/en/us/about/security-center/security-vulnerability-policy.html.

Temporal Metric Group. The Temporal metric group "measures the current state of exploit techniques or code availability, the existence of any patches or workarounds, or the confidence that one has in the description of a vulnerability".[71] Unlike the Base characteristics, these attributes may change over time - exploit kits may be developed if not already in existence, as is true for patches and workarounds.

Environmental Metric Group. This metric group reflects the importance of the affected component(s) within an organization. In other words, this metric takes into account the environment in which the vulnerability may be exploited and allows for the customization of the severity score based on the impact to Confidentiality (C), Integrity (I), and Availability (A)[72] of the potentially vulnerable systems. At a high-level, computation of this metric includes measurement of the security requirements of the potentially affected components and systems and incorporates a modified Base Impact score based on the environmental factors influencing the component's or system's operation.

We refrain from covering all of the individual elements within the sets of metric groups mentioned above and direct the reader to CVSS 3.0 specification document for further details. For the purposes of this work, a high-level understanding of the metric groups is sufficient.

In order for a finder and a vendor to agree on a publication timeline, as required in steps five through eight in the Communications Process outlined in the previous section, we propose employing what we term the Simple Vulnerability Classification (SVC) protocol, as depicted in Figure 1.

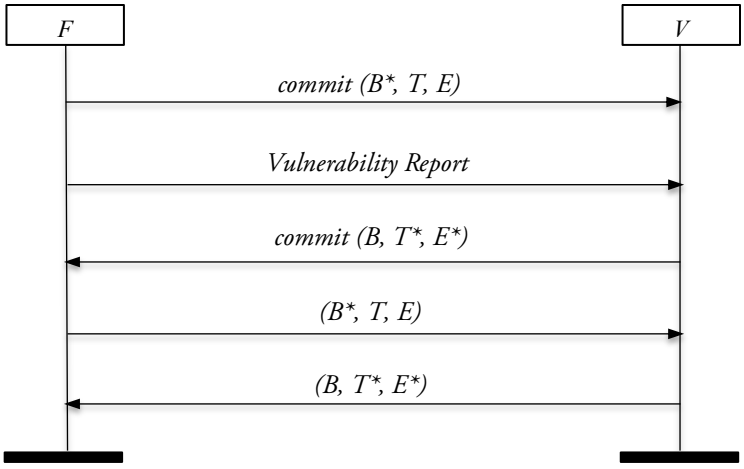

FIGURE 1: THE SVC PROTOCOL

In the SVC protocol, the finder determines the Base (B), the Temporal (T) and Environmental (E) scores relating to the discovered vulnerability. At this stage, each score is a vector indicating an agent's

[71] *Common Vulnerability Scoring System v3.0: Specification Document*, Forum of Incident Response and Security Teams, Inc. (FIRST.Org), available at https://www.first.org/cvss/specification-document.
[72] We defer to the definitions of confidentiality, availability and integrity as presented in the CVSS 3.0 specification document.

evaluation of the set of metrics within each metric group. The final severity score is a numerical value, ranging from 0 to 10, and is computed as a function on all three metric group vectors - calculation details can be found in the CVSS 3.0 specification document.[73] The finder (F) then blindly commits to her three scores (vectors), meaning that the vendor cannot see the scores, and sends a commitment value to the vendor (V), along with the detailed Vulnerability Report, which may include attack implementation code, or a Proof-of-Concept (PoC) of some sort. Upon receiving the detailed Vulnerability Report, the vendor computes the Base score, the Temporal score and the Environmental score. The vendor then blindly commits to the computed scores and sends the commitment value to the finder. Note that the exchange of these commitments, and all subsequent messages, will take place within the secure channel established between the finder and the vendor (see step two of the Communications Process).

After both commitments have been received, the finder and the vendor both send their score values to the respective receiving party. The final severity score is computed using the finder's Base score and the vendor's Temporal and Environmental scores, resulting in what we term the *contributive score*. The rationale for computing the severity score in this manner is that we believe both parties should have a degree of influence on the publication timeline, opposed to one party dictating public release of the vulnerability details, as is done in CVD, where vendors largely control the publication schedule, and in the case of finder- or coordinator-controlled release scenarios, as dictated by some policies and guidelines. Also, we believe that it is appropriate to assume that the finder has a detailed knowledge of the intrinsic characteristics of the vulnerability, thus making this party best suited to calculate B, and that the vendor most likely has superior knowledge with regards to the environmental, operational and customer usage factors relating to the vulnerability, hence placing the vendor in the best position to calculate Temporal and Environemnt.

As the final severity score is computed from a combination of inputs from both of the interested parties, we restrict one party's score from being visible prior to computation of the other party's score so as to prevent manipulation of subsequent scores with the purpose of favorably influencing the publication timeline. The exchange of the first four messages as described above is intended to achieve this, and what we have described, at a high-level, is known as a *commitment scheme*. The initial exchange of values is known as the *commit phase* and the subsequent exchange of the scores is known as the *reveal phase*. Such schemes have been well studied in the field of cryptography.[74] It is also important that both parties receive the commitment "(correct)" to their score values, as well as "(correct)" on the score values themselves. The general problem here is known as the problem of fair exchange: both parties want to ensure that they get the other party's promised item, or they both receive nothing. There are elegant solutions in the literature, including gradual release and Concurrent Signatures.[75] There is an impossibility result that says, roughly, that a third party is necessary.[76] In the "optimistic fair exchange"

---

[73] We note the existence of an easy-to-use CVSS 3.0 calculator on the FIRST website, available at https://www.first.org/cvss/calculator/3.0. This would allow for the convenient computation the final score by both parties.

[74] In the very basic case, one could consider the $commit(B, T, E)$ message to be a cryptographic hash value of (B, T, E).

[75] Liqun Chen, Kudla Caroline & Kenneth G. Paterson, *Concurrent Signatures*, EuroCrypt (2004).

[76] Henning Pagnia & Felix C. Gartner, *On the Impossibility of Fair Exchange Without a Trusted Third Party*, Darmstadt University of Technology Technical Report TUD-BS-1999-02 (1999).

setting, the third party needs only be involved in the event of a dispute. We consider disputes in the next section.

In CVSS' specification lists the T and E scores as optional. Calculation of these scores is mandatory in our proposed solution. Additionally, the final severity score could be calculated in a number of ways - scores could be weighted, for instance, based on the role of the agent computing them. Our solution offers one possible method for classification. It is of course also possible to create a different scoring system for the purposes of determining a publication timeline, one that might be more nuanced and disclosure-focused. Again, we highlight that our solution is one of many possible solutions. Use of this scoring system, however, does aid in the registration of a CVE identifier (at the time of public release), if applicable.

b. Score-Dependent Publication Schedule

The output of the CVSS is a score in the range of 0 to 10, inclusive. The CVSS specification maps the following qualitative severity ratings to the range of possible severity scores (see Table 1):

TABLE 1: CVSS 3.0 RATINGS

CVSS Score	Rating
0.0	None
0.1 - 3.9	Low
4.0 - 6.9	Medium
7.0 - 8.9	High
9.0 - 10.0	Critical

We now propose a publication schedule based on these ratings (Table 2), where publication refers to public release of the vulnerability details:

TABLE 2: PROPOSED PUBLICATION SCHEDULE

Rating	Time to release (days)
None	0
Low	35
Medium	62
High	76
Critical	90

Parties may choose to coordinate disclosure, as recommended in the CVD model, but in our scheme, after the suggested time period has lapsed either party is free to divulge the vulnerability details. As noted earlier, in many of the vulnerability disclosure guidelines, attacks may be ongoing in the wild at the time of discovery by the finder. These guidelines recommend the issuing of an advance security advisory with limited vulnerability details but including suggested workarounds and countermeasures to protect users. At this stage, we leave this decision regarding whether or not to take this action up to the parties involved and do not provision for this in our publication schedule.

c. *Contested Classification*

In the case of dispute regarding the B, T and E scores, our solution calls for the involvement of a Trusted Third Party (TTP). We recommend that a TTP agreeable to both parties be selected. To prevent the manipulation of scores by a party so as to influence the final score in a favorable manner, we introduce a penalty for exaggeration: if a party disagrees with the scores computed by the other party in the negotiation, and no settlement can be reached, the parties should submit all details to a TTP, including the vulnerability details, and request judgment on the B, T and E scores. Based on the vulnerability details provided, the TTP should compute all components of the CVSS severity score (B,T and E), and if the TTP finds that one of the parties has exaggerated their scores, then the final score is calculated using the B, T and E values of the other party only. This is why we call for all scores to be computed by both the finder and the vendor. We hope that the threat of losing influence over the publication schedule entirely will keep parties honest. In the event that the TTP rules that no exaggeration has taken place, then the score remains as originally calculated. We note that there are already agents in the disclosure ecosystem that can take on the role of TTPs, namely coordinators. It is also possible, if this solution gets adopted, that other agents will emerge as possible and willing TTPs, such as computer security experts or research and government institutions. We note that the TTP has no publication rights, and that it is bound by confidentiality surrounding the vulnerability details. Irresponsible release of details by the TTP would result in reputational harm as well as possible legal action by the vendor, as the TTP is not subject to the safe harbor in our proposed framework.

The classification scheme suggested above is to be considered the default means by which a finder and vendor agree on a publication timeline; it is entirely possible for these parties to agree to negotiate a publication timeline via some other means, if they so choose. In this case, if the parties agree to make use of the remaining safe harbor components as per the Communications Protocol, these components apply via contract. We are aware the CVSS is geared towards addressing software vulnerabilities. We consider this system, and therefore our classification scheme, to be equally applicable to most hardware vulnerabilities, potentially with some additional attributes and caveats, as discussed in Section 5.III.

5. COUNTER ARGUMENTS

The proposal for a responsible disclosure driven safe harbor is obviously not uncontroversial. In this section we will address several typical arguments that would be presented in opposition to such a regime. While there are a myriad of typical policy arguments that might be made to refute to our proposal, most are addressed by the specifics of our implementation. To address some here: the argument that the government does not have the expertise to regulate vulnerability disclosure is addressed by the private-party-focused deliberation structure. The argument that malicious actors will use a responsible disclosure regime to acquire a legal safe harbor for their activity is rebutted by carefully considered definitions that make clear that only designated research activities are captured within the scope of the safe harbor.

There are, however, some counter-arguments that merit more detailed analysis. The prominent legal argument that will be addressed is that government-mandated responsible disclosure might be a prior restraint on speech in violation of the First Amendment. We will also address the policy argument that a legislative reform is not practically feasible, and briefly consider alternatives such as a Department of

Justice prosecutorial discretion advisory or a consumer protection advisory from the Federal Trade Commission. The prominent technical counter-arguments include our choice of classification scheme, that the solution does not account for the nuanced differences between software and hardware, and that the solution is overly complex and laborious for security researchers. Each counter-argument will be addressed in turn.

I. Responsible Disclosure as a Violation of the First Amendment

Some have argued that the First Amendment of the U.S. Constitution, which protects free speech, among other things, makes government-driven responsible disclosure unconstitutional. The most powerful version of this argument is that a legislative regime of responsible disclosure, which necessarily requires security researchers to delay publication of information they already possess, qualifies as a prior restraint on protected speech, which is at the core of First Amendment protection. It is clear from jurisprudence that, in most cases, prior restraints violate the First Amendment.[77] Delays, not only prohibitions, on publication still trigger the doctrine of prior restraints.[78] Kristin Bergman argues that restricting the publication of security vulnerabilities in a responsible disclosure regime meets all the conditions required to qualify as a prior restraint doctrine. While the data acquired by exploiting a vulnerability is very likely not protected speech, disclosing the vulnerability itself to the public might be, according to Bergman. She cites a Fourth Circuit precedent in which the publication of Social Security Numbers was protected expression even though the numbers themselves were not, due to the fact that the context of publication was an attempt to convey a message about how the government had been handling personal information.[79] There are narrow exceptions to the prior restraints doctrine, famously laid out in *Near v. Minnesota* where, though the Supreme Court held a law unconstitutional, it presented the potential that a prior restraint for the legitimate purpose of protecting national security might be constitutionally permissible,[80] but that burden has proven highly difficult to meet.[81] Bergman also argues that data security is a matter of public concern, and that restricting publication of vulnerabilities can cause damage to users based on their not having the relevant information to take precautions, making disclosure of vulnerabilities more likely to be recognized as protected expression.[82] This argument, while compelling, rests on several empirical assumptions and refutable legal principles.

First, claiming that responsible disclosure causes more damage to users than full disclosure relies on two assumptions: that full disclosure will make manufacturers more likely to patch vulnerabilities than mandated responsible disclosure and that malicious actors taking advantage of fully disclosed, unpatched vulnerabilities do less damage than the unpublished vulnerabilities in the time between finding the vulnerability and disclosing it in the responsible disclosure context. While there has been some evidence[83] as to the assumption that vendors are more likely to patch in full disclosure scenarios, much of this evidence is from the older days of the online ecosystem, has several mitigating factors,

---

[77] Kristin M. Bergman, *A Target to the Heart of the First Amendment: Government Endorsement of Responsible Disclosure as Unconstitutional*, 13.2 Nw. J. L. & Tech. 117, 127–129 (2015).
[78] *See Neb. Press Ass'n v. Stuart*, 427 U.S. 539, 559 (1976).
[79] *See generally Ostergen v. Cuccinelli*, 615 F.3d 263 (4th Cir., 2010).
[80] *See Near v. Minnesota*, 283 U.S. 697, 716 (1931).
[81] *New York Times Co. v. United States*, 403 U.S. 713 (1971). *See United States v. Progressive*, 467 F. Supp. 990 (W.D. Wis., 1979).
[82] Bergman, *infra* note 77, at 133.
[83] Ashish Arora, Ramayya Krishnan, Rahul Telang & Yubao Yang, *An Empirical Analysis of Software Vendors' Patching Behavior: Impact of Vulnerability Disclosure*, Heinz School of Public Policy and Management, Carnegie Mellon (2006).

and, most importantly, is based on our current legal regime in which there is an established incentive for vendors to ignore responsible disclosure timelines because of their ability to bring lawsuits regardless of disclosure. The second point is even more objectionable. As vendors have accumulated more data and control over the lives of users, the danger of malicious actors exploiting vulnerabilities between full disclosure and the (potentially rushed and imperfect) patch has grown exponentially.

Our particular legislative implementation of responsible disclosure also likely avoids the constitutional pitfalls of the prior restraint doctrine. First of all, it is not in fact a legislative restriction or delay on protected expression. Rather than mandating that all researchers must delay publication until the relevant period has elapsed, the proposed regime grants a benefit in the form of a safe harbor to those who do. If the researcher opts for full disclosure and releases the details of the vulnerability immediately, they are not punished under the proposed statutory regime, they simply forego a potential benefit. Those researchers are subject to the same legal regime that they would have been before our safe harbor existed. Their position has not changed at all relative to the status quo. Our proposal simply seeks to add an incentive to engage in responsible disclosure. Second, our proposal does not involve a government-enforced solution. Only government restraints on speech are violations of the First Amendment. Like the safe harbors in §512 of the DMCA,[84] our proposal creates a private mechanism for the relevant parties to manage delays in publication and creates default rules for those interactions. Rather than the government setting direct timelines and penalties, the private parties go through a designed process to negotiate and decide on the outcome. It has been suggested that DMCA §512's safe harbor would be a violation of the First Amendment if the process went through a court, rather than the back and forth occurring between two private parties, but that the private nature of the mechanism shields it from that potential unconstitutionality.[85] Our proposal operates similarly. While the government might be implicitly agreeing not to take action, such as bringing prosecution under the CFAA and DMCA §1201, the government is not engaging in the restraint on speech. The private parties interact, decide on the vulnerability classification and therefore on the timeline, and a court is simply enforcing the agreement as they would if the process under DMCA §512 broke down. This points to a crucially important benefit of the proposed system: by eliminating liability under statutes that criminalize the investigation activity (as opposed to the publication), this reform does not seek to directly regulate the publication at all, focusing instead on the activity that happens well before any potential publication.

There are also larger, more system-oriented legal defenses of the proposed regime. As already discussed, both DMCA §1201 and the CFAA have been widely decried as restricting and chilling speech. §1201 has been used as the basis for content takedowns and harsh penalties, such as Sklyarov's arrest and a court ruling that DeCSS had to be removed from online circulation.[86] The CFAA has been used as grounds for temporary restraining orders and criminal prosecutions based on published research. Wendy Seltzer investigates in detail the chilling effects of the DMCA and their implication on the First Amendment and on free speech. A group of cryptographers and security researchers filed

---

[84] Section 512 of the DMCA provides a safe harbor for online service providers from being held liable for copyright infringement based on content that users upload, as long as the providers comply with several conditions, including requiring them to take down content when a copyright owner submits a good faith notice to the service provider that they believe the relevant content to be infringing. *See* 17 U.S.C. § 512.

[85] Wendy, Seltzer, *Free Speech Unmoored in Copyright's Safe Harbor: Chilling Effects of the DMCA on the First Amendment*, 24.1 Harv. J. L. & Tech. 171, 176 (2010).

[86] *Universal City Studios, Inc. v. Reimerdes*, 82 F. Supp. 2d 211 (S.D.N.Y., 2000).

an amicus brief in a case involving the publication of a cryptographic tool claiming that DMCA §1201 implicates a large portion of security research activity, creates chilling effects, and is unconstitutional if read as widely as some courts have suggested. There are strong arguments that the DMCA goes further in restricting speech than a responsible disclosure safe harbor does. If the DMCA system is constitutional, based on its implementation and it's placing the burden on private parties, then our proposal fits that same mold. Also, there is significant literature signaling that safe harbors based on clear conditions have had a positive impact on the development of the Internet and on free speech.[87] The two oft-cited safe harbors in the online ecosystem, §512 of the DMCA and §230 of the Communications Decency Act (which provides a safe harbor from defamation and other publisher-related claims for online services) have been hailed as having been "among the most important protections of free expression in the United States in the digital age."[88] The §512 system that is cited as having allowed the online ecosystem to thrive is a system in which a copyright owner can request for content to be immediately taken down, a request which the online service provider must honor. That regime is undoubtedly restricting speech, but it is doing so in order to protect the legal rights of one party and to allow the larger online system to operate smoothly. The claim that a safe harbor which, in some carefully calculated circumstances, temporarily delays speech has a necessarily detrimental impact on speech ignores historical precedent.

To return to Bergman's argument, she lays out conditions under which statutorily imposed responsible disclosure might survive constitutional scrutiny: (1) "place the burden on the data-holder to show the disclosure implicates unprotected expression," (2) ensure the holder's decision does not result in finality, (3) "limit the delay to a brief period," and (4) "provide for expedited judicial review."[89] Our proposal takes all of those concerns into account. The vendor should have the burden of proof in a contested classification dispute in front of a TTP, the proposed contested classification resolution system should ensure that the vendor does not have final decision-making power, all of our proposed timelines are intentionally limited, and we would suggest expedited review in relevant circumstances. Even under Bergman's framework, our proposal could pass scrutiny. Bergman's argument that responsible disclosure chills speech is based on the belief that, by responsibly disclosing, security researchers are at the mercy of companies who may or may not patch and who may or may not bring suit. She cites the case of a professor at Purdue who no longer discloses vulnerabilities after a student found a vulnerability in a school system and the authorities threatened the professor with suit months later, after a potentially unrelated hack.[90] Our safe harbors completely defeat that underlying necessary component of Bergman's logic. By providing a safe harbor, there is no longer the specter of suit, and no longer the chilling effect that Bergman attributes to responsible disclosure. The chilling effect on speech is produced by overbroad laws, not by responsible disclosure.

---

[87] *See* Jack M. Balkin, *Old School/New School Speech Regulation*, 127 Harv. L. Rev. 2296 (2014); Mark A. Lemley, *Rationalizing Internet Safe Harbors*, 6 J. of Telecom. and High Tech. L. 101 (2007); David Ardia, *Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity Under Section 230 of the Communications Decency Act*, 43 Loyola of L.A. L. Rev. 373 (2010) (discussing the safe harbor in another statute, the Communications Decency Act).

[88] Balkin, *infra* note 87, at 2313.

[89] Bergman, *infra* note 82, at 139

[90] Scott Berinato, *Software Vulnerability Disclosure: The Chilling Effect*, Christian Science Monitor (2007), available at http://www.csoonline.com/article/2121727/application-security/software-vulnerability-disclosure--the-chilling-effect.html.

II. Legislative Solution as Impractical

The difficulty of convincing the U.S. Congress to undertake any significant action, particularly in an area such as this, opens proposing a legislative reform solution to a clear and obvious objection: it is unrealistic as a practical matter to expect such a reform to actually take hold.[91] There is serious merit to that argument. It is unlikely that the current Congress would prioritize such a reform and it would be incredibly difficult to rally the political will necessary. Our aim is to propose what we believe to be the ideal and most complete solution. A statutory solution provides the most complete protection for security researchers, provides the best incentives for vendors to effectuate patches, and provides the most certainty for vendors that they have a period during which to mitigate the vulnerability without the public being aware of it. It is also worth noting that a legislative fix may be necessary, as evidenced by the fact that the Copyright Office decided against a broad security research exception to DMCA §1201 at least in part due to the fact that it felt it should allow multiple agencies of government to be involved in such a decision.[92]

There are, however, intermediate options by which a similar regime could be implemented without Congressional action. Three in particular bear emphasis. First, this could be effectuated by bringing the major private parties together in order to create a near-universal bug bounty program that resembles our proposed reform almost exactly. Second, the Department of Justice (DOJ) could issue an advisory to all federal prosecutors that they should not prosecute security researchers who comply with the set of best practices as we have described them. Third, the Federal Trade Commission could issue a regulation or advisory claiming that security researchers that comply with our responsible disclosure regime should be protected under a consumer protection regime. Each will be examined in turn.

If the major technology companies, the major security research groups, and other interested parties could convene and mutually agree to this regime, that might operate as a private solution to avoid requiring action from Congress. The stakeholders could agree that if security researchers comply with the same conditions laid out in our proposed statutory reform, the vendors would agree not to bring any civil actions or file any criminal complaints under either of the two statutes. The parties could also agree to use the same classification system proposed above.[93] Unfortunately, this approach faces limitations. First, the stakeholders are disparate and in no way united. Security researchers do not all operate under any umbrella organization and building consensus among them would be as practically unlikely as building consensus among members of Congress. Achieving consensus among technology industry actors is equally difficult, and any defectors in the system would undermine it completely. If any one major player declines to implement the safe harbor-style regime in a private context, all security researchers would be disincentivized from attempting to discover vulnerabilities in their system for fear of legal action. Second, federal prosecutors could bring charges against security researchers despite this private agreement. The private agreement does not change the overbroad interpretations of the

---

[91] Aside from the obvious current political issues, Congress has generally been passing fewer and fewer pieces of legislation overall for the last several years. The 114th Congress enacted 329 laws, whereas the 100th Congress enacted 761. Trend data available at https://www.govtrack.us/congress/bills/statistics.

[92] *See* U.S. Copyright Office, *Section 1201 Rulemaking: Sixth Triennial Proceeding to Determine Exemptions to the Prohibition on Circumvention*, 317 (Oct. 2015).

[93] There is some precedent for private stakeholders to assemble and agree to best practices in this manner. American University's Center for Media and Social Impact convened stakeholders to agree on best practices for fair use in documentary filmmaking. Available at http://cmsimpact.org/code/documentary-filmmakers-statement-of-best-practices-in-fair-use/.

two relevant statutes, nor does it change the murky legality of the research activities. While a CFAA prosecution would be made much more difficult because the security researcher would have been granted authorization by the vendor, therefore they would not have exceeded authorization,[94] a DMCA §1201 charge is still viable. Despite those limitations, this intermediate solution maintains appeal. It could remove the chilling effects on security research as it relates the products of every firm that joins the private agreement, which would still operate as an improvement on the status quo.

Another potential implementation of the responsible disclosure safe harbor-style regime is through the Department of Justice (DOJ) issuing a memorandum to all federal prosecutors informing them that the new policy is to not prosecute security researchers who comply with the disclosure regime laid out in our proposed statutory reform. The DOJ has the authority to issue prosecution memorandums to U.S. Attorney's offices around the country, meant as guiding documents for when prosecutors are making discretionary decisions about which cases to pursue.[95] A DOJ prosecution guideline advising prosecutors not to pursue security researchers who have complied with the responsible disclosure best practices we proposed above would accomplish many of the same results as the statutory safe harbor. It would avoid the majority of criminal prosecution of security researchers for their regular activity, reduce the chilling effects, and provide more certainty for researchers. Unfortunately, such a DOJ guideline does not eliminate all potential litigation risk for security researchers. First, DOJ prosecution memoranda are not universally followed by federal prosecutors.[96] U.S. Attorneys might ignore or push the boundaries of the guideline. If that occurs in even only a small number of cases, the information would likely be disseminated within the security research community, creating some fear of prosecution. Second, vendors could still initiate civil lawsuits using either relevant statute. The CFAA and DMCA §1201 both explicitly create private rights of action, meaning that private citizens can initiate lawsuits with monetary damages or an injunction as the remedy (rather than a criminal sentence).[97] Civil suits can create a significant chilling effect independent of the potential criminal conviction, based on the resources required to litigate and the potential injunction barring publication or monetary damages the researcher would be forced to pay, as well as the reputational harm of being a party to a lawsuit.

A third intermediate, non-Congressional avenue is to use administrative agency action to accomplish the same goals and strategies. The Federal Trade Commission (FTC), has a mandate of consumer protection including language in the Federal Trade Act, which empowers the FTC to regulate "unfair or deceptive acts or practices."[98] That has been construed by the FTC, and subsequently by courts, to include the authority to regulate issues of consumer privacy and data security.[99] The FTC could approach attempting to enforce a responsible disclosure safe harbor for security researchers. The commission could pursue privacy enforcement actions against vendors who file civil lawsuits against researchers who comply with the identified responsible disclosure best practices, aiming to disincentive the lawsuits that generate chilling effects. The commission could also publish the best practices and

---

[94] *See* 18 U.S.C. § 1030(a)(2).
[95] For a more detailed account of these guidelines, *see* Ellen S. Podgor, *Department of Justice Guidelines: Balancing "Discretionary Justice"*, 13 Cornell J. of L. and Pub. Policy 167 (2004).
[96] *See id.*, at 175.
[97] *See* 18 U.S.C. § 1030(g); 17 U.S.C. § 1201.
[98] 15 U.S.C. § 45(a).
[99] *See* Greg Dickenson, *Privacy Developments: TCPA Litigation, FTC Privacy Enforcement Actions, and the FTC's Internet of Things*, 71.1 The Business Lawyer 293 (2016); *see FTC v. Wyndham Worldwide Corp.*, 10 F. Supp. 3d 602, 612-15 (D.N.J., 2014).

advise other agencies to facilitate those practices. Unfortunately, the FTC has no binding authority over the DOJ, and therefore could not halt or prohibit prosecutions of security researchers who comply with the relevant best practices. FTC strategies for implementation likely work optimally in conjunction with the DOJ issuing a prosecution guideline.

While none of these intermediate solutions operates as a complete safe harbor in the manner we believe to be ideal, that does not mean they should not be pursued. In a political reality where Congress is unlikely to take action on any proposal approximating ours, these intermediate options should be adequately considered.

III. Choice of Classification Scheme

We next address the question of whether or not the CVSS (v 3.0) is most suited to our needs. We also consider the manner in which it is employed as part of the SVC protocol. As laid out in Section 4.II, the finder's computation of the Base score, and vendor's computations of the Temporal score and Environmental score, respectively, contribute to the calculation of the final severity score. This score, in turn, dictates the publication schedule. Our choice of the CVSS rests on the fact that the scoring system is a standard, well-understood and well-used system within the software security research community. There are also online tools available that document the scoring system and make scores easy to compute.[100] The Base score captures the intrinsic attributes of the vulnerability in question and the remaining scores capture the environment in which the vulnerability could be exploited, as well the ease of exploitation. These groups together make up a satisfactory collection of metrics for assessing the severity of a vulnerability. Our choice of split regarding the parties responsible for each respective computation is intended to fairly distribute the balance of power in determining the publication timeline, and is rooted in the assumption that the finder understands the intrinsic characteristics of the vulnerability, and that the vendor may have a better knowledge of the environmental consequences of the vulnerability. This, of course, may not be the case. Our solution does not prohibit the finder and the vendor from amicably running the protocol a number of times after discussing the vulnerability and the respective metric computations. In the case of dispute, our solution calls on a TTP for resolution. We feel that the requirement to compute all scores (B, T and E), also contributes to the sense of "fairness" in the negotiation process as the T and E scores are designed to allow for adjustment of the final score based on factors beyond the intrinsic characteristics of the vulnerability, thus allowing the vendor to highlight consequences of an exploit which may influence the mitigation development time.

As noted previously, the engaged parties may decide to forgo using the proposed negotiation protocol and determine a publication schedule via some other method. In this case, the parties may contract around the default solution. Safe harbor may still be provided if the parties agree that it is, and if the basic premise of the regime was still followed: the finder waits a predetermined, specified period before disclosing the vulnerability to any party other than the vendor. However, our solution provides a default means by which safe harbor will be provided; following our proposed protocol ensures safe harbor.

---

[100] *See* Common Vulnerability Scoring System Version 3.0 Calculator (FIRST.Org), available at https://www.first.org/cvss/calculator/3.0.

It is also possible that a more nuanced approach to determining the publication schedule exists, such a protocol that weights scores based on the role of the agents, or on some other characteristics, or indeed a protocol that employs more advanced cryptographic mechanisms, such as those suggested in the previous section. Our solution provisions one possible negotiation protocol to serve as the default means of timeline negotiation, it may be replaced by a more suitable protocol in the future.

It can be argued that this classification scheme, as used in the SVC protocol, is more suited to vulnerabilities of the software variety, given that the majority of CVEs concern such security vulnerabilities. However, we believe that the scheme is equally suitable to hardware security vulnerabilities, as remarked by the CVSS 3.0 specification itself: "Software, hardware and firmware vulnerabilities pose a critical risk to any organization operating a computer network, and can be difficult to categorize and mitigate. The Common Vulnerability Scoring System (CVSS) provides a way to capture the principal characteristics of a vulnerability, and produce a numerical score reflecting its severity".[101] We recognize that the publication timelines in the hardware case may differ from the software case as pushing an update to hardware may not always be possible, and the recall of devices may be infeasible; it is possible that the patch to a vulnerability will be only included in the next model of the device. We address the disclosure process in the hardware setting, as opposed to the software setting, in the next section.

a. Hardware vs. Software: Timeline Discrepancies and Disclosure Complications

As mentioned above, we find use of the CVSS to be reasonable in the classification of hardware (firmware) vulnerabilities. However, we do note that the publication schedules may differ for hardware owing to the difficulty in releasing mitigations. However, we would still like our solution to be satisfactory for finders when it comes to publication timelines. In the event that a vendor cannot provide a patch within the time window specified by our negotiation protocol citing legitimate reasons (such as device recall infeasibility, for instance), we propose the following:

The finder and the vendor should issue an advisory (with limited details) 14 days after the conclusion of the negotiation process. The advance advisory must include workarounds that protect users of the device, and the vendor must actively work to establish these workarounds within the 14 days leading up to publication of the advance advisory, potentially with help from the finder. The finder may publish the vulnerability details after the maximal time period specified in the publication schedule, 90 days.

Of course, again, if either party is not satisfied with the claims made by the respective peer in the negotiation process, a TTP should be called upon to settle disputes.

IV. **Is Publication Always Possible? The Case of Safety-Critical Devices**

We recognize that there may be conflicting opinions within the security community, and further afield, surrounding the publication of vulnerabilities connected to safety-critical devices. On the one hand, the risk of loss of life and/or significant financial losses to individuals may be deemed too great in the event that publication should proceed. On the other hand, publication of such vulnerability details

[101] *Common Vulnerability Scoring System v3.0: Specification Document*, Forum of Incident Response and Security Teams, Inc. (FIRST.Org), available at https://www.first.org/cvss/specification-document.

would better inform end-users of potentially dangerous products and may force manufacturers to address the problem with urgency, resulting in a safer security ecosystem.

For example, a pacemaker falls into this contested category. The "Early Stage" Coordinated Vulnerability Disclosure Template (v 1.1) of the NTIA Working Group[102] attempts to address disclosure for such devices: "[a]ny hard deadline for disclosure or remediation may both be too long and too short to safely address security vulnerabilities in safety-critical systems."[103] The infeasibility of establishing a hard deadline is also true in the case where significant financial loss may result as a consequence of exploiting a vulnerability. We believe that the solution presented above for hardware vulnerabilities also fits the class of safety-critical devices: A workaround should be put in place by the vendor and the finder may publish within the maximum timeframe of 90 days.

In order to address this class of devices, as well as other devices for which hard timelines may not be established, we adapt step (10) of the Communications Protocol to read:

> If the parties do agree, the vendor releases and advance advisory within 14 days and implements workarounds so as to ensure that the vulnerability is no longer deemed to be safety-critical, nor likely to result in significant financial loss to users, and the maximum timeline as dictated by the Classification protocol becomes binding.

The NTIA disclosure template recommends vendors publishing a list of devices which are open for investigation as part of their respective disclosure policies, thereby rendering certain devices out of bounds. Our solution deviates from this in that we wish to provide finders with safe harbor regardless of the device in question, and over and above the protections offered by §1201 of the DMCA.

In other words, our solution aims to encourage publication in all cases, following disclosure in a responsible fashion as per our framework. However, when it is absolutely infeasible to implement workarounds or mitigations of the vulnerability, and the vendor can prove as much, then the parties may consider negotiating to circumvent the defaults set by our proposed solution and coming to some other agreement.

V. A Laborious Solution

It is possible that the solution presented may be viewed as cumbersome and complicated by the parties wishing to engage in responsible disclosure. Again we state that parties are free to contract outside of the default terms suggested in our framework; we provide a default for safe harbor only. Also, we note that several of the steps mentioned in our Communications Process align with steps mentioned in vulnerability disclosure standards and guidelines, specifically steps (1) through (6) which involve the initial communication regarding a discovered vulnerability, as well as the sharing of vulnerability details. The additional computational element in our solution is the classification of the vulnerability and the determination of a publication schedule via the SCV protocol. This may give rises to fears of potentially unnecessary extra computation on the part of both agents. We note, however, that parties can make use of the publically available CVSS calculator, and that in the best case, the protocol will require only

---

[102] *"Early Stage" Coordinated Vulnerability Disclosure Template*, National Telecommunications and Information Administration (2016).
[103] *See id.*

a few additional commit and reveal messages as part of the commitment scheme, which can very simply be implemented via the use of a strong hash function. Many implementations of such hash functions exist.[104] More sophisticated and elegant fair exchange solutions would require the use of more advanced software libraries. In the worst case, several runs of the SVC protocol are needed and a dispute needs to be settled via use of a TTP. It is our hope that the Contested Classification process could be as streamlined as possible, and of course, parties are free to examine their 'dispute appetites' and may decide not to engage in the Contested Classification process.

VI. Bug Bounty Programs

It may be argued that bug bounty programs offer an effective market solution to the problem of vulnerability disclosure whereas our solution is of a legislative nature and perhaps unnecessary. A bug bounty program is a scheme by which software vendors offer individual finders financial reward and possibly recognition for reporting software bugs that pertain to security vulnerabilities and the potential exploits of those vulnerabilities. In such a scheme, the disclosure schedule is entirely determined by the vendor, assuming the finder would like to avoid having legal action levied against them. The Facebook bug bounty programs stipulates that the organization will not initiate a lawsuit or law enforcement investigation against a finder if the finder provides Facebook with reasonable time to investigate and mitigate the vulnerability prior to the finder publicly releasing the vulnerability details. We remark that if a finder actively agrees to participate in a bug bounty program, then she submits to the vendor-determined publication deadline and the conditions stated within the terms of such a program. Our solution does not aim to address the case in which bug bounty program terms are invoked - we focus solely on providing a legislative framework for vulnerability disclosure which offers safe harbor to finders and allows for them to have a degree of sway in determining the publication deadline. In other words, if finders agree to forgo participation in the bug bounty program and adhere to our solution, then they are granted safe harbor. Our solution is entirely separate from bug bounty programs, and indeed offers a disclosure framework to researchers when bug bounty programs are not available.

6. CONCLUSION

Neither of the relevant parties, namely the finder and the vendor, appear to be satisfied with the status quo. Vendors are sometimes subject to full disclosure by researchers, leaving them scrambling to effectuate mitigations in a rushed fashion. Their relationship with the security research community remains complex and less friendly than desirable. Researchers fare even worse. There is no clear disclosure mechanism that has achieved industry consensus, they are potentially subject to both civil and criminal liability for their everyday work, and they have absolutely no bargaining power in negotiating disclosure timelines. While it has been acknowledged that the CFAA and DMCA §1201 have been construed overly broadly, there has been little progress in slowing that trend. In order to provide the certainty and legal immunity necessary to remove the chilling effects on security research, a clear default rule is necessary. Our proposed safe harbor, though by no means the final iteration of such a statutory reform, aims to provide more certainty for both researchers and vendors, to create

[104] Many well-known cryptographic libraries implement hash functions, such as OpenSSL, for instance.

more parity in bargaining power, and thereby encourage more security research that will help maintain safety in the technologies used each and every day.

Our safe harbor is designed to take the complex dimensions of vulnerability disclosure into account. Our classification scheme leverages the relative expertise of each party. Generating associated timelines aims to ensure there is a credible threat of disclosure that incentivizes vendors to patch vulnerabilities within reasonable timeframes, and to avoid punishing researchers who are bound by publication and conference deadlines when they give vendors notice with a reasonable time to develop remediations. By operating as a default rule our proposal allows for continuing market convergence on reasonable terms. It allows parties to make private arrangements but provides a default rule that acts as a floor representing the minimal acceptable terms for researchers. As evidenced by all of the examples presented throughout the paper, from Edward Felten and the Sony rootkits to Flavio Garcia and the Megamos case, there is no substitute for a strong security research community. It is crucial that we eliminate chilling effects and enable good faith security research in order to keep users, vendors, and the general public safe.